Mixed Finite Element Methods

Guido Kanschat

July 28, 2017

Preface

These notes compile information from various sources, which should be consulted for more details on particular subjects. These are in particular the books by Daniele Boffi, Franco Brezzi, and Michel Fortin [BFB13] and by Alexandre Ern and Jean-Luc Guermond [EG04] on the general theory of mixed finite elements, the books by Dietrich Braess [Bra97; Bra13] and Philippe Ciarlet [Cia88] on elasticity, the book by Vivette Girault and Pièrre-Arnaud Raviart [GR86] on incompressible fluids, the book by Peter Monk [Mon03], and the works by Douglas Arnold, Richard Falk, and Ragnar Winther [AFW06; AFW10] on the finite element exterior calculus.

Special thanks go to Dr. Daniel Arndt for contributing problems, solutions, and improving the notes at many points. Furthermore, I thank Vladislav Olkhovskiy and Saurabh Mehta for valuable corrections and additions.

Contents

Chapter 1

From elliptic to mixed problems

We begin our course of mixed finite element methods by studying a vector-valued elliptic problem. Then, we study its dependence on its parameters and naturally arrive at a mixed formulation. We derive a few properties of mixed systems and then turn our attention to the first example: the Stokes equations of incompressible flow. We close the chapter by considering mixed systems as first order conditions for constrained minimization problems.

1.1 Linear elasticity

In this section, we study the simplest mathematical model for elastic deformation of solids based on Hooke's law. For comparison, consider [Bra97; Bra13]. For the full nonlinear model in all mathematical detail refer to [Cia88].

1.1.1 Notation: Differential operators for vector fields $u : \mathbb{R}^d \to \mathbb{R}^d$ are defined as follows:

$$\nabla u = \begin{pmatrix} \partial_1 u_1 & \cdots & \partial_d u_1 \\ \vdots & & \vdots \\ \partial_1 u_d & \cdots & \partial_d u_d \end{pmatrix} \qquad \text{(gradient)} \tag{1.1}$$

$$\nabla\cdot u = \sum_{i=1}^{d} \partial_i u_i \qquad \text{(divergence)} \tag{1.2}$$

For a tensor field $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$, the divergence is a vector defined column-wise as

$$\nabla\cdot\sigma = \left(\sum_{i=1}^{d} \partial_i \sigma_{ij}\right)_{j=1,\dots,d} \tag{1.3}$$

1.1.2. The deformation of a solid body is described as a mapping Φ from the **reference configuration** $\Omega \subset \mathbb{R}^d$ to a deformed configuration $\widehat{\Omega} \subset \mathbb{R}^d$, such that each undeformed point $x \in \Omega$ is mapped to the point $\widehat{x}$ after deformation. The domain d is 3 for physically relevant models, but we investigate two-dimensional problems in order to study mathematical properties and numerical methods more easily.

Actually, we are not quite interested in this mapping Φ, since it depends on the position of the points x. On the other hand, a basic principle of physical laws is frame invariance, namely, if we change from one Cartesian coordinate system to another, the physical law may only change by the same coordinate transformation, not in its physical implications. Therefore, only the differences $\widehat{x} - x$ should matter. Thus, we introduce the **displacement** u, such that

$$\Phi = \mathbb{I} + u.$$

The symbol $\mathbb{I}$ will refer to all occurrences of identical mappings and their matrices.

So far, by the introduction of u, we divide translations of the reference configuration out of our model. But, in addition, we have to eliminate the influence of rigid body rotations. These are operations, which leave all distances and angles unchanged. Thus, we investigate the change of the distance between x and $x+z$ under the mapping Φ. By definition of the derivative, we have

$$\begin{aligned} |\Phi(x+z) - \Phi(x)|^2 &= \|\nabla\Phi z\|^2 + o(|z|^4) \\ &= z^T \nabla\Phi^T \nabla\Phi z + o(|z|^4) \\ &= z^T (I + \nabla u^T + \nabla u + \nabla u^T \nabla u) z + o(|z|^4) \\ &= |z|^2 + 2z^T \varepsilon(u) z + o(|z|^4), \end{aligned}$$

where

$$\tilde{\varepsilon}(u) = \tfrac{1}{2}\big(\nabla u^T + \nabla u + \nabla u^T \nabla u\big) \tag{1.4}$$

is the **strain tensor**. From linear algebra, we know that a linear mapping which preserves all distances is orthogonal and thus also preserves angles. Thus, every deformation with $\varepsilon(u) = 0$ is a rigid body transformation, namely a combination of translation and rotation.

In this class we are concerned only with linear problems, which can be justified by the notion of infinitely small deformations u. Then, we only study first order effects in u, which implies that we are going to neglect the quadratic term in $\varepsilon(u)$. This is justified by the fact that we obtain a model, which is sufficiently accurate for small deformations.

1.1.3 Definition: The linearized **strain tensor** or **symmetric gradient** of u is

$$\varepsilon(u) = \frac{\nabla u + \nabla u^T}{2}. \tag{1.5}$$

1.1.4. Next, we have to consider the interplay of forces and deformations. The basic principle is Newton's axiom of force balance. If a body force f acts on a small volume V, there have to be surface forces acting against f in order to keep V at rest. Similarly, if a torque is applied inside this volume, there must be tangential forces on the surface balancing this torque. Due to Euler, we model these forces as a mapping t, which at each point x maps a direction vector n to a force vector $t(x, n)$. Thus, the balance of forces is written as

$$\begin{aligned} \int_V f \,\mathrm{d}x + \int_{\partial V} t(x,n)\,\mathrm{d}s &= 0 \\ \int_V x \times f \,\mathrm{d}x + \int_{\partial V} x \times t(x,n)\,\mathrm{d}s &= 0. \end{aligned}$$

Due to Euler and Cauchy, this mapping $t(x, n)$ can be expressed as $\sigma(x)n$ by the **stress tensor** σ. Without proof, we note that the balance of torque implies that σ is symmetric, while the force balance equation after integration by parts becomes

$$f + \nabla\cdot\sigma = 0. \tag{1.6}$$

What is missing now is a relation between the displacement u and the stress σ, which is not the result of fundamental principles, but of material properties.

Remark 1.1.5. At this point, we play again the card of small deformations, such that we do not have to distinguish whether equations are formulated on the reference domain Ω or on the deformed domain $\widehat{\Omega}$. Such a discussion becomes confusing easily and thus remains a subject for a more specialized class.

1.1.6 Definition: Hooke's law states that the stress depends linearly on the strain. Together with frame invariance, this implies the relation

$$\sigma = 2\mu\varepsilon(u) + \lambda \operatorname{tr} \varepsilon(u)\mathbb{I}, \tag{1.7}$$

where $\lambda \geq 0$ and $\mu > 0$ are material properties called **Lamé-Navier parameters**.

Remark 1.1.7. The trace of the strain operator is equal to the trace of the gradient. Thus, we have

$$\operatorname{tr} \varepsilon(u) = \nabla \cdot u\, \mathbb{I}. \tag{1.8}$$

1.1.8. Equations (1.6) and (1.7) together define a second order partial differential equation, for which we have to impose boundary conditions. A natural choice, which keeps the mathematical analysis simple is the **Dirichlet boundary condition** $u = 0$, corresponding to an elastic body whose boundary is fixed. The alternative is the traction free boundary condition $\sigma n = 0$ with vanishing normal traces. Combinations are possible, for instance $u \cdot n = 0$ for a boundary that allows sliding but no penetration. Constraining ourselves to Dirichlet condition on $\Gamma_D \subset \partial\Omega$ and traction free on $\gamma_N = \partial\Omega \setminus \Gamma_D$, we obtain the boundary value problem

$$\begin{aligned} -\nabla \cdot \sigma(x) &= f(x) && x \in \Omega, \\ u(x) &= 0 && x \in \Gamma_D, \\ \sigma(x)n &= 0 && x \in \Gamma_N, \end{aligned} \tag{1.9}$$

together with the material law (1.7). Once we test and integrate by parts to obtain our weak formulation, we obtain

$$\int_\Omega -(\nabla \cdot \sigma) \cdot v \,\mathrm{d}x = \int_\Omega \sigma : \nabla v \,\mathrm{d}x - \int_{\Gamma_N} \sigma n \cdot v \,\mathrm{d}s,$$

such that traction free is actually the natural boundary condition comparable to the Neumann condition for the Laplacian. Note that : is the double contraction or Frobenius product (see Problem 1.1.11 below) of the two tensors.

1.1.9. We now walk the missing steps to obtain a weak formulation. first, we enter Hooke's law for σ to obtain:

$$\int_\Omega \bigl[2\mu\varepsilon(u) : \nabla v + \lambda(\nabla \cdot u\mathbb{I}) : \nabla v\bigr] \,\mathrm{d}x = \int_\Omega f \cdot v \,\mathrm{d}x.$$

Then, we choose the space

$$V = H^1_{\Gamma_D}(\Omega; \mathbb{R}^d) = \bigl\{ v \in H^1(\Omega; \mathbb{R}^d) \big| v_{|\Gamma_D} = 0 \bigr\}. \tag{1.10}$$

We notice for the second term that

$$\mathbb{I} : \nabla v = \sum_{i=1}^{d} \partial_i v_i = \nabla\cdot v.$$

Furthermore, we use the result of Problem 1.1.11 to obtain

$$\varepsilon(u) : \nabla v = \varepsilon(u) : \varepsilon(v).$$

1.1.10 Definition: The weak formulation of the Lamé-Navier boundary value problem

$$\begin{aligned} -\nabla\cdot\sigma(x) &= f(x) \quad && x \in \Omega, \\ u(x) &= 0 && x \in \Gamma_D, \\ \sigma(x)n &= 0 && x \in \Gamma_N, \end{aligned}$$

is: find $u \in V = H^1_{\Gamma_D}(\Omega; \mathbb{R}^d)$ such that

$$a(u, v) \equiv 2\mu(\varepsilon(u), \varepsilon(v)) + \lambda(\nabla\cdot u, \nabla\cdot v) = (f, v) \quad \forall v \in V. \tag{1.11}$$

1.1.11 Problem: Given the vector space of square matrices $X = \mathbb{R}^{d\times d}$ with the Frobenius inner product

$$\langle A, B\rangle = A : B = \sum_{ij} a_{ij} b_{ij}. \tag{1.12}$$

Show that the subspaces of symmetric and skew-symmetric matrices, respectively, are orthogonal to each other and X is the direct sum of those.

1.1.12. The form $a(\cdot, \cdot)$ is symmetric and thus semi-definite on V. It can also be bounded easily by the H^1-norm. But, for well-posedness of the weak formulation, we also require ellipticity. This question is indeed not trivial and rests on the fact that for a function $u \in V$, such that ∇u is skew-symmetric everywhere, there holds $\varepsilon(u) \equiv 0$. Thus, such functions must be excluded by the boundary conditions. Note, that in particular for rigid body translations and rotations $\varepsilon(u) = 0$. Therefore, the Dirichlet boundary conditions must exclude such solutions.

The condition needed for well-posedness is called Korn inequality, and it will be posed as an assumption. We will give a proof for a simple case and refer the readers to a plethora of articles on more complicated cases.

1.1.13 Assumption: We assume that the boundary conditions defining the space V in the weak formulation of the Lamé-Navier equations are such that a **Korn inequality**

$$c_K^2 \|u\|^2_{H^1(\Omega;\mathbb{R}^d)} \le \|u\|^2_{L^2(\Omega;\mathbb{R}^d)} + \|\varepsilon(u)\|^2_{L^2(\Omega;\mathbb{R}^d)}, \tag{1.13}$$

holds for all $u \in V$ with a uniform constant $c_K > 0$.

Remark 1.1.14. We use the indefinite article 'a' for this inequality, because it comes in many different forms, for instance replacing the L^2-norm by a boundary term. It's mathematics is not trivial. There is nevertheless a very simple interpretation of this inequality: the kernel of the gradient consists of constant functions, in the context of elastic deformations of translations by a constant vector. The kernel of the strain tensor contains translation *and* rotations. Thus, a Korn inequality can only hold, if the boundary conditions rule out rotations.

1.1.15 Lemma: Let $V = H^1_0(\Omega;\mathbb{R}^d)$ for a Lipschitz domain Ω. Then, Korn's inequality holds on V with a constant $c_K > 0$.

Proof. While the inequality is of high importance in the mathematical and numerical analysis of problems in continuum mechanics, it a peripheral topic to this class. Therefore, we omit the proof and refer to Theorem 3.3 in [DL76, Section III.3.3]. □

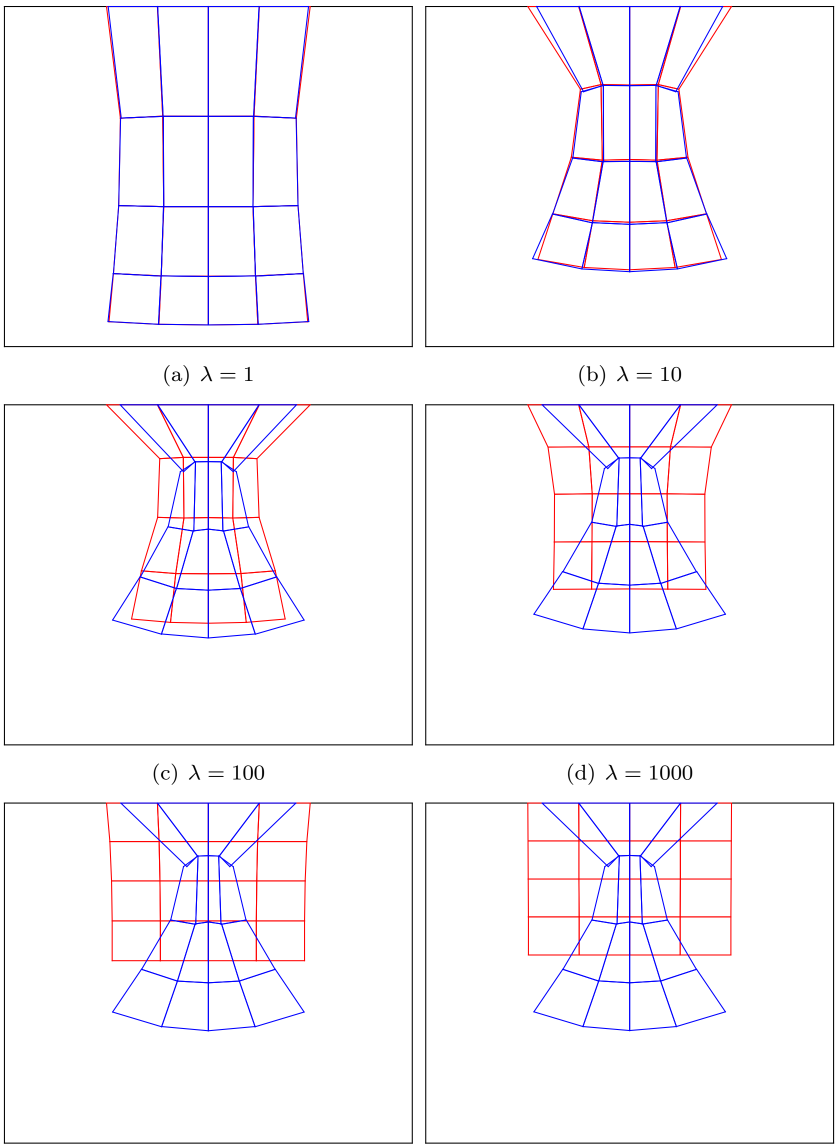
1.1.16 Problem: Let the space $V = H^1_0(\Omega;\mathbb{R}^d)$ be equipped with the inner product $\langle u, v\rangle = a(u, v)$ with the bilinear form of the Lamé-Navier equations and the corresponding norm $\|.\|_V$.
Show using techniques from the standard theory of elliptic partial differential equations:

1. The weak formulation has a unique solution for which there holds
$$\|u\|_V \le \|f\|_{V^*}.$$
2. The "energy estimate" for conforming finite element approximation with a space $V_h \subset V$
$$\|u - u_h\|_V = \inf_{v_h\in V_h} \|u - v_h\|_V.$$
3. The H^1-error estimate
$$\|u - u_h\|_{H^1} \le \frac{2\mu + d\lambda}{2c_K\mu} \inf_{v_h\in V_h} \|u - v_h\|_{H^1}. \tag{1.14}$$
Use the fact that the space V can be composed into the space V^0 of divergence-free functions ($\nabla\cdot v = 0$) and its complement.
4. For $\lambda \gg \mu$, the previous estimate is useless. Can it be improved easily? In view of the "energy estimate", can you think of conditions?

Example 1.1.17. We study the finite element approximation of the following problem: a square sheet of elastic material is hanging from the top, subject to gravity acting as body force pointing downward. We choose $\mu = 1$ and vary λ from 1 to 10^5. Figure 1.1 shows approximations with standard bilinear finite elements (red) and a "good" approximation (blue). In fact, the red solution for $\lambda = 10^5$ is almost identical with the undeformed configuration, although the material is not as hard. This phenomenon was discovered early in finite element history and is called "locking".

1.1.18. As we could see in the preceding problem and example, approximation of the solution to the Lamé-Navier equations becomes difficult, if $\lambda \gg \mu$. In this case, the material is called almost incompressible, since the divergence measures compression or dilation and the dominating divergence term forces the divergence of the solution to be small. These cases are important in engineering and they initiated a lot of the research that resulted in the topics of this class.

1.1.19. A way to approach this problem is the introduction of an auxiliary variable
$$p = -\lambda\nabla\cdot u.$$



(a) $\lambda = 1$ (b) $\lambda = 10$

(c) $\lambda = 100$ (d) $\lambda = 1000$

(e) $\lambda = 10000$ (f) $\lambda = 100000$

Figure 1.1: Approximation with standard finite elements (red) and “good” elements (blue) for $\mu = 1$ and different values of λ.

Entering this definition into the Lamé-Navier equations, we obtain the following weak formulation.

1.1.20 Definition: The **displacement-pressure formulation** of the Lamé-Navier equations reads: find a pair $(u,p) \in V \times Q$ such that

$$\begin{aligned} 2\mu(\varepsilon(u),\varepsilon(v)) - (p,\nabla\cdot v) &= (f,v) \quad && \forall v \in V \\ -(q,\nabla\cdot u) - \tfrac{1}{\lambda}(p,q) &= 0 && \forall q \in Q. \end{aligned} \tag{1.15}$$

Equivalently, we write this in a single equation as

$$2\mu(\varepsilon(u),\varepsilon(v)) - (p,\nabla\cdot v) - (q,\nabla\cdot u) - \tfrac{1}{\lambda}(p,q) = (f,v) \qquad \forall v \in V, q \in Q, \tag{1.16}$$

or in the nonsymmetric, (semi-)definite version

$$2\mu(\varepsilon(u),\varepsilon(v)) + (p,\nabla\cdot v) - (q,\nabla\cdot u) + \tfrac{1}{\lambda}(p,q) = (f,v) \qquad \forall v \in V, q \in Q, \tag{1.17}$$

Here, V is as before and $Q = L^2(\Omega)$.

Remark 1.1.21. The three forms have different purposes and will be used accordingly. The first one highlights the fact that we now have a system of equations, each equation tested with its own test function. The second and third stress the fact that we now have a bilinear form on the product space $X = V \times Q$.

The second form is symmetric, but we will see later that is indefinite. Thus, non of our tools from functional analysis apply. In contrast, the third form is nonsymmetric, but we have

$$\begin{aligned} &2\mu(\varepsilon(u),\varepsilon(u)) + (p,\nabla\cdot u) - (p,\nabla\cdot u) + \tfrac{1}{\lambda}(p,p) \\ &\qquad = 2\mu(\varepsilon(u),\varepsilon(u)) + \tfrac{1}{\lambda}(p,p) \ge \tfrac{2\mu}{c_K}\|u\|_{H^1}^2 + \tfrac{1}{\lambda}(p,p). \end{aligned}$$

Thus, we have ellipticity with respect to the norm

$$\|(u,p)\|_X^2 = \|u\|_{H^1}^2 + \|p\|_{L^2}^2. \tag{1.18}$$

Nevertheless, the ellipticity constant depends on λ, and for large λ, we loose sharpness of estimates again.

1.1.22 Definition: Integrating the first equation by parts, we obtain the **strong form** of the Lamé-Navier equations

$$\begin{aligned} -2\mu\nabla\cdot\varepsilon(u) + \nabla p &= f \\ \nabla\cdot u \quad + \tfrac{1}{\lambda}p &= 0 \end{aligned} \tag{1.19}$$

1.2 Abstract saddle-point systems

1.2.1. In order to study the mathematics not only of the Lamé-Navier equations but of more general systems of the form of equation (1.15), we introduce abstract bilinear forms

$$\begin{aligned} a(u,v) &= 2\mu(\varepsilon(u),\varepsilon(u)), && u,v\in V\\ b(u,p) &= -(p,\nabla\cdot u), && u\in V, p\in Q\\ c(p,q) &= \tfrac{1}{\lambda}(p,q) && p,q\in Q. \end{aligned} \tag{1.20}$$

1.2.2 Definition: For each of the bilinear forms, we define associated operators

$$\begin{aligned} A&\colon V\to V^* & \langle Au,v\rangle &= a(u,v),\\ B&\colon V\to Q^* & \langle Bu,q\rangle &= b(u,q),\\ B^T&\colon Q\to V^* & \left\langle B^Tp,v\right\rangle &= b(v,p),\\ C&\colon Q\to Q^* & \langle Cp,q\rangle &= c(p,q). \end{aligned} \tag{1.21}$$

Here, $\langle\cdot,\cdot\rangle$ is the canonical pairing between an element of the dual space and the space itself.

1.2.3 Definition: The abstract saddle-point problem in weak form reads: find a pair $(u,p)\in V\times Q$ such that

$$\begin{aligned} a(u,v)+b(v,p) &= f(v) && \forall v\in V,\\ b(u,q)-c(p,q) &= g(q) && \forall q\in Q. \end{aligned} \tag{1.22}$$

In operator notation, it reads

$$\begin{aligned} Au + B^Tp &= f && \text{in } V^*,\\ Bu - \ Cp &= g && \text{in } Q^*. \end{aligned} \tag{1.23}$$

1.2.4 Notation: In order to consider the whole bilinear form of the saddle-point system on the space $X = V\times Q$, we introduce

$$\mathcal{A}\left(\begin{pmatrix}u\\p\end{pmatrix},\begin{pmatrix}v\\q\end{pmatrix}\right) = a(u,v)+b(v,p)+b(u,q)-c(p,q) \tag{1.24}$$

1.2.5 Definition: Let the operator $A : V \to V^*$ in the saddle-point system be invertible. Then, we define the **Schur complement** operator $S\colon Q \to Q^*$ of the system as

$$S = -BA^{-1}B^T - C. \tag{1.25}$$

1.2.6 Lemma: Formally, the saddle-point system (1.23) can be solved in two steps by solving

$$Sp = g - BA^{-1}f, \tag{1.26}$$

$$Au = f - B^T p. \tag{1.27}$$

Proof. Formally solve the first equation of (1.23) for u and enter into the second. □

1.2.7 Lemma: Let $a(\cdot,\cdot)$ be elliptic and $c(\cdot,\cdot)$ be positive semi-definite, and let $\ker(B)^T \cap \ker(C) \neq \{0\}$. Then, the bilinear form $\mathcal{A}(\cdot,\cdot)$ is indefinite.

Proof. First, we note that because of ellipticity of $a(\cdot,\cdot)$

$$\mathcal{A}\left(\begin{pmatrix}u\\0\end{pmatrix},\begin{pmatrix}u\\0\end{pmatrix}\right) = a(u,u) \geq \gamma\|u\|_V^2,$$

for some positive constant γ. Furthermore, A is invertible and its inverse is positive definite. Furthermore, $A^{-1}B^T\colon Q \to V$. Then, choosing $v = -A^{-1}B^T p$ yields

$$\begin{aligned}\mathcal{A}\left(\begin{pmatrix}v\\p\end{pmatrix},\begin{pmatrix}v\\p\end{pmatrix}\right) &= a(A^{-1}B^T p, A^{-1}B^T p) - 2b(A^{-1}B^T p, p) - c(p,p)\\ &= \langle B^T p, A^{-1}B^T p\rangle - 2\langle A^{-1}B^T p, B^T p\rangle - c(p,p)\\ &= -\langle B^T p, A^{-1}B^T p\rangle - c(p,p).\end{aligned}$$

The first term is a quadratic term with the bilinear form associated with A^{-T}, which is positive definite. Since $b(\cdot,\cdot)$ is not the zero form, there is some p such that $v \neq 0$. Therefore, with the minus sign and the semi-definiteness of $c(\cdot,\cdot)$ we have found a vector such that

$$\mathcal{A}\left(\begin{pmatrix}v\\p\end{pmatrix},\begin{pmatrix}v\\p\end{pmatrix}\right) < 0.$$

□

Remark 1.2.8. The previous result holds in particular for $c(\cdot,\cdot) \equiv 0$.

1.3 Stokes equations

1.3.1. When we write the Lamé-Navier equations in mixed form according to Definition 1.1.20, there is no parameter λ tending to infinity when the material becomes more and more compressible. Instead, there is the parameter $1/\lambda$ tending to zero. Thus, we can simply consider the case of incompressible material by setting $1/\lambda = 0$ or, in our abstract framework (1.20) setting $c(\cdot,\cdot) = 0$. The resulting system is not only important for incompressible elasticity, but also models the slow flow of a very viscous liquid, so called creeping flow.

1.3.2 Definition: The **Stokes equations** in strong form are

$$\begin{aligned} -2\mu\nabla\cdot\varepsilon(u) + \nabla p &= f \\ \nabla\cdot u \qquad\qquad &= 0. \end{aligned} \tag{1.28}$$

In weak form, they are: find $u \in V \subset H^1(\Omega;\mathbb{R}^d)$ and $p \in Q \subset L^2(\Omega)$ such that

$$\begin{aligned} 2\mu(\varepsilon(u),\varepsilon(v)) - (\nabla\cdot v, p) &= (f,v) + \text{bdry} && \forall v \in V \\ -(\nabla\cdot u, q) &= 0 + \text{bdry} && \forall q \in Q. \end{aligned} \tag{1.29}$$

The subspaces V and Q are determined by boundary conditions.

1.3.3 Definition: A vector-valued function u is called **divergence-free** or **solenoidal**, if there holds

$$\nabla\cdot u = 0.$$

Flow described by a solenoidal function is called **incompressible**.

1.3.4 Lemma: Let $V = H^1_0(\Omega;\mathbb{R}^d)$. Then, for any solenoidal function $u \in V$ there holds

$$2\mu(\varepsilon(u),\varepsilon(v)) = \mu(\nabla u, \nabla v) \qquad \forall v \in V. \tag{1.30}$$

Proof. We have

$$\begin{aligned} \varepsilon(u) : \varepsilon(v) &= \frac{1}{4}\sum_{i,j=1}^{d} \left[(\partial_i u_j + \partial_j u_i)(\partial_i v_j + \partial_j v_i)\right] \\ &= \frac{1}{2}\sum_{i,j=1}^{d} \left[\partial_i u_j \partial_i v_j + \partial_i u_j \partial_j v_i\right]. \end{aligned}$$

The first term is the desired result, thus we have to eliminate the other one. We integrate by parts to obtain

$$\int_\Omega \partial_i u_j \partial_j v_i \,\mathrm{d}x = -\int_\Omega \partial_{ij} u_j v_i \,\mathrm{d}x = \int_\Omega \partial_j u_j \partial_i v_i \,\mathrm{d}x$$

Entering in the previous equation and summing over i and j, we obtain

$$2(\varepsilon(u), \varepsilon(v)) = (\nabla u, \nabla v) + (\nabla\cdot u, \nabla\cdot v) = (\nabla u, \nabla v).$$

□

1.3.5. In order to simplify subsequent discussion, it is customary to use the previous lemma to simplify the Stokes equations and to replace the strain tensor by the gradient. As a result, we can avoid the use of a Korn inequality and operate directly with the inner product of H^1. We note though that this formulation, while mathematically simpler, is physically wrong if $u \neq H^1_0(\Omega;\mathbb{R}^d)$.

1.3.6 Definition: The simplified **Stokes equations** in strong form are

$$\begin{aligned} -\nu\Delta u + \nabla p &= f \\ \nabla\cdot u &= 0. \end{aligned} \tag{1.31}$$

In weak form, they are: find $u \in V \subset H^1(\Omega;\mathbb{R}^d)$ and $p \in Q \subset L^2(\Omega)$ such that

$$\begin{aligned} \nu(\nabla u, \nabla v) - (\nabla\cdot v, p) &= (f, v) + \text{bdry} && \forall v \in V \\ -(\nabla\cdot u, q) &= 0 + \text{bdry} && \forall q \in Q. \end{aligned} \tag{1.32}$$

The subspaces V and Q are determined by boundary conditions.

1.3.7 Definition: Typical boundary conditions for the Stokes problem are

$$\text{no-slip:} \quad u = 0, \tag{1.33}$$

$$\text{free:} \quad \partial_n u + pn = 0, \tag{1.34}$$

$$\text{slip:} \quad u_n = 0 \qquad \partial_n u_\tau = 0, \tag{1.35}$$

$$\text{friction:} \quad u_n = 0 \qquad \partial_n u_\tau = \alpha u_\tau. \tag{1.36}$$

Here, u_n and u_τ are the normal and tangential components of u at the boundary.

Remark 1.3.8. Very much the same way as for elliptic equations, boundary conditions on the function u itself are essential boundary conditions which have to be incorporated into the space V, while those on the normal derivatives are the result of integration by parts and thus of the type of natural boundary conditions.

All of the conditions above can also be imposed with nonzero data, where the physical meaning of such a condition might be debatable in some cases. Mathematically, inhomogeneous essential boundary conditions are achieved by lifting an arbitrary function with this boundary condition and modifying the right hand side of the equation, while inhomogeneous conditions on the normal derivative are implemented by boundary integrals on the right hand side.

Remark 1.3.9. Physically, the condition $u_n = 0$ models an impermeable wall. It says in particular, that no mass is lost through this boundary and is thus related to the first principle of mass conservation.

The conditions on the tangential velocity model the fact that molecules very close to the wall stick to the wall. While this claim is not supported by a first principle, it has been verified by measurements to very high accuracy. Nevertheless, in the study of turbulent flow, a friction condition comes up quite naturally.

1.3.10 Lemma: For any solenoidal u function there holds

$$\int_{\partial\Omega} u \cdot n \, \mathrm{d}s = 0. \tag{1.37}$$

Furthermore, if the boundary condition $u_n = 0$ holds on the whole boundary $\partial\Omega$, then the pressure p is determined by the Stokes equations only up to a constant.

Proof. The first statement is simple application of the Gauss theorem

$$\int_\Omega \nabla \cdot u \, \mathrm{d}x = \int_{\partial\Omega} u \cdot n \, \mathrm{d}s.$$

For the second statement, we note that the only term in the equations which determines the pressure is $(\nabla \cdot v, p)$. Integrating by parts and using the boundary condition, we obtain

$$\int_\Omega \nabla \cdot vp \, \mathrm{d}x = -\int_\Omega v \nabla p \, \mathrm{d}x + \int_{\partial\Omega} v \cdot np \, \mathrm{d}s = -\int_\Omega v \nabla p \, \mathrm{d}x.$$

Since the gradient of a constant is zero, we can add any constant function to a given solution p without changing the term $(\nabla \cdot v, p)$, thus leaving p determined only up to a constant. □

1.3.11 Notation: If in Definition 1.3.6 the space V is chosen such that for all $v \in V$ there holds $v \cdot n = 0$ on the whole boundary $\partial\Omega$, then the pressure solution p cannot be determined uniquely in $L^2(\Omega)$. In such cases, we choose the pressure space

$$L^2_0(\Omega) = L^2(\Omega)/\mathbb{R} = \left\{ q \in L^2(\Omega) \middle| \int_\Omega q \,\mathrm{d}x = 0 \right\}. \tag{1.38}$$

Remark 1.3.12. As we could see from the preceding lemma, solvability of equation (1.32) depends on some compatibility of the spaces V and Q we had not seen in the elliptic case. Indeed, we need a whole new tool from functional analysis to replace the Lax-Milgram lemma. This tool will be studied in Chapter 2.

1.4 Relation to constrained minimization

1.4.1. We end our introduction by relating the Stokes equations to a constrained minimization problem very much like we had considered the solution of the Poisson equation as a minimization problem on the space V.

1.4.2 Theorem: Assume that $a(\cdot,\cdot)$ is symmetric and V-elliptic on the Hilbert space V. Let

$$J(v) = \tfrac{1}{2} a(v,v) - f(v). \tag{1.39}$$

Then, the minimization problem finding $u \in V$ such that

$$J(u) = \inf_{v\in V} J(v), \tag{1.40}$$

has a unique solution, which then is a minimum. It is determined by the first order necessary condition

$$a(u,v) = f(v) \quad \forall v \in V.$$

1.4.3. In the next step, we constrain the solution u to a subspace of V defined as the kernel of the linear operator $B : V \to Q^*$. Hence, we consider the minimization problem

$$J(u) = \min J(v) \quad \text{subject to} \quad Bu = 0.$$

Going back to formulations with bilinear forms, the constraint translates to

$$u \in \ker(B) = \bigl\{ v \in V \bigm| b(v,q) = 0 \quad \forall q \in Q \bigr\}. \tag{1.41}$$

Since $\ker(B)$ is a vector space, we can just consider the restriction of the minimization problem to this space. This is called the reduced problem.

1.4.4 Definition: The **reduced problem** of the constrained minimization problem above is: find $u \in \ker(B)$, such that

$$J(u) = \inf_{v \in \ker(B)} J(v). \tag{1.42}$$

1.4.5 Lemma: If under the assumptions of Theorem 1.4.2 there holds in addition that $\ker(B)$ is a closed subspace of V, then the reduced problem in Definition 1.4.4 has a unique solution.

Proof. A closed subspace of a Hilbert space is a Hilbert space itself. Then, the V-ellipticity of $a(\cdot,\cdot)$ is inherited on $\ker(B)$, and thus the Lax-Milgram lemma provides a unique solution for the first order necessary condition on $\ker(B)$. □

Remark 1.4.6. Here we already note that V-ellipticity of $a(\cdot,\cdot)$ is sufficient, but not necessary. Indeed, ellipticity on $\ker(B)$ would have been sufficient for well-posedness of the reduced problem.

1.4.7. While the solution theory for the reduced problem is particularly simple and purely elliptic, the actual solution requires a representation of functions in $\ker(B)$, for instance a basis of these functions. In practice, this is often inconvenient, and we seek a method that solves a problem on the whole space.

1.4.8 Theorem: If $u \in V$ is a solution of the constrained minimization problem

$$J(u) = \min_{v \in V} J(v), \qquad b(u,q) = 0 \quad \forall q \in Q,$$

the pair $(u,p) \in V \times Q$ is a stationary point of the **Lagrange functional**

$$\mathscr{L}(v,q) = \tfrac{1}{2} a(v,v) - f(v) + b(v,q). \tag{1.43}$$

Here, $p \in Q$ is called the **Lagrange multiplier**.

1.4.9 Problem: Verify: the first order necessary conditions of the Lagrange multiplier rule are

$$\begin{aligned} a(u,v)+b(v,p) &= f(v) \qquad \forall v\in V,\\ b(u,q) &= 0 \qquad \forall q\in Q. \end{aligned}$$

To this end, recall the type of objects that derivatives of linear functionals are and compute the derivatives of $\mathscr{L}$.

1.4.10 Corollary: The Stokes equations with no-slip boundary conditiond are the first order necessary conditions for the constrained minimization problem: find a velocity $u \in V = H^1_0(\Omega)$ such that

$$\begin{aligned} \tfrac{1}{2}(\nabla u, \nabla u) - f(u) &= \min,\\ (\nabla\cdot u, q) &= 0 \qquad \forall q\in Q. \end{aligned}$$

The pressure p assumes the role of a Lagrange multiplier.

Chapter 2

Conditions for well-posedness

In this chapter, we will first modify conditions for well-posedness in finite dimensions from positive definiteness to the general case. In particular, we will derive a quantitative formulation, which we will study for infinite dimensional problems in the second section. In the third section, we derive the inf-sup condition for mixed problems as a special case.

2.1 Finite-dimensional problems

2.1.1. So far, our power horse for well-posedness was the Lax-Milgram lemma, which can be applied under the conditions

$$a(u,v) \le M\|u\|\|v\| \qquad \forall u,v \in V \tag{2.1}$$

$$a(u,u) \ge \gamma\|u\|^2 \qquad \forall u \in V. \tag{2.2}$$

The second condition can also be rewritten in terms of the Rayleigh quotient as

$$0 < \gamma = \inf_{u\in V} \frac{a(u,u)}{\|u\|^2}.$$

Restricting this to a finite dimensional space, the notation usually changes from

$$a(u,v) = f(v) \qquad \text{to} \qquad v^T A u = v^T f, \tag{2.3}$$

where $A \in \mathbb{R}^{n\times n}$ is the matrix associated with the bilinear form. The bound for the Rayleigh quotient means nothing but that the real parts of all eigenvalues of A are bounded from below by γ. Thus, a matrix A for which we can apply the Lax-Milgram lemma is positive definite. And the statement of the lemma in finite dimension is, that a positive definite matrix is invertible. We know from linear algebra that this is true, but we also know that the condition is all but necessary.

2.1.2. Why did we replace this clear theorem by the weaker Lax-Milgram lemma, when we studied elliptic partial differential equations? For the first condition, it should be noted that spectral properties of operators between spaces of infinite dimension are much harder to obtain. Further, we do not need information on the whole spectrum, but only on the eigenvalue closest to zero. Therefore, we used a simple estimate in order to avoid discussing the spectrum at all. But, there is an important difference between Theorem 2.1.5 and the estimate (2.2): the assumption of the theorem is qualitative, $\lambda \neq 0$, while the assumption of Lax-Milgram is quantitative,

$$\Re\lambda \geq \gamma > 0.$$

The following problem shows why such a change is necessary.

2.1.3 Problem: On the space $\ell_2(\mathbb{R})$ define the operator A by its eigenvalue decomposition

$$\begin{aligned} A : \ell_2(\mathbb{R}) &\to \ell_2(\mathbb{R}) \\ e_k &\mapsto \tfrac{1}{k} e_k. \end{aligned}$$

Here, $\{e_k\}$ is the orthogonal basis of unit vectors of the form

$$e_k = (0\,, \ldots, 0\,, \underset{\substack{\uparrow \\ k}}{1}\,, 0\,, \ldots)^T.$$

1. Show that this operator does not have a bounded inverse, albeit its eigenvalues are positive.
2. Show that the range of A is not closed in $\ell_2(\mathbb{R})$

2.1.4 Problem: Find an invertible, symmetric matrix $A \in \mathbb{R}^{2\times 2}$ and a vector $v \in \mathbb{R}^2$ such that $v^T A v = 0$ and thus the Lax-Milgram lemma is inconclusive.

The question of well-posedness in finite dimensions can be answered by:

2.1.5 Theorem: A matrix $A \in \mathbb{R}^{n\times n}$ is invertible if and only if one of the following equivalent conditions holds:

1. all its (possibly complex) eigenvalues are nonzero,
2. all its singular values are nonzero,
3. for each nonzero $v \in \mathbb{R}^n$ holds $Av \neq 0$.

2.1.6. We focus on the second and third conditions, respectively, in Theorem 2.1.5. But, the problem above tells us that we will run into trouble, if we do not quantify this. Therefore, we start our attempt by requiring:

$$\|Au\|^2 \geq \gamma \|u\|^2 \qquad \forall u \in V.$$

But while this is a condition we can easily write down for matrices and operators, it does not work that well for bilinear forms. Thus, we first look at the singular value decomposition.

2.1.7 Theorem: Let $A \in \mathbb{R}^{m\times n}$ be a real matrix. Then, there exist two orthogonal matrices $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ as well as a real, nonnegative diagonal matrix $\hat{\Sigma}$, such that

$$A = U\Sigma V^T, \qquad \text{and } \Sigma = \begin{cases} \begin{bmatrix} \hat{\Sigma} & 0 \end{bmatrix} & \text{for } m < n \\ \hat{\Sigma} & \text{for } m = n \\ \begin{bmatrix} \hat{\Sigma} \\ 0 \end{bmatrix} & \text{for } m > n \end{cases} \tag{2.4}$$

This is the **singular value decomposition** (**SVD**) of A, the diagonal entries of $\hat{\Sigma}$ are the **singular values** of A and the column vectors of U and V are the left and right **singular vectors** of A, respectively. The same theorem holds for complex matrices with unitary U and V.

Proof. We prove constructively for the real case by induction. For $m = 1$ or $n = 1$ the theorem is obvious. Let now $m, n > 1$ and assume the theorem has been proven for $A \in \mathbb{R}^{(m-1)\times(n-1)}$. Let σ_1^2 be the largest eigenvalue of $A^T A$, which due to the symmetry of $A^T A$ is real and nonnegative. Actually, if it is zero, then $A^T A = 0$ and thus $A = 0$ and $\Sigma = 0$. Now assume $\sigma_1^2 > 0$. Choose x_1 as an eigenvector to the eigenvalue σ_1^2 of $A^T A$ and

$$y_1 = \frac{1}{\sigma_1} A x_1.$$

We can complete both x_1 and y_1 to an orthonormal basis X and Y, respectively. Then, there holds for $e_1 \in \mathbb{R}^n$ and $\bar{e}_1 \in \mathbb{R}^m$:

$$Y^T A X e_1 = Y^T A x_1 = \sigma_1 Y^T y_1 = \sigma_1 \bar{e}_1$$
$$(Y^T A X)^T \bar{e}_1 = X^T A^T Y \bar{e}_1, = X^T A^T y_1 = \tfrac{1}{\sigma_1} X^T A^T A x_1 = \sigma_1 X^T x_1 = \sigma_1 e_1.$$

Thus,

$$Y^T A X = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \tilde{A} \end{bmatrix}$$

with $\tilde{A} \in \mathbb{R}^{(m-1)\times(n-1)}$. By induction, there holds $\tilde{A} = \tilde{U}\tilde{\Sigma}\tilde{V}^T$ with orthogonal matrices U and V and $\tilde{\Sigma}$ of the same form as Σ and both dimensions reduced by 1. Now let

$$U = Y \begin{bmatrix} 1 & 0 \\ 0 & \tilde{U} \end{bmatrix}, \qquad V = X \begin{bmatrix} 1 & 0 \\ 0 & \tilde{V} \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \tilde{\Sigma} \end{bmatrix}$$

U and V as the product of orthogonal matrices are orthogonal, σ has the claimed structure and there holds $A = U\Sigma V^T$. □

2.1.8 Corollary: By the construction of the proof, the singular values are ordered by decreasing magnitude,

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0. \tag{2.5}$$

The number r of nonzero singular values is the dimension of the range of the matrix A.

2.1.9 Definition: Let $A : V \to W$ be a linear operator. Then, we define the **kernel** and the **range** of A as

$$\begin{aligned} \ker(A) &= \{v \in V \,|\, Av = 0\} \\ \operatorname{im}(A) &= \{w \in W \,|\, \exists v \in V : Av = w\}. \end{aligned}$$

2.1.10 Definition: Let $V \subset \mathbb{R}^n$ be a subspace. We define the **orthogonal complement** of V as

$$V^\perp = \{w \in \mathbb{R}^n \,|\, \forall v \in V \langle w, v\rangle = 0\}. \tag{2.6}$$

2.1.11 Lemma: Let $A \in \mathbb{R}^{m\times n}$ and A^T its transpose. Then, there holds

$$\begin{aligned} \ker(A) &= \operatorname{im}(A^T)^\perp \\ \operatorname{im}(A) &= \ker\left(A^T\right)^\perp \\ \ker\left(A^T\right) &= \operatorname{im}(A)^\perp \\ \operatorname{im}(A^T) &= \ker(A)^\perp \end{aligned} \tag{2.7}$$

Proof. Let $A = U\Sigma V^T$ be the singular value decomposition of A and r be the number of nonzero singular values. Then, the first r vectors of U span the range of A and the last $n-r$ vectors of V span its kernel. Furthermore,

$$A^T = \left(U\Sigma V^T\right)^T = V\Sigma U^T. \tag{2.8}$$

Therefore, the first r vectors of V span the range of A^T and the last $n-r$ vectors of U span its kernel. ☐

2.1.12 Corollary: Let $A \in \mathbb{R}^{m\times n}$ and A^T its transpose. Then, the restrictions $A\colon \ker(A)^\perp \to \operatorname{im}(A)$ and $A^T\colon \ker\left(A^T\right)^\perp \to \operatorname{im}(A^T)$ are isomorphisms.

Proof. We note that $\dim\operatorname{im}(A) = \dim\operatorname{im}(A^T)$. Thus, by Lemma 2.1.11 the dimensions of domain and range of each of the restricted operators are equal, say $\dim\operatorname{im}(A) = r$. The singular value decomposition of the operators is

$$A = U\Sigma V^T \qquad A^T = V\Sigma U^T,$$

where all matrices are in $\mathbb{R}^{r\times r}$ and

$$\Sigma = \operatorname{diag}(\sigma_1, \dots, \sigma_r),$$

and all singular values are positive. Thus, A and A^T are invertible. ☐

2.1.13 Corollary: Let $r = \dim\ker(A)^\perp$. Then, for the smallest nonzero singular value there holds

$$\sigma_r = \inf_{v\in\ker(A)^\perp}\sup_{w\in\mathbb{R}^m}\frac{w^TAv}{|v||w|} = \inf_{w\in\ker(A^T)^\perp}\sup_{v\in\mathbb{R}^n}\frac{w^TAv}{|v||w|}. \tag{2.9}$$

Proof. Since the Cauchy-Schwarz inequality turns into an equation if and only if the two vectors are coaligned, there holds for any $v \in \mathbb{R}^n$:

$$\sup_{w\in\mathbb{R}^m}\frac{w^TAv}{|w|} = \frac{v^TA^TAv}{|Av|}.$$

Therefore,

$$\inf_{v\in\ker(A)^\perp}\sup_{w\in\mathbb{R}^m}\frac{w^TAv}{|v||w|} = \inf_{v\in\ker(A)^\perp}\frac{|Av|}{|v|}.$$

Now, let $v = \sum \alpha_i v_i$ where v_i are the columns of V in the SVD of A. Then,

$$|Av|^2 = \left| A \sum_{i=1}^r \alpha_i v_i \right|^2 = \left| \sum_{i=1}^r \sigma_i \alpha_i u_i \right|^2 = \sum_{i=1}^r \sigma_i^2 \alpha_i^2.$$

The quotient

$$\frac{\sum_{i=1}^r \sigma_i^2 \alpha_i^2}{\sum_{i=1}^r \alpha_i^2}$$

clearly has its minimum if $\alpha_1 = \cdots = \alpha_{r-1} = 0$. □

2.1.14 Definition: A bilinear form $a(\cdot,\cdot)$ on $V \times W$ is said to admit the **inf-sup condition** or is called **inf-sup stable**, if there holds

$$\inf_{u\in V} \sup_{w\in W} \frac{a(u,w)}{\|u\|_V \|w\|_W} \geq \gamma > 0. \tag{2.10}$$

Remark 2.1.15. In this finite dimensional exposition, is clear that V and W must have the same dimension, and thus $V = W = \mathbb{R}^n$. This will be different, when we consider infinite dimensional spaces and indeed consider different spaces V and W.

2.1.16 Lemma: The following statements are equivalent to the inf-sup condition (2.10):

$$\forall u \in V \;\exists w \in W \;:\; a(u,w) \geq \gamma \|u\|_V \|w\|_W \tag{2.11}$$

$$\forall u \in V \;\exists w \in W \;:\; \begin{cases} \|w\|_W \leq \|u\|_V \\ a(u,w) \geq \gamma \|u\|_V^2 \end{cases} \tag{2.12}$$

$$\forall u \in V \;\exists w \in W \;:\; \begin{cases} \gamma\|w\|_W \leq \|u\|_V \\ Aw = u \end{cases} \tag{2.13}$$

2.1.17 Problem: Prove Lemma 2.1.16.

2.2 Infinite dimensional Hilbert spaces

2.2.1. In the previous section, we derived quantitative conditions to ensure the invertibility of a matrix A or its restriction to its cokernel $\ker(A)^\perp$. The arguments there have a natural extension to infinite dimensional Hilbert spaces, which we will derive in this section. We already saw in Problem 2.1.3 that we may run into trouble if the range of A is not closed. On the other hand, it turns out that most notions of linear algebra related to orthogonality can be maintained in Hilbert spaces if closed subspaces are considered. We begin by citing the most important results.

2.2.2 Definition: Let $W \subset V$ be a subspace of a Hilbert space V. We define its **polar space** $W^0 \subset V^*$ and its **orthogonal complement** $W^\perp \subset V$ by

$$\begin{aligned} W^0 &= \left\{ f \in V^* \middle| \langle f, w \rangle_{V^* \times V} = 0 \quad \forall w \in W \right\}, \\ W^\perp &= \left\{ v \in V \middle| \langle v, w \rangle_V = 0 \quad \forall w \in W \right\}. \end{aligned} \tag{2.14}$$

For a subspace $U \subset V^*$, we define its polar space

$$U^0 = \left\{ v \in V \middle| \langle u, v \rangle_{V^* \times V} = 0 \quad \forall u \in U \right\} \tag{2.15}$$

2.2.3 Lemma: The polar space W^0 and the orthogonal complement $W^\perp$ of a subspace $W \subset V$ are both closed. So is the polar space U^0 of a subspace $U \subset V^*$.

Proof. Consider the mapping

$$\begin{aligned} \Phi_w \colon V^* &\to \mathbb{R}, \\ v &\mapsto \langle v, w \rangle_{V^* \times V}. \end{aligned}$$

For any w, the kernel of Φ is closed as the pre-image of a closed set. W^0 is closed since it is the intersection of these kernels for all $w \in W$.

The inner product is continuous on $V \times V$. Therefore, the mapping

$$\begin{aligned} \varphi_w \colon V &\to \mathbb{R}, \\ v &\mapsto \langle v, w \rangle, \end{aligned}$$

is continuous. The argument continues as above. Similar for U^0. □

2.2.4 Theorem: Let W be a subspace of a Hilbert space V and $W^\perp$ its orthogonal complement. Then, $W^\perp = \overline{W}^\perp$. Further, $V = W \oplus W^\perp$ if and only if W is closed.

Proof. Clearly, $\overline{W}^{\perp} \subset W^{\perp}$ since $W \subset \overline{W}$. Let now $u \in W^{\perp}$. Then, $\varphi = \langle u, \cdot \rangle$ is a continuous linear functional on V. Therefore, if a sequence $w_n \subset W$ converges to $w \in \overline{W}$, we have

$$\langle u, w \rangle = \lim_{n\to\infty} \langle u, w_n \rangle = 0.$$

Hence, $u \in \overline{W}^{\perp}$ and $W^{\perp} = \overline{W}^{\perp}$.

Now, the "only if" follows by the fact, that if W is not closed, there is an element $w \in \overline{W}$ but not in W such that $\langle w, u \rangle = 0$ for all $u \in W^{\perp}$. Thus, $w \notin W^{\perp}$ and consequently $w \notin W^{\perp} \oplus W$.

Let now W be closed. We show that there is a unique decomposition

$$v = w + u, \qquad w \in W,\ u \in W^{\perp}, \tag{2.16}$$

which is equivalent to $V = W \oplus W^{\perp}$. Uniqueness follows, since

$$v = w_1 + u_1 = w_2 + u_2$$

implies that for any $y \in V$

$$0 = \langle w_1 - w_2 + u_1 - u_2, y \rangle = \langle w_1 - w_2, y \rangle + \langle u_1 - u_2, y \rangle.$$

Choosing $y = u_1 - u_2$ and $w_1 - w_2$ in turns, we see that one of the inner products vanishes for orthogonality and the other implies that the difference is zero.

If $v \in W$, we choose $w = v$ and $u = 0$. For $v \notin W$, we prove existence by considering that due to the closedness of W there holds

$$d = \inf_{w \in W} \|v - w\| > 0.$$

Let w_n be a minimizing sequence. Using the parallelogram identity

$$\|a + b\|^2 + \|a - b\|^2 = 2\|a\|^2 + 2\|b\|^2,$$

we prove that $\{w_n\}$ is a Cauchy sequence by

$$\begin{aligned}
\|w_m - w_n\|^2 &= \|(v - w_n) - (v - w_m)\|^2 \\
&= 2\|v - w_n\|^2 + 2\|v - w_m\|^2 - \|2v - w_m - w_n\|^2 \\
&= 2\|v - w_n\|^2 + 2\|v - w_m\|^2 - 4\left\|v - \frac{w_m + w_n}{2}\right\|^2 \\
&\le 2\|v - w_n\|^2 + 2\|v - w_m\|^2 - 4d^2,
\end{aligned}$$

since $(w_m + w_n)/2 \in W$ and d is the infimum. Now we use the minimizing property to obtain

$$\lim_{m,n\to\infty} \|w_m - w_n\|^2 = 2d^2 - 2d^2 - 4d^2 = 0.$$

By completeness of V, $w = \lim w_n$ exists and by the closedness of W, we have $w \in W$. Let $u = v - w$. By continuity of the norm, we have $\|u\| = d$. It remains to show that $u \in W^\perp$. To this end, we introduce the variation $w + \varepsilon\tilde{w}$ with $\tilde{w} \in W$ to obtain

$$\begin{aligned} d^2 &\le \|v - w - \varepsilon\tilde{w}\|^2 \\ &= \|u\|^2 - 2\varepsilon\langle u, \tilde{w}\rangle + \varepsilon^2\|\tilde{w}\|, \end{aligned}$$

implying for any $\varepsilon > 0$

$$0 \le -2\varepsilon\langle u, \tilde{w}\rangle + \varepsilon^2\|\tilde{w}\|,$$

which requires $\langle u, \tilde{w}\rangle = 0$. □

2.2.5 Definition: Let V be a Hilbert space and $W \subset V$ be a closed subspace. For a vector $v \in V$, let $v = w + u$ be the unique decomposition with $w \in W$ and $u \in W^\perp$. Then we call w and u the **orthogonal projections** of v into W and $W^\perp$, respectively. We write

$$\Pi_W = w, \qquad \Pi_{W^\perp} = u.$$

2.2.6 Lemma: Let V be a Hilbert space and $W \subset V$ be a closed subspace. Then, the polar space $W^0 \subset V^*$ and the orthogonal space $W^\perp$ can be isometrically identified by Riesz representation.

Proof. For every f in the dual of $W^\perp$, define $g \in V^*$ by

$$\langle g, v\rangle_{V^* \times V} = \langle f, \Pi_{V^\perp} v\rangle_{(V^\perp)^* \times V^\perp}.$$

Clearly, $g(v) = 0$ for $v \in W$, therefore $g \in W^0$. □

2.2.7 Theorem: Let V, W be Hilbert spaces and $A\colon V \to W$ a continuous linear operator. Then, the following statements are equivalent:

$$\begin{aligned} &\operatorname{im}(A) \text{ is closed in } W, \\ &\operatorname{im}(A^T) \text{ is closed in } V^*, \\ &\operatorname{im}(A) = \ker\left(A^T\right)^0, \\ &\operatorname{im}(A^T) = \ker\left(A\right)^0. \end{aligned} \tag{2.17}$$

Remark 2.2.8. This is the famous *closed range theorem* by Banach. It actually holds under weaker assumptions, for instance V, W only Banach spaces. The proof can be found for instance in [Yos80, p. 205–209].

2.2.9 Theorem: Let $A\colon V \to W$ be continuous and surjective. Then, the image $A(U) \subset W$ of any open set $U \subset V$ is open.

Remark 2.2.10. This is the *open mapping theorem* by Banach. The proof can be found for instance in [Yos80, p.75–76].

2.2.11 Lemma: Let $A\colon V \to W$ be continuous. Then, $\operatorname{im}(A)$ is closed in W if and only if there exists $\gamma > 0$ such that

$$\forall w \in \operatorname{im}(A)\ \exists v \in V \quad Av = w\ \wedge\ \gamma\|v\|_V \le \|w\|_W. \tag{2.18}$$

Proof. We first show that the inf-sup condition (2.18) implies $\operatorname{im}(A)$ closed. To this end, let $\{w_n\}$ be a Cauchy sequence in $\operatorname{im}(A)$ converging to a point $w \in W$. Thus, there is a sequence $\{v_n\}$ in V such that $Av_n = w_n$ and $\gamma\|v_n\| \le \|w_n\|$. Hence,

$$\|v_m - v_n\|_V \le \frac{1}{\gamma}\|w_m - w_n\|_W,$$

and $\{v_n\}$ is a Cauchy sequence in V. Therefore, $v_n \to v \in V$ and due to continuity of A we obtain $Av = w$ and thus $w \in \operatorname{im}(A)$.

Conversely, let $\operatorname{im}(A)$ be closed in W. Thus, it is a Banach space and the open mapping theorem applies to $A\colon V \to \operatorname{im}(A)$. We map the open unit ball $B_1(0) \subset V$ and obtain that $A(B_1(0))$ is open in $\operatorname{im}(A)$, implying that there is an open ball $B_\delta(0) \subset A(B_1(0))$. This is sufficient to construct v:

Let $w \in \operatorname{im}(A)$. Then,

$$\tilde{w}\frac{\delta}{2}\frac{w}{\|w\|} \in B_\delta(0) \subset A(B_1(0)).$$

Hence, there is $v \in V$ with $\|v\| < 1$ such that $Av = \tilde{w}$, which proves the lemma. $\square$

2.2.12 Theorem: Let $a(\cdot,\cdot)$ on $V \times W$ be a bounded bilinear form such that

$$a(v,w) \le M\|v\|_V\|w\|_W, \tag{2.19}$$

and $A\colon V \to W^*$ its associated operator. Then, the following statements are equivalent:

1. There exists $\gamma > 0$ such that

$$\inf_{w\in W}\sup_{v\in V}\frac{a(v,w)}{\|v\|_V\|w\|_W} \ge \gamma. \tag{2.20}$$

2. The operator $A^T\colon W \to \ker(A)^0$ is an isomorphism and

$$\|A^Tw\|_{V^*} \ge \gamma\|w\|_W \qquad \forall w \in W. \tag{2.21}$$

3. The operator $A\colon \ker(A)^\perp \to W^*$ is an isomorphism and

$$\|Av\|_{W^*} \ge \gamma\|v\|_V \qquad \forall v \in \ker(A)^\perp. \tag{2.22}$$

Proof. First, we show the equivalence of the first two statements. Let us rephrase the inf-sup condition to

$$\|A^Tw\|_{V^*} = \sup_{v\in V}\frac{\langle A^Tw, v\rangle}{\|v\|_V} = \sup_{v\in V}\frac{a(v,w)}{\|v\|_V} \ge \gamma\|w\| \qquad \forall w \in W.$$

Thus, equations (2.20) and (2.21) are equivalent and we have already proven that the second statement implies the first. It remains to show the A^T is an isomorphism from W onto $\ker(A)^0$. Equation (2.21) implies that $A^T\colon W \to \mathrm{im}(A^T)$ is an isomorphism and its inverse is bounded by $1/\gamma$ (multiply both sides by A^{-1}). Using Lemma 2.2.11, we obtain that $\mathrm{im}(A^T)$ is closed in V^* and the closed range theorem settles the issue.

In order to prove equivalence of the second and third statement, we use Lemma 2.2.6 to isometrically identify $(\ker(A)^\perp)^*$ with $\ker(A)^0$. Thus, A is an isomorphism from $\ker(A)^\perp$ onto W^* if and only if A^T is an isomorphism from W onto $(\ker(A)^\perp)^* = \ker(A)^0$. and

$$\|A\|_{W^*\to\ker(A)^\perp} = \|A^T\|_{\ker(A)^0\to W}.$$

□

2.2.13 Corollary: Let $a(\cdot,\cdot)$ on $V\times W$ be a bounded bilinear form such that

$$a(v,w) \le M\|v\|_V\|w\|_W. \tag{2.23}$$

Let the inf-sup-condition

$$\inf_{w\in W}\sup_{v\in V}\frac{a(v,w)}{\|v\|_V\|w\|_W} \ge \gamma > 0$$

hold. Then, the problem finding $w\in W$ such that

$$a(v,w) = f(v) \qquad \forall v\in V,$$

has a unique solution for $f\in \ker(A)^0$ and

$$\|w\|_W \le \frac{1}{\gamma}\|f\|_{V^*}. \tag{2.24}$$

Remark 2.2.14. Corollary 2.2.13 exhibits an asymmetry between the left and right argument. In particular, we obtain a unique solution only for the adjoint operator A^T, which is exactly what we need, when we compute say a pressure from the divergence of a velocity field. In general, we consider the restriction of f to the polar set of the kernel in the above well-posedness result detrimental and would prefer a result that holds for all $f\in V^*$. This on the other hand requires $\ker(A)=\{0\}$, or $\overline{\operatorname{im}(A^T)} = W^*$. Then, on the other hand, we see that $\operatorname{im}(A^T)$ is closed since $\operatorname{im}(A)$ is closed and the closed range theorem holds. Therefore, we obtain the following theorem for the case that we require a unique solution for all right hand sides.

2.2.15 Theorem: Let $a(\cdot,\cdot)$ on $V \times W$ be a bounded bilinear form such that

$$a(v,w) \le M \|v\|_V \|w\|_W. \tag{2.25}$$

Let for some $\gamma > 0$ the inf-sup-conditions

$$\inf_{w\in W} \sup_{v\in V} \frac{a(v,w)}{\|v\|_V \|w\|_W} \ge \gamma,$$
$$\inf_{v\in V} \sup_{w\in W} \frac{a(v,w)}{\|v\|_V \|w\|_W} \ge \gamma$$

hold. Then, the problem finding $v \in V$ such that

$$a(v,w) = f(w) \qquad \forall w \in W,$$

has a unique solution for $f \in W^*$ and

$$\|v\|_V \le \frac{1}{\gamma} \|f\|_{W^*}. \tag{2.26}$$

Remark 2.2.16. If we compare Theorem 2.2.15 with Corollary 2.2.13, we see that the only difference lies in the fact that the second inf-sup condition ensures surjectivity of A by injectivity of A^T. In some cases it may be difficult to prove both inf-sup conditions. Then, it is sufficient to prove one inf-sup condition, say the first, and then only

$$\inf_{v\in V} \sup_{w\in W} \frac{a(v,w)}{\|v\|_V \|w\|_W} > 0,$$

thus, injectivity of A^T. Although we verify less than the assumptions of Theorem 2.2.15, the closed range theorem saves us from the additional work. We further note that this notion is symmetric between A and A^T, that is, it is sufficient to prove inf-sup for either operator and injectivity for the other.

2.3 The inf-sup condition for mixed problems

2.3.1. In the previous section, we have developed a framework for well-posedness of problems which are not V-elliptic. In principle, this theory can be applied to the bilinear form $\mathcal{A}((u,p),(v,q))$ as a whole. On the other hand, we can formally split the solution of a constrained minimization problem into the reduced problem and then computing the Lagrange multiplier, which more clearly exhibits the relation of the two spaces V and Q involved in the mixed formulation. Here are the resulting theorems.

2.3.2 Theorem: Let V and Q be Hilbert spaces and let the mixed bilinear form

$$\mathcal{A}\left(\begin{pmatrix}u\\p\end{pmatrix},\begin{pmatrix}v\\q\end{pmatrix}\right) = a(u,v) + b(v,p) + b(u,q)$$

be defined and bounded for any $u, v \in V$ and $p, q \in Q$. Then, the problem

$$\mathcal{A}\left(\begin{pmatrix}u\\p\end{pmatrix},\begin{pmatrix}v\\q\end{pmatrix}\right) = \langle f, v\rangle + \langle g, q\rangle \quad \forall v \in V, q \in Q,$$

has a unique solution for any $f \in V^*$ and any $g \in Q^*$ if and only if there exists $\gamma > 0$ such that

$$\forall \begin{bmatrix}u \in V\\p \in Q\end{bmatrix} \; \exists \begin{bmatrix}v \in V\\q \in Q\end{bmatrix} : \quad \mathcal{A}\left(\begin{pmatrix}u\\p\end{pmatrix},\begin{pmatrix}v\\q\end{pmatrix}\right) \geq \gamma \|(u,p)\|_{V\times Q}\|(v,q)\|_{V\times Q},$$

and vice versa.

Proof. Straight application of Theorem 2.2.15. □

2.3.3 Theorem: Let V and Q be Hilbert spaces and let

$$\ker(B) = \left\{v \in V \middle| b(v,q) = 0 \;\forall q \in Q\right\}. \tag{2.27}$$

Then, the problem finding $(u,p) \in V \times Q$ such that

$$a(u,v) + b(v,p) + b(u,q) = f(v) \quad \forall v \in V, q \in Q, \tag{2.28}$$

is well-posed if and only if the problem finding $u \in \ker(B)$ such that

$$a(u,v) = f(v) \quad \forall v \in \ker(B) \tag{2.29}$$

is well-posed for any $f \in V^*$ and there is a positive constant β such that

$$\inf_{q\in Q}\sup_{v\in V}\frac{b(v,q)}{\|v\|_V\|q\|_Q} \geq \beta. \tag{2.30}$$

Proof. By requiring well-posedness of the reduced problem, $u \in V$ is well-determined and bounded by the data $f \in V^*$ without knowledge of the Lagrange multiplier. Hence, with $u \in V$ given and $b(u,q) = 0$, the problem of determining the Lagrange multiplier p reduces to

$$b(v,p) = f(v) - a(u,v), \qquad \forall v \in V. \tag{2.31}$$

Applying Corollary 2.2.13 to the bilinear form $b(.,.)$, we deduce that this equation has a unique solution $p \in Q$ if and only if $f(v) - a(u,v) \in \ker(B)^0$, which is the reduced problem. □

Remark 2.3.4. Since V is a Hilbert space, the decomposition $V = \ker(B) \oplus \ker(B)^\perp$ is uniquely determined and there is a corresponding decomposition $V^* = \ker(B)^0 + (\ker(B)^\perp)^0$, such that $f = f^0 + f^\perp$ above. The way we solve the reduced problem first and then compute the Lagrange multiplier implies that the solution u only depends on $f^\perp$, while the Lagrange multiplier p only depends on f^0.

Remark 2.3.5. We have imposed well-posedness of the reduced problem only in an abstract way. Depending on $a(.,.)$ we can formulate two conditions: ellipticity on $\ker(B)$ or inf-sup stability on $\ker(B)$. Indeed, most problems considered in this class will have symmetric bilinear forms $a(.,.)$, such that ellipticity serves as our usual assumption. In these cases, note that V-ellipticity already implies the well-posedness on $\ker(B)$.

2.3.6 Problem: Show that Theorem 2.3.3 can be extended to the case with right hand side $f(v) + g(q)$ with $g \in Q^*$.

2.3.7. We summarize the result of this section in an assumption for well-posedness which will be the basis for further results in this course. We know from the discussion above that this assumption is only sufficient and weaker conditions may be imposed on $a(.,.)$. But indeed, it helps us through a lot of problems and is a good compromise between generality and ease of use.

2.3.8 Assumption: Let V and Q be Hilbert spaces and let $a(.,.)$ and $b(.,.)$ be bounded bilinear forms on $V \times V$ and $V \times Q$, respectively. We define their norms as the smallest constants such that for all arguments there holds

$$a(u,v) \le \|a\| \|u\|_V \|v\|_V, \qquad b(v,q) \le \|b\| \|v\|_V \|q\|_Q. \tag{2.32}$$

With these forms, we associate bounded operators A, B, and B^T according to Definition 1.2.2. With $b(.,.)$ we associate the spaces

$$\begin{aligned} \ker(B) &= \left\{ v \in V \middle| b(v,q) = 0 \; \forall q \in Q \right\}, \\ \ker(B)^T &= \left\{ q \in Q \middle| b(v,q) = 0 \; \forall v \in V \right\}. \end{aligned} \tag{2.33}$$

Furthermore, we assume that $a(.,.)$ is positive semi-definite on V and elliptic on $\ker(B)$,

$$a(u,u) \ge 0 \quad \forall u \in V, \qquad a(u,u) \ge \gamma \|u\|_V^2 \quad \forall u \in \ker(B). \tag{2.34}$$

2.4 Galerkin approximation of mixed problems

2.4.1. The Galerkin approximation of mixed problems starts out the same way as for elliptic problems, namely, choose discrete subspaces $V_h \subset V$ and $Q_h \subset Q$. There is a fundamental difference though: the inf-sup condition is not inherited automatically on the subspaces like V-ellipticity. It actually becomes an additional requirement on the choice of V_h and Q_h. We will thus work our way in several steps towards the final result.

2.4.2 Definition: Let $V_h \subset V$ and $Q_h \subset Q$. Then, we define the subspace

$$\ker(B_h) = \left\{ v_h \in V_h \middle| b(v_h, q_h) = 0 \quad \forall q_h \in Q_h \right\}. \tag{2.35}$$

Furthermore, we define the affine space

$$V_h^g = \left\{ v_h \in V_h \middle| b(v_h, q_h) = g(q) \quad \forall q_h \in Q_h \right\}. \tag{2.36}$$

2.4.3 Definition: We introduce the mixed discrete problem: find $(u_h, p_h) \in V_h \times Q_h$ such that

$$a(u_h, v_h) + b(v_h, p_h) + b(u_h, q_h) = f(v_h) + g(q_h), \quad \forall v_h \in V_h, q_h \in Q_h, \tag{2.37}$$

and the discrete reduced problem: find $u_h \in V_h^g$ such that

$$a(u_h, v_h) = f(v_h), \qquad \forall v_h \in \ker(B_h). \tag{2.38}$$

2.4.4 Theorem: Let V_h^g be nonempty and let $a(.,.)$ and $b(.,.)$ be bounded with norms $\|a\|$ and $\|b\|$, respectively, and let there be constant and γ_h such that

$$a(v_h, v_h) \geq \gamma_h \|v_h\|_V^2, \qquad \forall v_h \in \ker(B_h). \tag{2.39}$$

Let furthermore the continuous mixed problem be well-posed with solution $(u,p) \in V \times Q$. Then, the discrete reduced problem (2.38) has a unique solution and there holds

$$\|u - u_h\|_V \leq \left(1 + \frac{\|a\|}{\gamma_h}\right) \inf_{w_h \in V_h^g} \|u - w_h\|_V + \frac{\|b\|}{\gamma_h} \inf_{q_h \in Q_h} \|p - q_h\|_Q \tag{2.40}$$

Proof. Let $u_h^g \in V_h^g$ arbitrary. By the ellipticity assumption, there is a unique function $u_h^0 \in \ker(B_h)$ such that

$$a(u_h^0, v_h) = f(v_h) - a(u_h^g, v_h), \qquad \forall v_h \in \ker(B_h)\,.$$

Hence, $u_h = u_h^g + u_h^0$ is the unique solution to (2.38). Choose now $w_h \in V_h^g$ arbitrarily. Then, $v_h = u_h - w_h \in \ker(B_h)$ and using

$$f(v_h) = a(u, v_h) - b(v_h, p),$$

we obtain

$$\begin{aligned} a(v_h, v_h) &= f(v_h) - a(u_h - v_h, v_h) \\ &= a(u - w_h, v_h) - b(v_h, p) \\ &= a(u - w_h, v_h) - b(v_h, p - q_h) \end{aligned} \tag{2.41}$$

for any $q_h \in Q_h$, yielding

$$\gamma_h \|v_h\|_V^2 \le |a(v_h, v_h)| \le \|a\| \|u - w_h\|_V \|v_h\|_V + \|b\| \|p - q_h\|_Q \|v_h\|_V.$$

We conclude by the standard triangle inequality argument

$$\begin{aligned} \|u - u_h\|_V &\le \|u - w_h\|_V + \|u_h - w_h\|_V \\ &\le \|u - w_h\|_V + \frac{\|a\|}{\gamma_h} \|u - w_h\|_V + \frac{\|b\|}{\gamma_h} \|p - q_h\|_Q. \end{aligned}$$

The estimate follows since $w_h \in V_h^g$ and $q_h \in Q_h$ were chosen arbitrarily. □

Remark 2.4.5. Note that we only used ellipticity of $a(.,.)$ on the subspace $\ker(B_h)$ for the discrete problem and on $\ker(B)$ for the continuous problem. Since the union of two vector spaces is not a vector space, this is a strange condition. In practice, we will encounter two situations: either ellipticity holds on the whole space V or $\ker(B_h) \subset \ker(B)$, where the latter is again a vector space.

2.4.6 Corollary: If in addition to the assumptions of Theorem 2.4.4 there holds

$$\ker(B_h) \subset \ker(B)\,, \tag{2.42}$$

then

$$\|u - u_h\|_V \le \left(1 + \frac{\|a\|}{\gamma_h}\right) \inf_{w_h \in V_h^g} \|u - w_h\|_V \tag{2.43}$$

Proof. Consider that in equation (2.41) there holds $v_h \in \ker(B)$. □

2.4.7 Theorem: Assume in addition the assumptions of Theorem 2.4.4 that there are constants β_h, possibly depending on the parameter h, such that

$$\inf_{q_h \in Q_h} \sup_{v_h \in V_h} \frac{b(v_h, q_h)}{\|v_h\|_V \|q_h\|_Q} \geq \beta_h. \tag{2.44}$$

Then, there is a unique solution $p_h \in Q_h$ such that (u_h, p_h) is the unique solution to the discrete mixed problem (2.37). There are a constants $c_h^{(i)}$ only depending on $\|a\|$, $\|b\|$, γ_h and β_h such that

$$\|u - u_h\|_V \leq c_h^{(1)} \inf_{v_h \in V_h} \|u - v_h\|_V + c_h^{(2)} \inf_{q_h \in Q_h} \|p - q_h\|_Q \tag{2.45}$$

$$\|p - p_h\|_Q \leq c_h^{(3)} \inf_{v_h \in V_h} \|u - v_h\|_V + c_h^{(4)} \inf_{q_h \in Q_h} \|p - q_h\|_Q. \tag{2.46}$$

Proof. Applying Theorem 2.3.3 to the discrete problem (2.37), we conclude that there is a unique solution $(u_h, p_h) \in V_h \times Q_h$. Let us begin estimating the error by establishing the bound

$$\inf_{w_h \in V_h^g} \|u - w_h\|_V \leq \left(1 + \frac{\|b\|}{\beta_h}\right) \inf_{v_h \in V_h} \|u - v_h\|_V. \tag{2.47}$$

By the third condition in Theorem 2.2.12, there is a unique $z_h \in \ker(B_h)^\perp$ such that

$$B_h z_h = B_h(u - v_h), \qquad \forall v_h \in V_h,$$

and

$$\|z_h\|_V \leq \frac{1}{\beta_h} \|B_h(u - v_h)\|_{Q_h^*} \leq \frac{\|b\|}{\beta_h} \|u - v_h\|_V.$$

Let $w_h = z_h + v_h$. Then, $w_h \in V_h^g$ since

$$b(w_h, q_h) = b(u - v_h, q_h) = b(u, q_h) = g(q_h), \qquad \forall q_h \in Q_h.$$

Furthermore,

$$\|u - w_h\|_V \leq \|u - v_h\|_V + \|z_h\|_V \leq \left(1 + \frac{\|b\|}{\beta_h}\right) \|u - v_h\|_V.$$

Since $v_h \in V_h$ was chosen arbitrarily, we have proven (2.47) and thus by Theorem 2.4.4 the estimate for $\|u - u_h\|_V$ with

$$c_h^{(1)} = \left(1 + \frac{\|b\|}{\beta_h}\right)\left(1 + \frac{\|a\|}{\gamma_h}\right), \qquad c_h^{(2)} = \left(1 + \frac{\|b\|}{\beta_h}\right)\frac{\|b\|}{\gamma_h}. \tag{2.48}$$

It remains to prove the estimate for $\|p - p_h\|_Q$. Using Galerkin orthogonality for the test function v_h, we obtain

$$a(u - u_h, v_h) + b(v_h, p - p_h) = 0. \tag{2.49}$$

Hence, for any $q_h \in Q_h$ there is by the inf-sup condition $v_h \in V_h$ with $\|v_h\|_V = 1$ such that

$$\begin{aligned}\beta_h \|p_h - q_h\|_Q &\le b(v_h, p_h - q_h)\\ &= a(u - u_h, v_h) + b(v_h, p - q_h)\\ &\le \|a\| \|u - u_h\|_V + \|b\| \|p - q_h\|_Q.\end{aligned}$$

Again, the estimate for $\|p - p_h\|_Q$ follows by triangle inequality. □

Remark 2.4.8. Indeed, if (2.44) holds, we do not have to require that V_h^g is nonempty anymore, since $B\colon V \to Q^*$ is surjective.

Remark 2.4.9. We purposely proved the preceding theorems with γ_h and β_h depending on the parameter h, typically the mesh size. This is the minimal condition for well-posedness of the discrete problems. Nevertheless, this well-posedness is not uniform in h, which causes loss of approximation, as the following problem shows. Therefore, we will only be satisfied with uniform inf-sup constants in applications.

2.4.10 Problem: Let the following interpolation estimates hold:

$$\inf_{v_h \in V_h} \|u - v_h\|_V = \mathcal{O}(h^k), \qquad \inf_{q_h \in Q_h} \|p - q_h\|_V = \mathcal{O}(h^k).$$

Then, the estimates in Theorem 2.4.7 are asymptotically optimal if and only if there are constants $\tilde{\gamma} > 0$ and $\tilde{\beta} > 0$ independent of h such that

$$\gamma_h \ge \tilde{\gamma}, \quad \beta_h \ge \tilde{\beta}, \tag{2.50}$$

independent of h.

2.4.11. As we can see from the form

$$\forall q_h \in Q_h\ \exists v_h \in V_h\colon \quad B_h v_h = q_h \quad \wedge \quad \|v_h\|_V \le \|q_h\|_Q,$$

the uniform, discrete inf-sup condition introduces a compatibility condition between the spaces V_h and Q_h. An immediate necessary condition is

$$\dim V_h \ge \dim Q_h. \tag{2.51}$$

We often say that the space V_h is "rich enough" to control functions in Q_h. Obviously, counting dimensions is not sufficient, since we could have added basis functions in $\ker(B_h)$. Even the condition

$$\dim \ker(B_h)^\perp = \dim Q_h$$

is necessary and sufficient only for the existence of an inf-sup constant β_h depending on h. Therefore, we need a stronger argument in order to ensure compatibility of the discrete spaces. Such an argument is the following lemma by Fortin. The projection operator Π_{V_h} introduced there is usually referred to as **Fortin projection**.

2.4.12 Lemma: Let the inf-sup condition for the bilinear form $b(.,.)$ hold on $V \times Q$ with a constant $\beta > 0$. Then, it holds on $V_h \times Q_h$ uniformly with a constant $\tilde{\beta} > 0$ if and only if there exists a linear operator $\Pi_{V_h} : V \to V_h$ satisfying for any $v \in V$

$$b(v - \Pi_{V_h} v, q_h) = 0, \qquad \forall q_h \in Q_h, \tag{2.52}$$

$$\|\Pi_{V_h} v\|_V \le c\|v\|_V, \tag{2.53}$$

with c independent of h. There holds $\tilde{\beta} = \beta/c$.

Proof. Assume first that Π_{V_h} exists. then, there holds for any $q_h \in Q_h$

$$\sup_{v_h \in V_h} \frac{b(v_h, q_h)}{\|v_h\|_V} \ge \sup_{v \in V} \frac{b(\Pi_{V_h} v_h, q_h)}{\|\Pi_{V_h} v_h\|_V} = \sup_{v \in V} \frac{b(v, q_h)}{\|\Pi_{V_h} v_h\|_V} \ge \frac{\beta}{c}\|q_h\|_Q.$$

Conversely, we assume the existence of a uniform, discrete inf-sup constant $\tilde{\beta} > 0$. Then, for any $v \in V$ let $g(.) = b(v, .) \in Q_h^*$. By Theorem 2.2.12, there is a unique element $\Pi_{V_h} v \in \ker(B_h)^\perp$ such that

$$b(\Pi_{V_h} v, q_h) = b(v, q_h), \qquad \forall q_h \in Q_h$$

and

$$\|\Pi_{V_h} v\|_V \le \frac{1}{\tilde{\beta}}\|B_h v\|_{Q_H^*} \le \frac{\|b\|}{\tilde{\beta}}\|v\|_V.$$

Thus, Π_{V_h} is bounded and (2.52) holds with $c = \|b\|/\tilde{\beta}$. □

2.4.13 Problem: Let $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{k\times n}$, $k \le n$. Moreover, assume that B has full rank and that A is symmetric and positive definite. Consider the problem

$$\begin{pmatrix} A & B^* \\ B & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} F \\ G \end{pmatrix} \tag{*}$$

1. Prove that then $S := BA^{-1}B^*$ is symmetric and positive definite, too. How can this matrix be used to solve (*)?
2. Show that

$$P := I - B^*(BB^*)^{-1}B.$$

is a projector on the kernel of B with $\|P\|_2 = 1$.

3. Show for the case $G = 0$ that x is a solution of

$$PAPx = PF$$

if (x, y) is a solution of (*).

2.5 Bringing back $c(p, q)$

2.5.1. The key to the mixed analysis which is also underlying our quasi-best-approximation result was a splitting of the solution process into the reduced problem for u and then applying the inf-sup condition for $b(.,.)$ in order to estimate p. This way, we will be able to obtain estimates for the Stokes problem, but we have tacitly abandoned weakly compressible elasticity. Indeed, the mixed form of the Lamé-Navier equations is not a constrained minimization problem. In this section, we will fill the gap and derive estimates for the solution of this problem which are robust in λ.

In the Lamé-Navier equations, we had

$$c(p, q) = -\tfrac{1}{\lambda}\langle q, p\rangle_{L^2(\Omega)},$$

which suggests assuming symmetric and Q-elliptic. But, we want estimates independent of λ! Therefore, we should only require semi-definite, which on the other hand turns out a bit too weak.

2.5.2 Assumption: In addition to Assumption 2.3.8, let $c(.,.)$ be positive semi-definite and elliptic on $\ker\left(B^T\right)$,

$$c(q, q) \ge u \quad \forall q \in Q, \qquad c(q, q) \ge \gamma_c \|q\|_Q^2 \quad \forall q \in \ker\left(B^T\right). \tag{2.54}$$

Remark 2.5.3. Again, this assumption is not necessary for the analysis, but it yields a convenient and useful theorem which goes far beyond weakly compressible elasticity and covers stabilized methods for spaces where the inf-sup condition for $b(.,.)$ does not hold for the whole space Q.

2.5.4 Theorem: Let Assumption 2.5.2 hold and let $a(.,.)$ and $c(.,)$ be symmetric. In addition, let there be $\beta > 0$ such that

$$\begin{aligned} \inf_{q\in\ker(B^T)^\perp} \sup_{v\in V} \frac{b(v,q)}{\|v\|_V \|q\|_Q} \geq \beta \\ \inf_{v\in\ker(B)^\perp} \sup_{q\in Q} \frac{b(v,q)}{\|v\|_V \|q\|_Q} \geq \beta \end{aligned} \tag{2.55}$$

Then, the problem finding $(u,p) \in V \times Q$ such that

$$a(u,v) + b(v,p) + b(u,q) - c(p,q) = f(v) + g(q) \qquad \forall v \in V, q \in Q \tag{2.56}$$

has a unique solution for all $f \in V^*$ and $g \in Q^*$ and there is a constant C such that

$$\|u\|_V + \|p\|_Q \leq C\big(\|f\|_{V^*} + \|g\|_{Q^*}\big). \tag{2.57}$$

Proof. First, note that the ellipticity assumptions as well as the inf-sup conditions are symmetric in V and Q. Indeed, replacing the test functions and the form $b(.,.)$ by their negatives, we can transform the problem into one where V and Q have exchanged their roles. Thus, it is sufficient to show well-posedness for $f = 0$. The same result then holds for $g = 0$ and it holds for both nonzero by linearity.

We note that by the inf-sup conditions both $\operatorname{im}(B)$ and $\operatorname{im}(B^T)$ are closed. Thus, we can decompose $u = u^0 + u^\perp$ with $u^0 \in \ker(B)$ and $u^\perp$ in its orthogonal complement. Assuming $f = 0$ and testing with $q = 0$ we obtain the equation

$$a(u,v) = -b(v,p) = 0. \tag{2.58}$$

In particular, testing with $v = u^0$ yields

$$a(u,u^0) = -b(u^0,p) = 0.$$

Hence,

$$\gamma\|u^0\|_V^2 \leq a(u^0,u^0) = -a(u^\perp,u^0) \leq \|a\|\|u^\perp\|\|u^0\|,$$

which implies

$$\|u^0\|_V \leq \frac{\|a\|}{\gamma} \|u^\perp\|_V. \tag{2.59}$$

Testing with $v = u$ and $q = -p$ yields

$$a(u,u) + c(p,p) \leq g(p) = g^0(p^0) + g^\perp(p^\perp),$$

where $p^0 \in \ker\left(B^T\right)$, $g^\perp \in \ker\left(B^T\right)^0$ and the other two are defined by orthogonality in Q and Q^*, respectively. Let first $g^0 = 0$. Then, by (2.58) and the inf-sup condition for p, there is $v \in V$ with $\|v\|_V = 1$ such that

$$\beta\|p^\perp\| \leq |b(v,p^\perp)| = |b(v,p)| = |a(u,v)| \leq \sqrt{a(u,u)}\sqrt{a(v,v)},$$

by the Bunyakovsky-Cauchy-Schwarz inequality for symmetric bilinear forms. Therefore, squaring and using the definition of the operator norm of $g^\perp$ yields

$$\|p^\perp\|_Q \leq \frac{\|a\|}{\beta^2} \|g^\perp\|_{Q^*}.$$

Furthermore, we have

$$c(p,p^0) = b(u,p^0) - g^\perp(p^0) = 0.$$

Hence,

$$\gamma_c \|p^0\|_Q^2 \leq c(p^0,p^0) = c(p^\perp,p^0) \leq \|c\|\|p^0\|_Q\|p^\perp\|_Q,$$

concluding

$$\|p^0\|_Q \leq \frac{\|a\|\|c\|}{\gamma_c\beta^2} \|g^\perp\|_{Q^*}.$$

We continue our proof for $g^\perp = 0$ and $g^0 \neq 0$. Testing with $q = p^0$, we obtain

$$\begin{aligned}
c(p^0,p^0) &= c(p,p^0) - c(p^\perp,p^0) \\
&= b(u,p^0) - g^0(p^0) - c(p^\perp,p^0) \\
&\leq \|g^0\|_{Q^*}\|p^0\|_Q + \|c\|\|p^\perp\|_Q\|p^0\|_Q,
\end{aligned}$$

yielding

$$\|p^0\|_Q \leq \frac{1}{\gamma_c}\left(\|g^0\|_{Q^*} + \|c\|\|p^\perp\|_Q\right).$$

$p^\perp$ is estimated as before by

$$\begin{aligned}
\|p^\perp\|_Q^2 &\leq \frac{\|a\|}{\beta^2}\|g^0\|_{Q^*}\|p^0\|_Q \\
&\leq \frac{\|a\|}{\gamma_c\beta^2}\|g^0\|_{Q^*}^2 + \frac{\|a\|\|c\|}{\gamma_c\beta^2}\|g^0\|_{Q^*}\|p^\perp\|_Q \\
&\leq \frac{\|a\|}{\gamma_c\beta^2}\|g^0\|_{Q^*}^2 + \frac{1}{2}\|p^\perp\|_Q^2 + \frac{\|a\|^2\|c\|^2}{2{\gamma_c}^2\beta^4}\|g^0\|_{Q^*}^2.
\end{aligned}$$

We conclude that $p^\perp$ and p^0 are bounded by g^0. Summing up, we obtain for $f = 0$ and $g \in Q^*$ the estimate

$$\|p\|_Q \le c\|g\|_{Q^*}.$$

We estimate u^0 by $u^\perp$ using (2.59) and $u^\perp$ by the inf-sup condition, choosing $q \in Q$ with $\|q\|_Q = 1$ such that

$$\beta\|u^\perp\| = b(u,q) = c(p,q) + g(q) \le \|c\|\|p\|_Q + \|g\|_{Q^*}.$$

Thus, we have estimated all components of the solution by the norm of g, assuming $f = 0$. Now we conclude the proof by reverting the roles of u and p, respectively f and g. □

2.5.5. The extension of Theorem 2.4.7 to the saddle-point problem of Definition 1.2.3 with bilinear form $c(.,.)$ is even more cumbersome than this theorem. Nevertheless, the use of residual operators as a technique to structure the proof of convergence is instructive and may come handy at some point.

2.5.6 Definition: For the saddle-point problem

$$a(u,v) + b(v,p) + b(u,q) - c(p,q),$$

and functions $w_h \in V_h$ and $r_h \in Q_h$ we introduce the the residual operators $R_f \in V_h^*$ and $R_g \in Q_h^*$ as

$$\begin{aligned} R_f(v_h) &= a(u - w_h, v_h) + b(v_h, p - r_h) \\ R_g(q_h) &= b(u - w_h, q_h) - c(p - r_h, q_h). \end{aligned}$$

2.5.7 Corollary: Under Assumption 2.5.2, we have

$$\begin{aligned} |R_f(v_h)| &\le \big(\|a\|\|u - w_h\|_V + \|b\|\|p - r_h\|\big)\|v_h\|_V \\ |R_g(q_h)| &\le \big(\|b\|\|u - w_h\|_V + \|c\|\|p - r_h\|\big)\|q_h\|_Q. \end{aligned} \tag{2.60}$$

2.5.8 Lemma: Let the assumptions of Theorem 2.5.4 hold. Then, there are constants c_1 to c_4 independent of the solutions u, p, u_h, and p_h and the discretization parameter h, such that for any $v_h \in V_h$ and $q_h \in Q_h$

$$\begin{aligned} \|u_h - v_h\| &\le c_1\|R_f\|_{V_h^*} + c_2\|R_g\|_{Q_h^*} \\ \|p_h - q_h\| &\le c_3\|R_f\|_{V_h^*} + c_4\|R_g\|_{Q_h^*}. \end{aligned} \tag{2.61}$$

Proof. The proof is lengthy and follows the lines of the proof of well-posedness for Theorem 2.5.4. It is obtained by considering the components $u_h^0 - v_h^0$ and $u_h^\perp - v_h^\perp$ as well as $p_h^0 - q_h^0$ and $p_h^\perp - q_h^\perp$ separately. □

2.5.9. In spite of the bad treatment the proof of the previous lemma received in these notes, it contains the main parts of the convergence proof, and whenever a saddle-point problem including $c(.,.)$ is solved, it has to be reproduced. It is just the fact that the proof is overwhelmingly technical that led to the decision to leave this experience to the first time the reader actually needs this result.

2.5.10 Corollary: Let the assumptions of Theorem 2.5.4 hold. Then, there are constants c_1 to c_4 independent of the solutions u, p, u_h, and p_h and the discretization parameter h, such that

$$\begin{aligned} \|u-u_h\| &\le c_1 \inf_{v_h\in V_h} \|u-v_h\|_V + c_2 \inf_{q_h\in Q_h} \|p-q_h\|_Q \\ \|p-p_h\| &\le c_3 \inf_{v_h\in V_h} \|u-v_h\|_V + c_4 \inf_{q_h\in Q_h} \|p-q_h\|_Q. \end{aligned} \tag{2.62}$$

Proof. The proof begins with the standard approach with triangle inequality

$$\begin{aligned} \|u-u_h\| &\le \|u-v_h\| + \|v_h-u_h\| \\ \|p-p_h\| &\le \|p-q_h\| + \|q_h-p_h\|. \end{aligned}$$

Then, we employ Lemma 2.5.8 on the terms on the right and use the estimate of Corollary 2.5.7. □

Chapter 3

The Stokes problem

3.1 Well-posedness of the continuous problem

3.1.1. We begin our investigation into the Stokes problem by investigating the well-posedness of the continuous problem. This is particularly simple, since we have

$$a(u, v) = (\varepsilon(u), \varepsilon(v))$$

for the original Stokes problem in Definition 1.3.2 and

$$a(u, v) = (\nabla u, \nabla v)$$

for the simplified Stokes equations in Definition 1.3.6. From the standard theory for the Laplacian, we know that the second one is V-elliptic on $V = H_0^1(\Omega; \mathbb{R}^d)$. For the first one, we conclude this by using a Korn inequality. Therefore, we can already conclude a first result:

3.1.2 Lemma: Let $V = H_0^1(\Omega, \mathbb{R}^d)$ and $V_h \subset V$ a finite dimensional subspace. Then, $a(.,.)$ is elliptic on $\ker(B)$ and on $\ker(B_h)$ independent of the choices of Q and Q_h.

Remark 3.1.3. We focus here on no-slip boundary condition on the whole boundary as the exemplary case. Other boundary conditions are possible, but as soon as the Dirichlet boundary for one velocity component becomes too small, the ellipticity of $a(.,.)$ on V must be established by new arguments known for instance for Robin boundary conditions. In the extreme case of natural boundary conditions all around, V is the subspace of $H^1(\Omega, \mathbb{R}^d)$ obtained by dividing by the space of all translations for the simplified form and by the space of all rigid body movements.

Note that we have established already in Lemma 1.3.10 that the condition $V = H^1_0(\Omega, \mathbb{R}^d)$ implies the reduction of the pressure to the space $Q = L^2_0(\Omega)$ from Notation 1.3.11.

3.1.4. The previous lemma guarantees well-posedness of the reduced problem in all possible cases. Therefore, the remainder of this section is only concerned with the inf-sup condition for the divergence operator. We follow [GR86] in this presentation.

3.1.5 Lemma: Let $V = H^1_0(\Omega, \mathbb{R}^d)$. Then, the divergence operator $\nabla\cdot\colon V \to L^2(\Omega)$ is continuous and the subspace

$$V^0 = \ker(\nabla\cdot) = \left\{v \in V \middle| \nabla\cdot v = 0 \text{ a.e.}\right\}$$

is closed in V and V admits the orthogonal decomposition

$$V = V^0 \oplus V^\perp.$$

Proof. We have that

$$\|\nabla\cdot v\|^2_{L^2(\Omega)} = \int_\Omega \left(\sum \partial_i v_i\right)^2 \,\mathrm{d}x \le d \int_\Omega \sum |\partial_i v_i|^2 \,\mathrm{d}x \le d\|v\|^2_{H^1(\Omega;\mathbb{R}^d)}.$$

Thus, the divergence operator is a continuous mapping from V to $L^2(\Omega)$. The definition of V^0 is equivalent to the definition of zero in $L^2(\Omega)$. Finally, since the kernel is the pre-image of a closed set under a continuous map, it is closed. The existence of the decomposition follows from Theorem 2.2.4. □

3.1.6 Lemma: If $f \in V^* = H^{-1}(\Omega; \mathbb{R}^d)$ satisfies

$$f(v) = 0 \quad \forall v \in V^0,$$

then, there exists $p \in L^2(\Omega)$ such that

$$f = \nabla p.$$

If Ω is connected, then p is unique up to an additive constant.

Proof. First, we identify $L^2(\Omega)$ with its dual. Then, by

$$\langle -\nabla p, v\rangle_{V^*\times V} = \langle p, \nabla\cdot v\rangle_{L^2(\Omega)}, \qquad \forall v \in V,$$

we see that $-\nabla\colon L^2(\Omega) \to V^*$ is the dual to the divergence operator. Using the Cauchy-sequence argument, we see that $\operatorname{im}(\nabla\cdot)$ is closed in $L^2(\Omega)$ and the closed range theorem applies. Thus, $\operatorname{im}(-\nabla)$ is closed in V^* and

$$\operatorname{im}(\nabla) = (V^0)^0 \cong V^\perp$$

is the polar set of the kernel V^0. This implies the statement that there is a p for every f. Uniqueness follows by the fact that the only differentiable functions on a connected domain with $\nabla p = 0$ are the constant functions, and by density of such functions in $L^2(\Omega)$. □

3.1.7 Corollary: Let Ω be connected. Then,

1. $\nabla\colon L^2_0(\Omega) \to V^0$ is an isomorphism
2. $\nabla\cdot\colon V^\perp \to L^2_0(\Omega)$ is an isomorphism

3.1.8 Theorem: Let $\Omega \subset \mathbb{R}^d$ be a Lipschitz-domain, $V = H^1_0(\Omega, \mathbb{R}^d)$ and $Q = L^2_0(\Omega)$. Then, there is a constant $\beta > 0$ depending only on the geometry of Ω such that

$$\inf_{q\in Q} \sup_{v\in V} \frac{(\nabla\cdot v, q)}{\|v\|_V \|q\|_Q} \ge \beta. \tag{3.1}$$

Furthermore, the problem finding $(u, p) \in V \times Q$ such that

$$a(u, v) + (\nabla\cdot v, p) + (\nabla\cdot u, q) = f(v) + g(q) \quad \forall v \in V, q \in Q, \tag{3.2}$$

has a unique solution for any right hand side $f \in V^*$ and $g \in \operatorname{im}(\nabla\cdot)$.

3.2 Stable discretizations

3.2.1. We begin by application of the generic theory of the previous chapter to the Stokes problem in order to obtain a generic error estimate based on the concrete choice of norms and a single assumption. Guided by this theorem, we spend the remaining part of this section exploring different options for the discrete spaces.

3.2.2 Theorem: Let $V = H_0^1(\Omega;\mathbb{R}^d)$ and $Q = L_0^2(\Omega)$. Let furthermore $V_h \subset V$ and $Q_h \subset Q$ be discrete subspaces such that there exists $\beta > 0$ independent of h such that

$$\inf_{q_h\in Q_h} \sup_{v_h\in V_h} \frac{(\nabla\cdot v_h, q_h)}{\|v_h\|_V \|q_h\|_Q} \geq \beta. \tag{3.3}$$

Then, the Galerkin approximation of (3.2) admits a unique solution $(u_h, p_h) \in V_h \times Q_h$ with the quasi-bestapproximation property

$$\begin{aligned}
\|u - u_h\|_1 &\leq c_1 \inf_{v_h\in V_h} \|u - v_h\|_1 + c_2 \inf_{q_h\in Q_h} \|p - q_h\|_0\\
\|p - p_h\|_1 &\leq c_3 \inf_{v_h\in V_h} \|u - v_h\|_1 + c_4 \inf_{q_h\in Q_h} \|p - q_h\|_0.
\end{aligned} \tag{3.4}$$

3.2.3 Corollary: Under the assumptions of Theorem 3.2.2, let there be in addition interpolation operators I_{V_h} and I_{Q_h} such that

$$\begin{aligned}
\|u - I_{V_h} u\|_1 &\leq ch^k |u|_{k+1}\\
\|p - I_{Q_h} p\|_0 &\leq ch^k |p|_k.
\end{aligned} \tag{3.5}$$

Then, there is a constant c independent of the approximation spaces such that

$$\begin{aligned}
\|u - u_h\|_1 &\leq ch^k \big(|u|_{k+1} + |p|_k\big)\\
\|p - p_h\|_1 &\leq ch^k \big(|u|_{k+1} + |p|_k\big).
\end{aligned} \tag{3.6}$$

3.2.4. We continue showing that the most natural discretizations in two dimensions are not inf-sup stable. This holds for the discretization using continuous linear or bilinear elements for both velocity components and the pressure as well as for continuous linear or bilinear velocity functions combined with piecewise constant pressure functions.

Example 3.2.5. We begin with a one-dimensional example. Piecewise linear velocity and piecewise linear pressure ($P_1 - P_1$). Both continuous. Then, $\nabla\cdot v_h$ is piecewise constant. Consequently, a pressure function which has zero mean value on each cell is in the kernel of B_h^T.

Example 3.2.6. Take a patch of four quadrilaterals or triangles meeting in a common vertex. Let Ω be the union of these grid cells. Choose linear and bilinear shape functions for V_h, respectively. Then, $\dim V_h = 2$, since we have one interior vertex with one basis function for each velocity component. Choose piecewise constant pressure functions. Dividing out the global constant, we conclude that $\dim Q_h = 3$. Thus, the statement

$$\forall q_h \in Q_h\ \exists v_h \in V_h : \quad \|v_h\|_1 = \|q_h\|_0 \ \wedge\ b(v_h, q_h) \geq \beta \|q_h\|^2$$

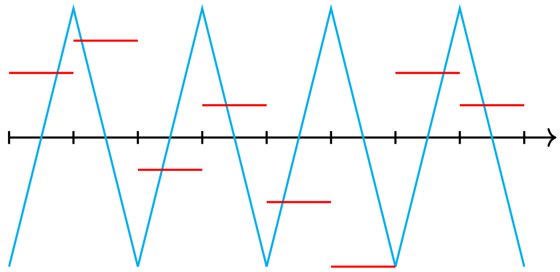


Figure 3.1: Piecewise linear pressure (—) and divergence (—) of piecewise linear velocity.

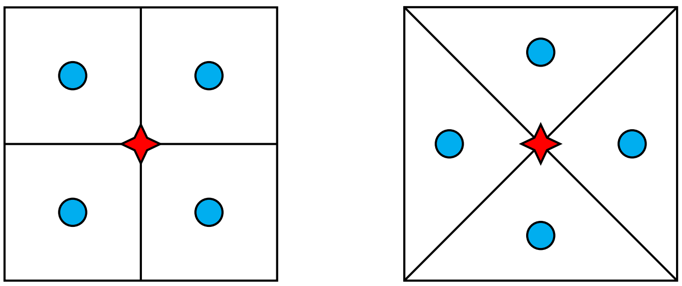


Figure 3.2: Very coarse meshes with Dirichlet boundary. Degrees of freedom for pressure (●) and for both velocity components(✦).

cannot hold true. Therefore, the inf-sup condition does not hold. In fact, $\ker(B_h) = \{0\}$.

Thus, we conclude that for this combination of shape function spaces, there is a mesh such that they are not suited for the approximation of the Stokes problem. But, this may be a problem of a mesh with too few cells. In fact, asymptotically, a triangular mesh contains twice as many vertices as cells, a quadrilateral mesh as many. Therefore, $\dim V_h > \dim Q_h$ as soon as the mesh is sufficiently fine. Will this be sufficient?

3.2.7 Problem: The domain $\Omega = [0,1]^2$ is decomposed into $N \times N$ congruent squares where each of them is again divided into two triangles. The decomposition $\mathbb{T}_h$ is given by these triangles.
We again choose piecewise linear ansatz functions for the velocity for V_h (vanishing on $\partial\Omega$) and piecewise constant ansatz functions for Q_h.
Is there a N and an orientation of the triangles such that $V_h \times Q_h$ is inf-sup stable?

3.2.8 Problem: Let $\Omega = (0,1)^2$ be the unit square and let the mesh consist of Cartesian squares of side length $1/n$. Choose $V_h \subset V$ based on bilinear shape functions. Show that the piecewise constant pressure function $p_c = \pm 1$ in a checkerboard fashion is in the kernel of B_h^T, that is

$$b(v_h, p_c) = 0 \quad \forall v_h \in V_h.$$

3.2.1 Bubble stabilization and the MINI element

3.2.9 Definition: A simplex $T \in \mathbb{R}^d$ with vertices $x_0, \dots, x_d$ is described by a set of $d+1$ **barycentric coordinates** $\lambda_0, \dots, \lambda_d$ such that

$$0 \le \lambda_i(x) \le 1 \qquad i = 0, \dots, d; \quad x \in T \tag{3.7}$$

$$\lambda_i(x_j) = \delta_{ij} \qquad i, j = 0, \dots, d \tag{3.8}$$

$$\sum \lambda_i(x) = 1. \tag{3.9}$$

Remark 3.2.10. The functions $\lambda_i(x)$ are the shape functions of the linear P_1 element on T. They allow us to define basis functions on the cell T without use of a reference element $\widehat{T}$.

Note that $\lambda_i \equiv 0$ on the face opposite to the vertex x_i.

Example 3.2.11. We can use barycentric coordinates to define shape functions on simplicial meshes easily, as in Table 3.1.

3.2.12 Notation: We denote by

$$H_h^k(\mathcal{P}) = \left\{ v \in H^k(\Omega) \middle| v_{|T} \in \mathcal{P} \ \forall T \in \mathbb{T}_h \right\} \tag{3.10}$$

the finite element space which is based on the shape function space $\mathcal{P}$, the mesh $\mathbb{T}_h$ and is a subspace of $H^k(\Omega)$. Examples are the continuous spaces of piecewise polynomials or tensor product polynomials of degree k

$$H_h^1(\mathbb{P}_k) \qquad H_h^1(\mathbb{Q}_k),$$

and the discontinuous spaces

$$H_h^0(\mathbb{P}_k) \qquad H_h^0(\mathbb{Q}_k).$$

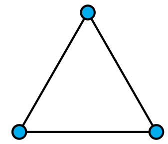
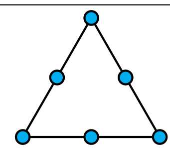
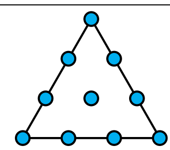
Degrees of freedom	Shape functions
	$\varphi_i = \lambda_i, \quad i = 0,1,2$
	$\varphi_{ii} = 2\lambda_i^2 - \lambda_i, \qquad i = 0,1,2$ $\varphi_{ij} = 4\lambda_i\lambda_j \qquad j \neq i$
	$\varphi_{iii} = \frac{1}{2}\lambda_i(3\lambda_i - 1)(3\lambda_i - 2) \quad i = 0,1,2$ $\varphi_{ij} = \frac{9}{2}\lambda_i\lambda_j(3\lambda_j - 1) \qquad j \neq i$ $\varphi_0 = 27\lambda_0\lambda_1\lambda_2$

Table 3.1: Degrees of freedom and shape functions of simplicial elements in terms of barycentric coordinates

3.2.13 Definition: An H^1**-bubble function** on a mesh cell T is a function $b \in H_0^1(T)$. A **bubble space** b_T on T is a discrete vector space of such bubble functions. We denote the space of bubble functions on the mesh $\mathbb{T}_h$ by

$$B_h(b_T) = \left\{ v \in H^1(\Omega) \middle| v_{|_T} \in b_T \ \forall T \in \mathbb{T}_h \right\}.$$

If there is no confusion about the local bubble space b_T, we also write just B_h.

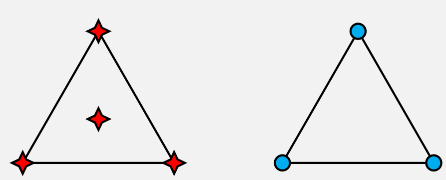
Example 3.2.14. A bubble function on a triangle T is easily defined by

$$b_3 = b_{3,T} = \lambda_0\lambda_1\lambda_2. \tag{3.11}$$

3.2.15 Definition: The **MINI element** consists of the spaces

$$V_h = \left(H_h^1(\mathbb{P}_1) \oplus B_h(b_3)\right)^2 \cap V, \qquad Q_h = H_h^1(\mathbb{P}_1) \cap Q. \tag{3.12}$$

Its degrees of freedom are:



3.2.16. We will show now that the MINI element is indeed inf-sup stable. To this end, we construct the Fortin projection according to Lemma 2.4.12. Since the construction of such a projection operator turns out a bit complicated, we first introduce a construction principle, which will help us in our further analysis. The idea of this principle is separating the interpolation into V_h from the preservation of the divergence.

3.2.17 Lemma: Let there be operators $\Pi_1, \Pi_2 \colon V \to V_h$ such that

$$\begin{aligned} \|\Pi_1 v\|_V &\le c_1 \|v\|_V && \forall v \in V, && (3.13)\\ \|\Pi_2(\mathbb{I} - \Pi_1) v\|_V &\le c_2 \|v\|_V && \forall v \in V, && (3.14)\\ b(v - \Pi_2 v, q_h) &= 0 && \forall v \in V,\ q_h \in Q_h, && (3.15) \end{aligned}$$

with constants c_1 and c_2 independent of the discretization parameter h. Then, the operator

$$\Pi_h = \Pi_1 + \Pi_2 - \Pi_2 \Pi_1 \qquad (3.16)$$

is a Fortin projection, that is, it is bounded on V and

$$b(v - \Pi_h v, q_h) = 0 \qquad \forall q_h \in Q_h.$$

Proof. Boundedness of Π_h is obvious, such that we only focus on preservation of the kernel ob B:

$$\begin{aligned} b(v - \Pi_h v, q_h) &= b(v - \Pi_1 v - \Pi_2 v + \Pi_2 \Pi_1 v, q_h) \\ &= b(v - \Pi_2 v, q_h) - b(\Pi_1 v - \Pi_2 \Pi_1 v, q_h) = 0 - 0 = 0. \end{aligned}$$

□

3.2.18 Assumption: There exists an H^1-stable interpolation operator $I_h : V \to V_h$ such that for each cell $T \in \mathbb{T}_h$ there holds for $m = 0, 1$

$$|v - I_h v|_{m,T} \le c \sum_{T' \cap T \neq \emptyset} h_{T'}^{1-m} |v|_{1,T'}, \qquad (3.17)$$

with a constant c independent of the mesh parameter h.

Remark 3.2.19. The interpolation operators of Clément, Scott and Zhang, Schöberl or Ern and Guermond fullfil these assumptions.

3.2.20 Definition: A family of meshes $\{\mathbb{T}_h\}$ is called **locally quasi-uniform**, if there is a constant c such that

$$\forall h\, \forall T, T' \in \mathbb{T}_h \quad T \cap T' \neq \emptyset \Rightarrow h_T \leq c h_{T'}. \tag{3.18}$$

3.2.21 Assumption: We assume of all families of meshes that they are shape regular and locally quasi-uniform, such that with Assumption 3.2.18 there holds for $m = 0, 1$

$$|v - I_h v|_{m,T} \leq c h_T^{1-m} |v|_{1,\Omega_T}, \tag{3.19}$$

where Ω_T is the union of all cells with nonempty intersection with T.

3.2.22 Theorem: Under Assumption 3.2.21, the MINI element is inf-sup stable.

Proof. We construct a Fortin projection by choosing $\Pi_1 = I_h$, where $I_h : V \to \big(H_h^1(\mathbb{P}_1)\big)^2$ is an H^1-stable interpolation operator into the standard linear finite element space. Now, we construct $\Pi_2 : V \to \big(B_h\big)^2$ such that for all $q_h \in Q_h$

$$\int_\Omega \nabla\cdot(\Pi_2 v - v) q_h \,\mathrm{d}x = \int_\Omega (v - \Pi_2 v) \cdot \nabla q_h \,\mathrm{d}x = 0.$$

Indeed, $\Pi_2 v$ can be chosen on each cell. Since ∇q_h is constant on a cell T, we choose

$$\int_T \Pi_2 v_i \,\mathrm{d}x = \alpha_{T,i} \int_T b_{3,T} \,\mathrm{d}x = \int_T v_i \,\mathrm{d}x,$$

where $i = 1, 2$ enumerates the velocity components. This is possible, since the mean value of b_3 is strictly positive. Assuming shape regularity, we can use the inverse estimate for b_3 to obtain

$$\|\Pi_2 v\|_{1,T} \leq c h_T^{-1} \|\Pi_2 v\|_{0,T} \leq c h_T^{-1} \|v\|_{0,T}.$$

Finally, we use the estimates for I_h to obtain

$$\|\Pi_2(\mathbb{I} - \Pi_1) v\|_{1,T} \leq c h_T^{-1} \|v - I_h v\|_{0,T} \leq c |v|_{1,\Omega_T}.$$

Since the number of intersecting of cells of shape regular meshes is bounded, the final term is bounded by $\|v\|_{1,\Omega}$. □

3.2.23 Notation: We use the abbreviation

$$(f,g)_{\mathbb{T}_h} = \sum_{T\in\mathbb{T}_h} (f,g)_T, \tag{3.20}$$

for so called **broken bilinear forms**, where instead of integrating over the union of subsets, we sume the integrals.

3.2.24 Lemma: The discretization of the Stokes problem (3.2) with the MINI element is equivalent to solving

$$\begin{aligned}(\nabla u,\nabla v) + (\nabla\cdot v, p) + (\nabla\cdot u, q) - (c_T\nabla p, \nabla q)_{\mathbb{T}_h}\\ = f(v) + g(q) + (c_T f_T, \nabla q)_{\mathbb{T}_h}\end{aligned} \tag{3.21}$$

with standard, continuous linear finite elements for velocity and pressure. Here,

$$f_T = \int_T f\,\mathrm{d}x, \qquad c_T = \frac{(b_3,1)_T}{\|\nabla b_3\|_T^2}.$$

Proof. Let $V_h^1 = H_h^1(\mathbb{P}_1)^2$ be the linear, vector-valued velocity space and $V_h^b = B_h(b_3)^2$ the bubble function space, such that the MINI element space is

$$V_h = V_h^1 \oplus V_h^b.$$

Accordingly, we split the solution with the MINI element into $u_h = u_h^1 + u_h^b$. By integration by parts, we obtain for the cubic bubble b_3

$$(\nabla v, \nabla b_{3,T})_T = (-\Delta v, b_{3,T})_T = 0 \qquad \forall v \in \mathbb{P}_1,$$

such that

$$(\nabla v, \nabla b) = (\nabla b, \nabla v) = 0 \qquad \forall v \in V_h^1, b\in V_h^b.$$

Thus, testing (3.2) with $v_h \in V_h^1$ yields

$$\big(\nabla u_h^1, \nabla v_h\big) + (\nabla\cdot v_h, p_h) = f(v_h) \qquad \forall v_h \in V_h^1. \tag{3.22}$$

Testing the same equation with $v \in V_h^b$, we obtain

$$\big(\nabla u_h^b, \nabla v_h\big) = f(v_h) - (\nabla\cdot v_h, p_h) = f(v_h) + (v_h, \nabla p_h)_{\mathbb{T}_h}. \tag{3.23}$$

Choosing more specifically v_h as the bubble function $b_{3,T}$ of the cell T for each vector component yields

$$\mu_T^{(i)} = \frac{1}{\|\nabla b_3\|_T^2}(f + \partial_i p_h, b_3)_T, \tag{3.24}$$

where $\mu_T^{(i)}$ is the coefficient in front of the basis function $b_{3,T}$ on cell T in the basis representation of $u_h^{(i)}$, such that

$$u_h^b = \sum_{T\in\mathbb{T}_h} \begin{pmatrix} \mu_T^{(1)} b_{3,T} \\ \mu_T^{(2)} b_{3,T} \end{pmatrix} \tag{3.25}$$

Testing the Stokes equations with $q_h \in Q_h$, we obtain the divergence equation

$$\left(\nabla\cdot u_h^1 + \nabla\cdot u_h^b, q_h\right) = \left(\nabla\cdot u_h^1, q_h\right) - \left(u_h^b, \nabla q_h\right)_{\mathbb{T}_h} = g(q_h). \tag{3.26}$$

Using (3.24), (3.25) and using $f_T^{(i)} = \left(f^{(i)}, 1\right)_T$ yields

$$\begin{aligned}\left(u_h^b, \nabla q_h\right)_{\mathbb{T}_h} &= \sum_{T\in\mathbb{T}_h} \frac{1}{\|\nabla b_3\|_T^2} \sum_{i=1,2} \left(f_T^{(i)} + \partial_i p_h, b_3(T)\partial_i q_h\right)_T \\ &= \sum_{T\in\mathbb{T}_h} \frac{(b_3,1)_T}{\|\nabla b_3\|_T^2} (f_T + \nabla p_h, \nabla q_h)_T\end{aligned}$$

□

Remark 3.2.25. The constant c_T in the previous lemma was computed by the formula

$$c_T = \frac{(b_3,1)_T}{\|\nabla b_3\|_T^2}.$$

This formula is complicated and we would rather like to avoid computing c_T for every mesh cell, since we have to evaluate integrals of cubic functions. On the other hand, the same constant c_T appears on the left and on the right of the modified equation (3.21). Therefore, we can replace both by a constant of similar size without affecting consistency or the characteristic properties of the equation. Therefore, we estimate

$$c_T = \frac{(b_3,1)_T}{\|\nabla b_3\|_T^2} \simeq \frac{\|b_3\|_T^2}{\|\nabla b_3\|_T^2} \simeq h_T^2, \tag{3.27}$$

where "$\simeq$" indicates equality up to a constant independent of h, but depending on the constant in shape regularity.

Remark 3.2.26. The method introduced in Lemma 3.2.24 is an example for a **stabilized method**, here in particular **pressure stabilization**. Such methods were particularly popular in the early decades of finite element computation, since they only involve simple shape function spaces. They are still widely used due to their simplicity. The method constructed this way is consistent, i. e. , the continuous solution (u, p) solves the discrete problem.

3.2.27 Problem: Show that the MINI element can be generalized to quadrilateral meshes. Design a bubble space b_Q of minimal tensor degree such that

$$V_h = \left(H_h^1(\mathbb{Q}_1) \oplus B_h(b_Q) \cap V\right)^2, \qquad Q_h = H_h^1(\mathbb{Q}_1) \cap Q.$$

Discuss extensions to tetrahedra and hexahedra in three dimensions.

3.2.28 Problem: By introducing barycentric coordinates $\lambda_0, \dots, \lambda_3$ for a tetrahedron $T \subset \mathbb{R}^3$ and the quartic bubble

$$b_{4,T} = \lambda_0 \lambda_1 \lambda_2 \lambda_3, \tag{3.28}$$

show that the MINI element has a natural generalization to three dimensional problems.

3.2.29. The reasoning behind the MINI element can be applied easily to pressure spaces of higher order. Take for instance the pair $P_k - P_k$, generalized from Example 3.2.5. There holds $\nabla \cdot v_h \in \mathbb{P}_{k-1}$ on each cell, and the term

$$\int_T \nabla \cdot v_h q_h \, \mathrm{d}x$$

does not control the function in $\hat{p}_T \in \mathbb{P}_k$ which is orthogonal to $\mathbb{P}_{k-1}$. The only function $p_h \in Q_h$ such that $p_{h|T} = \hat{p}_T$ for each cell $T \in \mathbb{T}_h$ may be zero or not, depending on the mesh geometry. Thus, the element is not stable on arbitrary shape regular meshes. But, as we prove below, the same enrichment process by bubble functions can be employed for its stabilization.

3.2.30 Definition: With any pressure space Q_h we associate the **bubble space**

$$B_h(b_T \nabla Q_h) = \left\{ v \in V \,\middle|\, \exists q_h \in Q_h \colon v_{|T} = b_T \nabla q_h \right\}. \tag{3.29}$$

Here, b_T is a bubble function on T like the cubic buble $b_{3,T}$ of a triangle, the quartic bubble $b_{4,T}$, the biquadratic bubble $b_{2^2,T}$ of a quadrilateral or the triquadratic bubble $b_{2^3,T}$ of a hexahedron.
We also define the cell bubble space

$$B_T(\nabla Q_h) = \left\{ v \in L^2(T) \,\middle|\, \exists q_h \in Q_h \colon v = b_T \nabla q_{h|T} \right\}. \tag{3.30}$$

3.2.31 Theorem: Assume that the pair $V_h \times Q_h$ is chosen such that there is an H^1-stable interpolation operator according to Assumption 3.2.18, such that $Q_h \subset C^0(\Omega)$, piecewise differentiable, and such that

$$B_h(b_T \nabla Q_h) \subset V_h. \tag{3.31}$$

Then, the pair $V_h \times Q_h$ is inf-sup stable.

Proof. We construct the Fortin projection by Lemma 3.2.17 choosing Π_1 as the H^1-stable interpolation operator. The operator Π_2 is constructed cell-wise such that $\Pi_2 \colon H^1(T) \to B_T(\nabla Q_h)$ fulfills

$$\int_T (\Pi_2 u - u) \cdot \nabla q = 0, \qquad \forall q \in Q_{h|T}. \tag{3.32}$$

Clearly, the dimension of $B_T(\nabla Q_h)$ equals the dimension of $Q_{h|T}$. Then, since the bubble functions are strictly positive inside T, equation (3.32) defines $\Pi_2 u$ uniquely. It remains to show the H^1-stability of $\Pi_2(\mathbb{I} - \Pi_1)$, which is done by the standard scaling argument

$$|\Pi_2 v|_{1,T} = |\widehat{\Pi_2 v}|_{1,\widehat{T}} \le c\|\widehat{v}\|_{1,\widehat{T}} \le c\big(h_T^{-1}\|v\|_{0,T} + |v|_{1,T}\big).$$

□

3.2.32 Corollary: Let $Q_h \subset Q$ be continuous and cell-wise differentiable. If

$$H_h^1(\mathbb{P}_1)^d \oplus B_h(b_T \nabla Q_h) \subset V_h \subset V,$$

then the pair $V_h \times Q_h$ is inf-sup stable. The same holds on quadrilateral and hexahedral meshes replacing $\mathbb{P}_1$ by $\mathbb{Q}_1$.

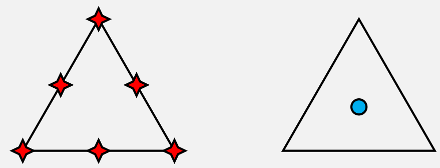
3.2.2 Elements with discontinuous pressure

3.2.33. In this section, we consider a second stable element, which like the MINI element is not so much of practical use, but exhibits typical properties of the analysis of finite element spaces for the Stokes problem.

3.2.34 Definition: The $\mathbf{P_2 - P_0}$ **element** on triangles consists of the finite element spaces

$$V_h = H^1_h(\mathbb{P}_2)^2 \cap V, \qquad Q_h = H^0_h(\mathbb{P}_0) \cap Q. \tag{3.33}$$

Its degrees of freedom are:



3.2.35 Lemma: The $P_2 - P_0$ element is inf-sup stable.

Proof. We again prove stability by constructing a Fortin projection using the two step algorithm of Lemma 3.2.17. Again, we choose for Π_1 an H^1-stable interpolation according to Assumption 3.2.18. It remains therefore to construct Π_2. First, since $q_{h|T}$ is constant on each cell $T \in \mathbb{T}_h$, we can apply the Gauss theorem to the divergence condition to obtain

$$\int_T \nabla\cdot(u - \Pi_2 u)\,\mathrm{d}x = \int_{\partial T} (u - \Pi_2 u)\cdot n\,\mathrm{d}s. \tag{3.34}$$

Hence, the following interpolation conditions on each cell T define a divergence preserving operator Π_2:

$$\Pi_2 u(x) = 0 \qquad \forall x \text{ is vertex of } T \tag{3.35}$$

$$\int_E \Pi_2 u\,\mathrm{d}s = \int_E u\,\mathrm{d}s \qquad \forall E \text{ is edge of } T \tag{3.36}$$

This is true, since (3.36) implies the right hand side of (3.34). It remains to show the H^1-stability of $\Pi_2(\mathbb{I} - \Pi_1)$. Let us first observe that the interpolation operator only involves edge integrals of u, which are well-defined on H^1. Thus, we have by the standard scaling argument

$$|\Pi_2 v|_{1,T} = |\widehat{\Pi_2 v}|_{1,\widehat{T}} \le c\|\widehat{v}\|_{1,\widehat{T}} \le c\big(h_T^{-1}\|v\|_{0,T} + |v|_{1,T}\big).$$

Entering $v = u - \Pi_1 u$ and the estimates (3.19) of Assumption 3.2.21, we obtain

$$\|\Pi_2(\mathbb{I} - \Pi_1)u\|_1^2 = \sum_{T\in\mathbb{T}_h} \|\Pi_2(\mathbb{I} - \Pi_1)u\|_{1,T}^2 \le c\|u\|_1^2$$

□

Remark 3.2.36. The proof shows, that from a mathematical point of view degrees of freedom on edges are more reasonably defined by integrals along the edge than by values in the mid points. This is something, we will encounter again and again. Nevertheless, we will not change the cartoons for the degrees of freedom and just note that a degree of freedom on an edge, while drawn as a point, may be an integral value.

3.2.37 Theorem: Let $(u, p) \in V \times Q$ be a solution to the Stokes problem and let the pair $(u_h, p_h) \in V_h \times Q_h$ be the approximation on a mesh $\mathbb{T}_h$ of mesh size h with the $P_2 - P_0$ element of Definition 3.2.34. Then, we have the error estimate

$$\|u - u_h\|_1 + \|p - p_h\|_0 \le c\big(h^2 |u|_3 + h|p|_1\big).$$

Remark 3.2.38. While this theorem is optimal with respect to our analysis, it is not optimal with respect to the approximation properties of V_h.

Remark 3.2.39. Let us review the construction principles behind the MINI element and the $P_2 - P_0$ element. The uncontrolled pressure modes in $\ker (B)_h^T$ of the $P_1 - P_1$ element in Example 3.2.5 were those pressures with alternating signs at neighboring vertices, such that the mean value of p_h is zero on each cell. Therefore, p_h is orthogonal to the constant derivatives of the linear velocity space. Thus, we add a local function on each cell with nonconstant gradient, and the mean value of the pressure can be controlled.

The kernel of B_h^T for the element $H^1(\mathbb{P}_1)^2 - H^0(\mathbb{P}_0)$ on the other hand contains functions that are constant on each cell, but jump over cell boundaries. By integration by parts, we have

$$\int_T \nabla \cdot b_T q_h \,\mathrm{d}x = -\int_T b_T \cdot \nabla q_h + \int_{\partial T} b_T q_h \,\mathrm{d}s = 0.$$

Hence, no kind of bubble function helps controlling the jump of p_h over an edge. Instead, we introduce a degree of freedom on the edge. Integrating by parts on two neighboring cells T_1 and T_2, we obtain on the common edge E_{12} a term of the form

$$\int_{E_{12}} \left[u \cdot n_1 q_1 + u \cdot n_2 q_2\right],$$

which by the continuity of $u \cdot n$ translates to

$$\int_{E_{12}} u \cdot n_1 (q_1 - q_2). \tag{3.37}$$

Thus, we can use the interpolation operator Π_2 to obtain a function u such that

$$\int_{E_{12}} u \cdot n_1 \,\mathrm{d}s = (q_1 - q_2),$$

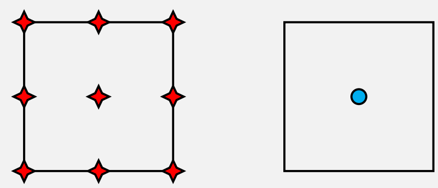
such that

$$(\nabla\cdot u, q_h) = \int_{E_{12}} |q_1 - q_2|^2 \,\mathrm{d}s + \text{ other terms.}$$

3.2.40 Problem: Show that the quadrilateral element

$$V_h = H^1_h(\mathbb{Q}_2)^2 \cap V, \qquad Q_h = H^0_h(\mathbb{P}_0) \cap Q, \tag{3.38}$$

with degrees of freedom



is inf-sup stable. Does the proof translate to the $P_2 - P_0$ element on tetrahedra or the $Q_2 - P_0$ element on hexahedra?

3.2.41. In spite of our remarks above, there is a generalization of the $P_2 - P_0$ element involving bubble functions. We will discuss it in an abstract theorem first and then derive a family of inf-sup stable pairs.

3.2.42 Lemma: Given a space $Q_h \subset Q$ possibly discontinuous, choose $V_h \subset V$ such that

$$B_h(b_T \nabla Q_h)^d \subset V_h.$$

If there is an operator Π_1 such that

$$\begin{aligned} \|\Pi_1 v\|_V &\le \|v\|_V && \forall v \in V, \\ \int_T \nabla\cdot(v - \Pi_1 v)\,\mathrm{d}x &= 0 && \forall v \in V, T \in \mathbb{T}_h, \end{aligned}$$

then the pair $V_h \times Q_h$ is inf-sup stable.

Proof. We construct the Fortin projection using Π_1 and define Π_2 only on $V^0 = \ker(\nabla\cdot)$. This is sufficient, since for any $v \in V$ there holds $v - \Pi_1 v \in V^0$. Therefore, define cell-wise $\Pi_2 \colon V^0_{|T} \to B_T(b_T \nabla Q_h)$ by the conditions

$$\int_T \nabla\cdot(\Pi_2 v - v) q_h \,\mathrm{d}x = 0 \qquad \forall q_h \in Q_{h|T}. \tag{3.39}$$

By this condition, $\Pi_2 v$ is divergence free itself. Note that by the Gauss theorem, the divergence of a bubble function has always zero mean. Therefore, we have unique solvability and $\Pi_h v$ is well defined. It remains to apply the standard scaling argument to prove

$$\|\Pi_2 v\|_1 \le c\|v\|_1.$$

□

Remark 3.2.43. The divergence condition in the previous lemma is different from the condition on Fortin projections, since it only involves piecewise constant pressure. Therefore, the lemma in effect splits the pressure space into a piecewise constant part and its complement. Then, the pressure in the complement is controlled by the bubble functions. It still remains to guarantee the existence of the operator Π_1. In one case, we have verified the existence of such an operator: the Fortin operator for the $P_2 - P_0$ element. Therefore, we have

3.2.44 Corollary: Let $Q_h \subset Q$ be a space of piecewise differentiable functions. If for $V_h \subset V$ holds

$$H^1_h(\mathbb{P}_2)^2 \oplus B_h(b_T \nabla Q_h)^2 \subset V_h,$$

then the pair $V_h \times Q_h$ is inf-sup stable.

3.2.45 Corollary: Let the space dimension be $d = 2$ and $k \ge 2$. Then, the spaces

$$V_h = H^1_h(\mathbb{P}_k)^2 \cap V, \qquad Q_k = H^0_h(\mathbb{P}_{k-2}) \cap Q, \tag{3.40}$$

form an inf-sup stable pair.

Proof. For $k = 2$, this is the $P_2 - P_0$ element. For $k > 2$, we have for all $q_h \in Q_h$ on every cell $\nabla q_h \in \mathbb{P}_{k-3}$. Therefore, $(b_{3,T} q_h)_{|T} \in \mathbb{P}_k$. □

3.2.46. Studying the proof of Lemma 3.2.35 in more detail, we can find a much more general result with much weaker assumptions. Indeed, we only need to be able to have a single degree of freedom on each edge or face which allows to adjust the average normal velocity over this edge of face. We summarize this in:

3.2.47 Theorem: Let $Q_h \subset Q$ and let $V_h \subset V$ be such that there is an H^1-stable interpolation according to Assumption 3.2.18. Furthermore, let there be a degree of freedom on each edge (face in 3D) controlling the average normal derivative of $u \in V_h$ on this face. Finally, let V_h contain the bubble space for ∇Q_h,

$$B_h(b_T \nabla Q_h)^d \subset V_h.$$

Then, the pair $V_h \times Q_h$ is inf-sup stable.

Proof. With the assumptions made, it is sufficient to construct the operator Π_1 in Lemma 3.2.42. Then, we can apply this lemma and the result is proven. Going back to (3.34), we see that the interpolation condition (3.36) is more than needed.

Now, let $\{\mathcal{N}_{T,i}\}$ be the n_T node values for the discrete velocity space $V_T = V_{h|T}$ on the cell T. For convenience, let them be ordered in a way, that the first ones control the normal derivatives of u_h on the faces of the cell, that is,

$$\mathcal{N}_{T,i} = \int_{F_i} u \cdot n \,\mathrm{d}s \qquad i = 1, \dots, n_F, \tag{3.41}$$

where n_F is the number of faces per cell. Given the H^1-stable interpolation operator I_h, define Π_2 cell-wise such that

$$\mathcal{N}_{T,i}(\Pi_2 u) = \mathcal{N}_{T,i}(u) \qquad i = 1, \dots, n_F \tag{3.42}$$

$$\mathcal{N}_{T,i}(\Pi_2 u) = 0 \qquad i = n_F + 1, \dots, n_T. \tag{3.43}$$

Choosing the basis on V_T which is dual to $\{\mathcal{N}_{T,i}\}$, we see that this indeed implies

$$0 = \int_{\partial T} (\Pi_2 u - u) \,\mathrm{d}s = \int_T \nabla \cdot (\Pi_2 u - u) \,\mathrm{d}x.$$

Therefore, $\Pi_1 = I_h + \Pi_2(\mathbb{I} - I_h)$ is divergence preserving. Boundedness follows by the standard scaling argument. □

3.2.48 Corollary: The finite-element pair $V_h \times Q_h$ with

$$V_h = H_h^1(\mathbb{Q}_k)^2 \cap V, \qquad Q_h = H_h^0(\mathbb{P}_{k-1}) \cap Q,$$

called the $Q_k - P_{k-1}$ element is inf-sup stable for any $k \geq 2$.

Proof. First, we prove that $B_h(b_T \nabla Q_h)^d \subset V_h$. To this end, we note that on each cell, we have that the gradient of a discrete pressure q_h restricted to this cell is in $\mathbb{P}_{k-2} \subset \mathbb{Q}_{k-2}$. The bubble function b_T is in $\mathbb{Q}_2$, such that $b_T \nabla q_T \in \mathbb{Q}_k$, which was to be proven.

For the assumption on the degrees of freedom, we refer to the following definition. Once the degrees of freedom for each velocity component are determined by this definition, we can simply select the normal component on Cartesian meshes. On meshes with straight interfaces, it is clear that we can choose a linear combination of the components of u splitting into normal and tangential and thus get the desired result. In general, we refer to the Piola transformation in the next chapter. □

3.2.49 Definition: The shape function space $\mathbb{P}_k$ on the reference element $\widehat{T} = [-1, 1]$ in one dimension can be split into

$$\mathbb{P}_k^0 \oplus \overline{\mathbb{P}_k}, \tag{3.44}$$

where $\mathbb{P}_k^0 = \mathbb{P}_k \cap H_0^1(\widehat{T})$. We choose an orthogonal basis for $\mathbb{P}_k^0$ with respect to the H_0^1-inner product $\langle p, q \rangle = \int p' q' \, \mathrm{d}x$ by

$$\varphi_i(x) = \int_{-1}^{x} L_i(t) \, \mathrm{d}t \quad i = 1, \dots, k-1, \tag{3.45}$$

where L_i is the Legendre polynomial of degree i. The two basis functions for $\overline{\mathbb{P}_k}$ are chosen such that

$$\begin{aligned} \varphi_0(-1) &= 1, & \langle \varphi_0, \varphi_i \rangle &= 0 \quad i = 1, \dots, k-1, \\ \varphi_k(1) &= 1, & \langle \varphi_k, \varphi_i \rangle &= 0 \quad i = 1, \dots, k-1. \end{aligned}$$

3.2.50 Lemma: The degrees of freedom

$$\begin{aligned} \mathcal{N}_0(\varphi) &= \varphi(-1), & \mathcal{N}_k(\varphi) &= \varphi(1), \\ \mathcal{N}_i(\varphi) &= \frac{1}{\int {\varphi_i'}^2} \int_{-1}^{1} \varphi' \varphi_i' \, \mathrm{d}x & i &= 1, \dots, k-1, \end{aligned} \tag{3.46}$$

are the dual basis for the basis described in Definition 3.2.49.

Proof. The proof is left to the reader. □

3.2.51 Definition: Let $\widehat{T} = [-1,1]^d$ be the reference square in $\mathbb{R}^d$. We define the space

$$\mathbb{Q}_k^0 = \mathbb{Q}_k \cap H_0^1(\widehat{T}). \tag{3.47}$$

A basis for $\mathbb{Q}_k^0$ consists of the functions

$$\varphi_{i_1 \cdots i_d}(x) = \varphi_{i_1}(x_1) \cdots \varphi_{i_d}(x_d), \tag{3.48}$$

where $i_j = 1, \dots, k-1$. For a tensor product mesh cell T, the space $\mathbb{Q}_k^0(T) = \mathbb{Q}_k(T) \cap H_0^1(T)$ is defined through $\mathbb{Q}_h^0$ by mapping.

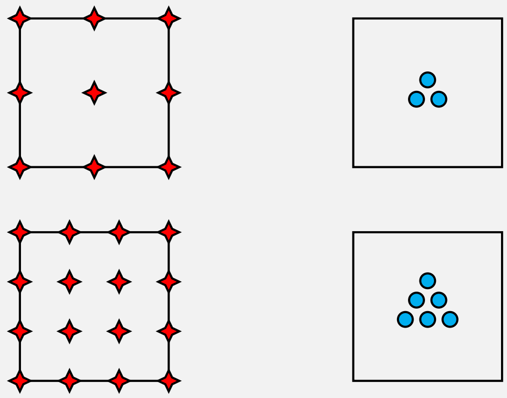
3.2.52 Definition: The moment degrees of freedom of the $\mathbb{Q}_k$ element are defined on the reference cell $\widehat{T}$ in two space dimensions as

$$\begin{aligned}
\mathcal{N}_{0,i}(u) &= u(x_i) & & x_i \text{ is vertex of } \widehat{T} \\
\mathcal{N}_{1,i,j}(u) &= \int_{E_i} u\varphi_j \,\mathrm{d}s & \varphi_j \in \mathbb{Q}_k^0(E_i) \quad & E_i \text{ is edge of } \widehat{T} \\
\mathcal{N}_{2,j}(u) &= \int_{\widehat{T}} u\varphi_j \,\mathrm{d}x & \varphi_j \in \mathbb{Q}_k^0(\widehat{T}). &
\end{aligned}$$

3.2.53 Definition: The moment degrees of freedom of the $\mathbb{Q}_k$ element are defined on the reference cell $\widehat{T}$ in three space dimensions as

$$\begin{aligned}
\mathcal{N}_{0,i}(u) &= u(x_i) & & x_i \text{ is vertex of } \widehat{T} \\
\mathcal{N}_{1,i,j}(u) &= \int_{E_i} u\varphi_j \,\mathrm{d}s & \varphi_j \in \mathbb{Q}_k^0(E_i) \quad & E_i \text{ is edge of } \widehat{T} \\
\mathcal{N}_{2,i,j}(u) &= \int_{F_i} u\varphi_j \,\mathrm{d}s & \varphi_j \in \mathbb{Q}_k^0(F_i) \quad & F_i \text{ is face of } \widehat{T} \\
\mathcal{N}_{3,j}(u) &= \int_{\widehat{T}} u\varphi_j \,\mathrm{d}x & \varphi_j \in \mathbb{Q}_k^0(\widehat{T}). &
\end{aligned}$$

3.2.54 Example: The first two members of the $Q_k - P_{k-1}$ family have the nodal representations



Remark 3.2.55. When we map the reference square to a quadrilateral mesh cell, this mapping may be affine for parallelograms or bilinear for general quadrilaterals. At some point, we have proven that the mapped $\mathbb{Q}_k$ space has optimal approximation properties, that is, approximation of order k in H^1 and of order $k+1$ in L^2. Such a thing has not been proven for a bilinearly mapped $\mathbb{P}_k$ element. And, unfortunately it is not true. We therefore have to distinguish between a mapped and an unmapped pressure space. In [ABF02], it is proven that the mapped polynomial space has worse approximation, in the worst case one order less than the unmapped.

3.2.3 The family of Hood-Taylor elements

3.2.56 Definition: The family of Hood-Taylor elements on simplices in dimension $d = 2, 3$ for polynomial degrees $k \geq 1$ consists of the pairs

$$V_h = H_h^1(\mathbb{P}_k)^d \cap V, \qquad Q_h = H_h^1(\mathbb{P}_{k-1}) \cap Q. \tag{3.49}$$

On quadrilaterals and hexahedra, it consists of the pairs

$$V_h = H_h^1(\mathbb{Q}_k)^d \cap V, \qquad Q_h = H_h^1(\mathbb{Q}_{k-1}) \cap Q. \tag{3.50}$$

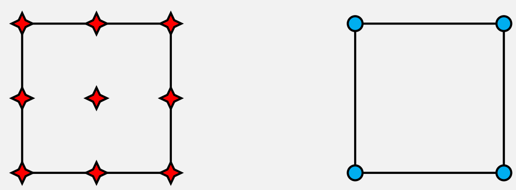
3.2.57 Example: The first two members of the Hood-Taylor family on triangles have the nodal representations

$\mathbb{P}_2 - \mathbb{P}_1$

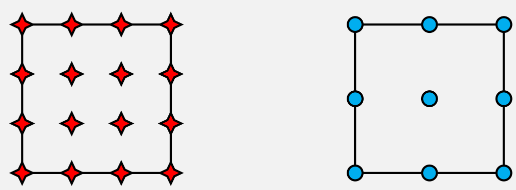
$\mathbb{P}_3 - \mathbb{P}_2$

3.2.58 Example: The first two members of the Hood-Taylor family on quadrilaterals have the nodal representations

$\mathbb{Q}_2 - \mathbb{Q}_1$



$\mathbb{Q}_3 - \mathbb{Q}_2$



3.2.59. The stable elements of the previous section featured discontinuous pressure spaces. Therefore, it was natural to split the analysis into cell-wise constant pressure and higher order. This is the basic technique behind Lemma 3.2.42 and Theorem 3.2.47. Here, the pressure is continuous, such that a cell-wise analysis is not feasible anymore. The solution is looking at patches, so called macro elements. The analysis is due to [Ste90] and consists of three parts: first, the covering of the domain with a macro element partitioning, then the local stability on each macro element with respect to an auxiliary norm on the pressure space, and finally the application of an abstract argument known as Verfürth's trick [Ver84].

3.2.60 Definition: A **macro element** $M \subset \mathbb{T}_h$ is a union of cells $T_i \in \mathbb{T}_h$ connected by their boundary faces. Given the mappings $\Phi_T : \widehat{T} \to T$, there is a reference macro element $\widehat{M}$ and a mapping $\Phi_M : \widehat{M} \to M$ such that $\Phi_M(\widehat{M}) = M$. We say that M is equivalent to $\widehat{M}$.
We will use the symbol M for the set of cells constituting a macro element as well as for the subset of Ω covered by their union.

3.2.61 Problem: Suggest reference macro elements $\widehat{M}$ for the following situations:

1. Two triangles sharing an edge
2. Two quadrilaterals sharing an edge
3. Two hexahedra sharing a face
4. Three quadrialterals at the edge between a coarse cell and a refined cell

Based on the mappings Φ_T from the reference cell to the actual mesh cells, define a continuously invertible mapping from $\Psi_M : \widehat{M} \to M$ (it is sufficent to describe the construction without writing a closed formula). Argue, that under the assumption of shape regularity, all macro elements M with the same connectivity between their cells are equivalent to the same reference macro element $\widehat{M}$.

3.2.62 Definition: For a macro element M, we introduce the spaces

$$V_M = \left\{ u \in H_0^1(M;\mathbb{R}^d) \middle| \exists v_h \in V_h : u = v_{h|M} \right\}, \tag{3.51}$$

$$Q_M = \left\{ p \in L^2(M) \middle| \exists q_h \in Q_h : p = q_{h|M} \right\}, \tag{3.52}$$

the kernel of the discrete, local gradient operator

$$\ker\left(B_M^T\right) = \left\{ q \in Q_M \middle| \forall v \in V_M : (\nabla\cdot v, q) = 0 \right\}. \tag{3.53}$$

3.2.63 Definition: Let $\mathbb{F}_h^i$ be the set of interior faces of the mesh $\mathbb{T}_h$. We introduce the alternative norm on Q_h defined by

$$\|p\|_h^2 = \sum_{T\in\mathbb{T}_h} h_T^2 \|\nabla p\|_T^2 + \sum_{F\in\mathbb{F}_h^i} \|[\![p]\!]\|_F^2, \tag{3.54}$$

where for a face F between two cells T_1 and T_2 we define the **jump operator**

$$[\![p]\!] = p_1 - p_2. \tag{3.55}$$

Remark 3.2.64. For continuous pressure spaces, the norm $\|p\|_h$ is simply the norm of the gradient taken in the interior of all cells.

3.2.65 Definition: On each macro element M, let $\mathbb{F}_M^i$ be the set of interior faces of M. We define the seminorm

$$|p|_M = \sum_{T\in M} h_T^2 \|\nabla p\|_T^2 + \sum_{F\in\mathbb{F}_M^i} \|[\![p]\!]\|_F^2. \tag{3.56}$$

It is not a norm because Q_M contains constant functions.

3.2.66 Lemma: Assume that there is a covering of Ω by macro elements such that every interior face $F \in \mathbb{F}_h^i$ is an interior face of one macro element and each cell $T \in \mathbb{T}_h$ is an element of not more than n_O macro elements. Then, the local stability estimate

$$\sup_{v\in V_M} \frac{(\nabla\cdot v, q)_M}{|v|_{1,M}} \ge \widehat{\beta}|q|_M \qquad \forall q \in Q_M, \tag{3.57}$$

implies the stability estimate

$$\sup_{v\in V_h} \frac{(\nabla\cdot v, q)}{\|v\|_1} \ge \beta\|q\|_h \qquad \forall q \in Q_h, \tag{3.58}$$

with a constant β independent of the mesh size h.

Proof. For arbitrarily chosen $q \in Q_h$, choose for each M according to assumption (3.57) functions $v_M \in V_M$ with $\|v_M\|_1 \le |q|_M$ such that

$$(\nabla\cdot v_M, q) = (\nabla\cdot v_M, q)_M \ge \widehat{\beta}|q|_M^2.$$

Define $v = \sum v_M$. Since every face is an interior face of a macro element, every cell is element of at least one macro element. Hence,

$$(\nabla\cdot v, q) = \sum_M (\nabla\cdot v_M, q) \geq \widehat{\beta} \sum_M |q|_M^2 \geq \widehat{\beta}\|q\|_h^2.$$

Furthermore, there holds by Poincaré inequality

$$c_S\|v\|_1 \leq |v|_1 \leq \sum_M |v_M|_1 \leq \sum_M |q|_M \leq n_O\|q\|_h.$$

Thus, the estimate holds with

$$\beta = \frac{c_S\widehat{\beta}}{n_O}.$$

□

3.2.67 Lemma: Let $\{M\}$ with $M \subset \mathbb{T}_h$ be a set of macro elements equivalent to the same reference macro element $\widehat{M}$. Let the family $\{\mathbb{T}_h\}$ be shape regular and assume that on each macro M the set $\ker\left(B_M^T\right)$ only contains the constant functions. Then, there is a constant $\beta_M > 0$ independent of h such that for all M there holds

$$\inf_{p\in Q_M} \sup_{v\in V_M} \frac{(\nabla\cdot v, q)_M}{|v|_{1,M}|q|_M} \geq \beta_{\widehat{M}}. \tag{3.59}$$

3.2.68 Problem: Prove Lemma 3.2.67. Furthermore, prove that under the assumption that there is a finite set of reference macro elements $\widehat{M}_i$, such that all macro elements in a family are equivalent to one of them, the estimate holds with a uniform constant $\beta > 0$.

Remark 3.2.69. Depending on the technique of proof being used, we also may decide to impose (3.59) directly for each macro element.

Remark 3.2.70. So far, we have proven that under the assumption that the kernel of the macro problems only contains the constant functions, we have an inf-sup condition with the pressure norm $\|.\|_h$. It remains to use Verfürth's trick to prove the condition for $\|.\|_0$.

3.2.71 Lemma: Assume that there is an H^1-stable interpolation operator $I_h : V \to V_h$ according to Assumption 3.2.21. Then, there are positive constants c_1 and c_2 independent of h such that for any $q_h \in Q_h$ there holds

$$\sup_{v_h\in V_h} \frac{(\nabla\cdot v_h, q_h)}{\|v_h\|_V} \geq c_1\|q_h\|_Q - c_2\|q_h\|_h. \tag{3.60}$$

Proof. We begin by using the continuous inf-sup condition to deduce that for arbitrary $q_h \in Q_h$ there is $v \in V$ with $\|v\|_V \le \|q_h\|_Q$ such that

$$(\nabla\cdot v, q_h) \ge c_1 \|q_h\|_Q^2.$$

Now, let $v_h = I_h v$. Hence,

$$\begin{aligned}
(\nabla\cdot v_h, q_h) &= (\nabla\cdot v_h - \nabla\cdot v, q_h) + (\nabla\cdot v, q_h) \\
&\ge \sum_{T\in\mathbb{T}_h} (v - v_h, \nabla q_h)_T + \sum_{F\in\mathbb{F}_h^i} ([v_h - v]\cdot n_1, [\![q_h]\!])_F + c_1\|q_h\|_Q \\
&\ge -\left[\sum_{T\in\mathbb{T}_h} h_T^{-2}\|v - v_h\|_T^2 + \sum_{F\in\mathbb{F}_h^i} h_F^{-1}\|v - v_h\|_F\right] \|q_h\|_h + c_1\|q_h\|_Q \\
&\ge -c|v|_1\|q_h\|_h + c_1\|q_h\|_Q \\
&\ge \Big[c_1\|q_h\|_Q - c_2\|q_h\|_h\Big] \|q_h\|_Q.
\end{aligned}$$

Furthermore, we have by the interpolation estimate

$$|v_h| \le c|v| \le c\|q_h\|_Q,$$

which proves the result by dividing $(\nabla\cdot v_h, q_h)$ by the norm of v_h. □

3.2.72 Lemma: Let the assumptions of Lemma 3.2.71 hold, and assume that there is a constant $\tilde\beta$ such that

$$\sup_{v\in V_h} \frac{(\nabla\cdot v, q)}{\|v\|_1} \ge \tilde\beta \|q\|_h \qquad \forall q \in Q_h.$$

Then, the inf-sup condition

$$\sup_{v\in V_h} \frac{(\nabla\cdot v, q)}{\|v\|_1} \ge \beta \|q\|_Q \qquad \forall q \in Q_h,$$

holds with β determined by c_1, c_2 and $\tilde\beta$.

Proof. For any $q_h \in Q_h$ and $\vartheta \in [0,1]$ we have

$$\begin{aligned}
\sup_{v_h\in V_h} \frac{(\nabla\cdot v_h, q_h)}{\|v_h\|_V} &= \vartheta \sup_{v_h\in V_h} \frac{(\nabla\cdot v_h, q_h)}{\|v_h\|_V} + (1-\vartheta) \sup_{v_h\in V_h} \frac{(\nabla\cdot v_h, q_h)}{\|v_h\|_V} \\
&\ge \vartheta c_1\|q_h\|_Q - c_2\vartheta\|q_h\|_h + (1-\vartheta)\tilde\beta\|q_h\|_h \\
&\ge \frac{c_1\tilde\beta}{c_2 + \tilde\beta}\|q_h\|_Q,
\end{aligned}$$

by choosing $\vartheta = \tilde\beta/(c_2 + \tilde\beta)$. □

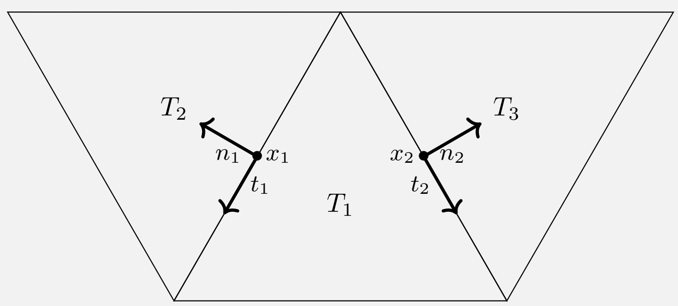
3.2.73 Theorem: The Hood-Taylor families are inf-sup stable. Thus, for solutions $u \in H^{k+1}(\Omega;\mathbb{R}^d) \cap V$ and $q \in H^k(\Omega) \cap Q$, there holds

$$\|u-u_h\|_1 + \|p-p_h\|_0 \le ch^k\big(|u|_{k+1} + |p|_k\big). \tag{3.61}$$

Proof. Summarizing all results of this section, the only thing that is left is defining a covering of $\mathbb{T}_h$ with macro elements, such that $\ker\left(B_M^T\right)$ contains only the constant functions. We do this at the example of the lowest order elements on quadrilaterals and triangles in the lemmas below. Both kinds of patches can be used to cover the whole mesh, such that we can use Lemma 3.2.66 and Lemma 3.2.72 to prove the inf-sup condition.

A general proof for higher order elements can be found in [SS96]. □

3.2.74 Lemma: For the $P_2 - P_1$ element choose the patch M as in



Then,

$$\ker\left(B_M^T\right) = \left\{q \in Q_M \,\middle|\, \forall v \in V_M\colon b(v,q) = 0\right\} = \mathbb{P}_0. \tag{3.62}$$

Proof. First, we observe that ∇q_h is constant on each cell and that the tangential derivatives $t_i \cdot \nabla q_h$ coincide for both adjacent cells due to the continuity of q_h. Now, we will derive conditions for $\ker\left(B_M^T\right)$ by choosing special test functions in V_M defined through interpolation in the points x_1 and x_2.

Furthermore, note that the shape function φ in $\mathbb{P}_2$ associated with the center of an edge is of the form $\lambda_1\lambda_2$ using the barycentric coordinates associated to the vertices at the ends of this edge. This function is positive everywhere inside the triangle T_i. Hence, there are positive numbers

$$w_1 = \int_{T_1} \varphi \,\mathrm{d}x, \quad w_1 = \int_{T_2} \varphi \,\mathrm{d}x,$$

Now, let $u(x_1) \cdot t_1 = 1$, $u(x_1) \cdot n_1 = 0$, and $u(x_2) = 0$. Then,

$$(\nabla\cdot u, q_h)_M = -(u, \nabla q_h)_M = -(w_1 + w_2)\nabla q_h \cdot t_1.$$

Hence, $q_h \in \ker\left(B_M^T\right)$ implies $\nabla q_h \cdot t_1 = 0$ in T_1 and T_2.

Exchanging x_2 for x_1, there holds $\nabla q_h \cdot t_2 = 0$ in T_2 and T_3. Since t_1 and t_2 are not collinear, we obtain

$$\nabla q_h|_{T_1} = 0. \tag{3.63}$$

Now we choose the test function $u(x_1) \cdot n_1 = 1$, $u(x_1) \cdot t_1 = 0$, and $u(x_2) = 0$. We get

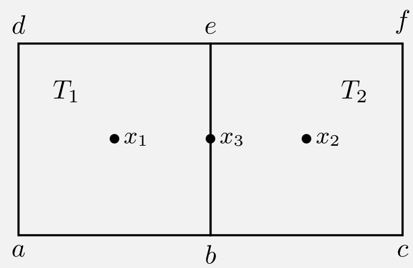
$$0 = (\nabla\cdot u, q_h) = -w_1 \nabla q_h|_{T_1} \cdot n_1 - w_2 \nabla q_h|_{T_2} \cdot n_1.$$

Due to (3.63), the first term vanishes and together with the tangential condition before, we obtain

$$\nabla q_h|_{T_2} = 0.$$

Exchanging again x_2 for x_1, we have the same for T_3, which proves the result. □

3.2.75 Lemma: For the $Q_2 - Q_1$ element choose the patch $\widehat{M}$ as in



Then,

$$\ker\left(B_M^T\right) = \left\{q \in Q_M \,\middle|\, \forall v \in V_M : b(v,q) = 0\right\} = \mathbb{P}_0.$$

Proof. Choose macro elements consisting of two quadrilateral charing an edge. Then, the reference macro element $\widehat{M}$ consists of the cells $[-1,0] \times [0,1]$ and $[0,1]^2$. We note that the velocity degrees of freedom are in x_1, x_2, and x_3, while those for the pressure are in a to f.

We do the analysis on the reference patch first. There, we have $u \in \mathbb{Q}_2^2$ and $\nabla q \in \mathbb{Q}_1^2$. Therefore, $u \cdot \nabla q \in \mathbb{Q}_3$ and the Simpson rule is exact on each cell. Hence

$$-(\nabla\cdot u, q)_{\widehat{M}} = (u, \nabla q)_{\widehat{M}} = \frac{4}{9} u(x_1)\nabla q(x_1) + \frac{4}{9} u(x_2)\nabla q(x_2) + \frac{4}{9} u(x_3)\nabla q(x_3).$$

We first test with velocites such that $u(x_3) = 0$ and one of $u_{1/2}(x_{1/2})$ is equal to one. Take for instance $u_1(x_1) = 1$. Then, the equation above implies for $q \in \ker\left(B_M^T\right)$ that $\partial_1 q(x_1) = 0$. Traversing through all four coimbinations, we obtain

$$\nabla q(x_1) = \nabla q(x_2) = 0.$$

On the cell T_2, we have by bilinear interpolation

$$q(x,y) = q(b)(1-x)(1-y) + q(x)x(1-y) + q(e)(1-x)y + q(f)xy,$$

and a similar representation on T_1. Thus, the conditions above translates to the system

$$\begin{aligned} q(c) + q(f) &= q(b) + q(e) \\ q(b) + q(c) &= q(e) + q(f) \\ q(b) + q(e) &= q(a) + q(d) \\ q(d) + q(e) &= q(a) + q(b), \end{aligned}$$

which in tern has solutions given for any $\alpha, \beta \in \mathbb{R}$ by

$$\begin{array}{ccccccc} q(a) &=& q(c) &=& q(e) &=& \alpha \\ q(b) &=& q(d) &=& q(f) &=& \beta \end{array}$$

The kernel of B_M^T is a subspace of the space generated by α and β. Now we choose $u_2(x_3) = 1$ and all other degrees of freedom zero. Again by simpson rule, we have

$$0 = -(\nabla\cdot u, q) = \frac{4}{9} u(x_3)\cdot\nabla q(x_3) = \frac{2}{9}\big(q(e) - q(b)\big).$$

Hence, $\alpha = q(e) = q(b) = \beta$ and

$$q \in \ker\left(B_M^T\right) \quad\Rightarrow\quad q \in \mathbb{P}_0.$$

For a patch M equivalent to $\widehat{M}$, we observe that

$$-(\nabla\cdot u, q)_M = (u, \nabla q)_M = \sum_{i=1,2} \int_{\widehat{T}_i} \hat{u}^T (\nabla\Phi)^{-T} \nabla\widehat{q} \, \det(\nabla\Phi)\, d\widehat{x}.$$

Cramer's rule implies

$$(\nabla\Phi_i)^{-T} \det(\nabla\Phi) = \begin{pmatrix} \partial_2\Phi_2 & -\partial_1\Phi_2 \\ -\partial_2\Phi_1 & \partial_1\Phi_1, \end{pmatrix}$$

where the mapping Φ is bilinear on each cell. Hence, on $\widehat{T}$

$$(\nabla\Phi_i)^{-T}\det(\nabla\Phi)\nabla\widehat{q}\in Q_1^2,$$

such that the integrand above is bicubic and the Simpson rule argument still applies. □

3.2.4 Almost incompressible elasticity

3.2.76. If we discretize the mixed formulation of almost incompressible elasticity with any of the stable Stokes pairs of the preceding sections, we can apply Corollary 2.5.10 to obtain optimal error estimates. Nevertheless, our problem with almost incompressible elasticity was not the approximation of the pressure (which was introduced artificially anyway), but locking. So, how does the choice of an inf-sup stable pair avoid locking?

Locking, in the terminology developed in the previous chapter, can be described as the fact that the kernel of the discrete divergence operator is too small, in the example presented even

$$\ker(B)_h = \{0\}.$$

Note though, that there might be more subtle locking effects, where the approximation is reduced but not destroyed.

In view of Theorem 2.4.4, locking means that V_h^g is too small or even the zero space, and therefore the quasi best-approximation result of this theorem is useless, since

$$\inf_{w_h\in V_h^g}\|u-w_h\|_V \not\to 0 \quad \text{as } h\to 0.$$

The additional assumption of the inf-sup condition inTheorem 2.4.7 on the other hand guarantees that an approximation of the kernel is possible, and thus, locking becomes impossible.

integration.tex integration.tex

3.2.77 Lemma: Let $V_h\times Q_h$ be a stable pair for the Stokes problem admitting a Korn inequality. Let furthermore Π_Q be the L^2-projection onto Q_h. Then, the solution $u_h\in V_h$ to the weak formulation

$$2\mu(\varepsilon(u_h),\varepsilon(v)) + \lambda(\Pi_Q\nabla\cdot u_h, \Pi_q\nabla\cdot v) = (f,v) \quad \forall v\in V_h, \tag{3.64}$$

admits the quasi-optimality estimate

$$\|u-u_h\|_1 \le c\sup_{v_h\in V_h}\|u-v_h\|_1, \tag{3.65}$$

with a constant c independent of the quotient λ/μ.

Proof. We introduce the auxiliary variable $p_h \in Q_h$ by the condition

$$\int_\Omega \nabla\cdot u_h q_h \,\mathrm{d}x = \int_\Omega p_h q_h \,\mathrm{d}x \quad \forall q_h \in Q_h.$$

By definition of the L^2-projection, we have

$$(\nabla\cdot u_h, q_h) = (\Pi_Q \nabla\cdot u_h, q_h) = (p_h, q_h).$$

Since $\Pi_Q \nabla\cdot u_h$ and p_h are in the same space, this implies that $p_h = \Pi_Q \nabla\cdot u_h$ pointwise. In addition, we observe that

$$(p_h, \Pi_Q \nabla\cdot v) = (p_h, \nabla\cdot v).$$

Hence, the formulation (3.64) is equivalent to

$$\begin{array}{cccccc} 2\mu(\varepsilon(u_h),\varepsilon(v)) & + & (p_h,\nabla\cdot v) & = & (f,v) & \forall v \in V_h \\ (\nabla\cdot u_h, q) & - & \frac{1}{\lambda}(p_h, q) & = & 0 & \forall q \in Q_h. \end{array}$$

This is the Stokes problem augmented by a positive definite bilinear form $c(.,.)$, such that Theorem 2.5.4 and Corollary 2.5.10 apply. □

Remark 3.2.78. The technique in the previous lemma is often called **reduced integration**. This refers to the fact, that we can replace the explicit projection Π_Q by using a quadrature formula which is exact on Q_h, but zero for all higher order polynomials occuring in $\nabla\cdot u_h$.

Chapter 4

Mixed formulation of elliptic problems

4.1 Modeling diffusion problems

4.1.1. Diffusion problems arise when a balance law, for instance for mass in ground water flow or for energy in temperature conduction is coupled with a constitutive equation relating the direction of movement to the gradient of the quantity of interest.

4.1.2. Let ϱ be the density of a conserved quantity. Then, for any given volume we have the "mass"

$$m = \int_V \varrho \,\mathrm{d}x.$$

Changes of this mass can be due to two processes:

1. Generation of additional mass by a source g,
2. Flow of mass over the boundary of V at a velocity v.

In formulas, we have

$$\tfrac{d}{dt} m = \int_V g \,\mathrm{d}x - \oint_{\partial V} J \cdot n \,\mathrm{d}s,$$

also known as **Reynolds transport theorem**. Here, J is the **flux**. The exact form of the flux will be modeled later. The formula above is somewhat unwieldy, since it combines volume and surface integrals. Therefore, we apply the Gauss theorem to obtain

$$\frac{d}{dt}\int_V \varrho \,\mathrm{d}x = \int_V g \,\mathrm{d}x - \int_V \nabla\cdot J \,\mathrm{d}x. \tag{4.1}$$

Concentrating and assuming sufficient regularity, we arrive at the equation

$$\partial_t \varrho + \nabla\cdot J = g. \tag{4.2}$$

As before in these notes, we ignore the time dependence and only look at stationary limits. In this case, this reduces to

$$\nabla\cdot J = g. \tag{4.3}$$

Example 4.1.3. Next we consider constitutive relations between ϱ and J such that we can complement equation (4.3) by a second equation and obtain a solvable system. To this end, we consider thermal diffusion and ground water flow.

Heat conduction: Here, the conserved quantity is not the density ϱ, but the temperature T. **Fourier's law** states that the flux is proportional to the gradient of the temperature, pointing in opposite direction:

$$J = -k\nabla T.$$

The constant of proportionality k is the heat conductivity.

Porous media flow: The conserved quantity is the amount of fluid, represented by the hydraulic head or pressure p. **Darcy's law** says that the flux is the product of the hydraulic **permeability** of the media and the gradient of the pressure:

$$J = -K\nabla p.$$

Here, the permeability K is either a positive scalar function or a symmetric, positive definite matrix. Note that in the latter case, J and ∇p do not point in the same direction.

General diffusion processes: Fick's law states, that the flux of a diffusion process is determined by the gradient of the diffusing quantity p by the relation

$$J = -D\nabla p.$$

D is the symmetric, positive definite **diffusion tensor**.

4.1.4. From the two equations for J, we derive the following system of PDE, where we replace the letter J by the more familiar u:

$$\begin{aligned} K^{-1}u + \nabla p &= 0 \\ \nabla\cdot u &= f. \end{aligned} \tag{4.4}$$

This system is closed by boundary conditions. Let Γ_D be the Dirichlet boundary and Γ_N be the Neumann boundary such that $\Gamma_D \cap \Gamma_N = \emptyset$ and $\Gamma_D \cup \Gamma_N = \partial\Omega$. Then, we let

$$\begin{aligned} p(x) &= p^D(x) & x &\in \Gamma_D, \\ u(x)\cdot n &= u^N(x)\cdot n & x &\in \Gamma_N. \end{aligned} \tag{4.5}$$

Following the concept of finding spaces such that we have an inf-sup condition, we are looking for a pair with minimal regularity, such that we have a stable and bounded inf-sup condition. We begin the usual way by multiplying with a test function and integrating:

$$\begin{aligned} \int_\Omega K^{-1}u\cdot v\,\mathrm{d}x + \int_\Omega \nabla p\cdot v\,\mathrm{d}x &= 0 \\ \int_\Omega \nabla\cdot u q\,\mathrm{d}x &= \int_\Omega f q\,\mathrm{d}x. \end{aligned} \tag{4.6}$$

It turns out, we have two immediate options: first, we can integrate the first equation by parts, having all derivatives on u and v. On the other hand, we can integrate by parts in the second equation, leaving all derivatives on p and q. In the second case, we obtain the equation

$$-\int_\Omega u\cdot\nabla q\,\mathrm{d}x + \int_{\partial\Omega} u\cdot n q\,\mathrm{d}s = \int_\Omega f q\,\mathrm{d}x.$$

Applying the boundary condition, we first follow the recipe of elliptic partial differential equations and implement $p = p^D$ as an essential boundary condition, that is, the test function space has zero trace on Γ_D. Then, we can swap in u^N for u on Γ_N, such that the boundary term ends up on the right hand side.

4.1.5 Definition: The **primal mixed formulation** of the mixed diffusion problem (4.4) reads: find $(u,p) \in V\times Q$ such that for all $v\in V$ and $q \in Q$ holds

$$\begin{aligned} \left(K^{-1}u, v\right) + (\nabla p, v) &= 0 \\ -(u,\nabla q) &= (f,q) - \left\langle u^N\cdot n, q\right\rangle_{\Gamma_N}. \end{aligned} \tag{4.7}$$

The spaces are

$$\begin{aligned} V &= L^2(\Omega;\mathbb{R}^d), \\ Q &= H^1_{\Gamma_D}(\Omega) = \left\{q\in H^1(\Omega) \middle|\, q_{|\Gamma_D} = 0\right\}. \end{aligned} \tag{4.8}$$

Remark 4.1.6. Since the first equation is tested with the test function v itself in all terms, we can eliminate this equation and there holds $u = K\nabla p$ in $L^2(\Omega;\mathbb{R}^d)$. Entering this into the second equation, we obtain the well-known primal formulation

$$(K\nabla p, \nabla q) = (f, q) - \left\langle u^N \cdot n, q\right\rangle_{\Gamma_N}.$$

Just keep in mind that the “natural boundary condition” in this case is

$$K\nabla p \cdot n = 0.$$

Hence, the primal mixed formulation does not provide any advantages compared to the primal formulation, and we are not going to pursue it further.

4.1.7. Now we return to the first alternative, namely integrating by parts in the first equation of (4.6):

$$\int_\Omega K^{-1}u \cdot v \,\mathrm{d}x - \int_\Omega p\nabla\cdot v \,\mathrm{d}x + \int_{\partial\Omega} v \cdot np \,\mathrm{d}s = 0.$$

Ensuing is a formulation multiplying and integrating the divergences of u and v, respectively, with functions in Q. In order to fit this into our standard framework, we have to introduce a new Sobolev space. In addition, since $u \cdot n$ does not appear as a boundary integral, we must make this an essential boundary condition. Thus, we require that the test functions have zero normal trace on Γ_N (and justify this below). Note that now the Dirichlet condition $p = 0$ has become a “natural boundary condition”!

4.1.8 Definition: Let $\Omega \subset \mathbb{R}^d$ be a domain. We define the Sobolev space

$$H^{\mathrm{div}}(\Omega) = \left\{v \in L^2(\Omega;\mathbb{R}^d) \middle| \nabla\cdot v \in L^2(\Omega)\right\}, \tag{4.9}$$

and its inner product

$$\langle u, v\rangle_{H^{\mathrm{div}}} = (u, v)_0 + (\nabla\cdot u, \nabla\cdot v)_0. \tag{4.10}$$

Furthermore, let $C^\infty_{00}(\Omega)$ be the space of smooth functions with compact support in Ω. Then, we define its closure in $H^{\mathrm{div}}(\Omega)$:

$$H^{\mathrm{div}}_0(\Omega) = \overline{C^\infty_{00}(\Omega)}. \tag{4.11}$$

For subset $\Gamma \subset \partial\Omega$, the space $H^{\mathrm{div}}_\Gamma(\Omega)$ is defined accordingly (compare to $H^1_\Gamma(\Omega)$)

Using the space H^{div} and for the moment the assumption, that H^{div}_0 and H^{div}_Γ serve to set boundary conditions, we can write down our second weak formulation of the mixed diffusion problem:

4.1.9 Definition: The **dual mixed formulation** of the mixed diffusion problem (4.4) reads: find $(u,p) \in V \times Q$ such that for all $v \in V$ and $q \in Q$ holds

$$\begin{aligned} \left(K^{-1}u, v\right) - (p, \nabla\cdot v) &= \left\langle p^D, v\cdot n\right\rangle_{\Gamma_D} \\ (\nabla\cdot u, q) &= (f, q). \end{aligned} \tag{4.12}$$

The spaces are

$$V = H^{\mathrm{div}}_{\Gamma_N}(\Omega), \qquad Q = L^2(\Omega). \tag{4.13}$$

4.1.1 Properties of $H^{\mathbf{div}}(\Omega)$

4.1.10 Theorem: Let Ω be a bounded Lipschitz domain. Then, the space $C^\infty(\overline{\Omega}; \mathbb{R}^d)$ is dense in $H^{\mathrm{div}}(\Omega)$.

Proof. Either by a standard mollifier argument [AF03] or following [GR86, Theorem 2.4] □

Remark 4.1.11. The condition of boundedness entered the assumptions since we use the space $C^\infty(\overline{\Omega})$. It could be dropped, if we used a more appropriate space (cf. [GR86, Theorem 2.4]).

4.1.12 Theorem: The trace operator $\gamma_n\colon C^\infty(\overline{\Omega}; \mathbb{R}^d) \to C^\infty(\overline{\partial\Omega})$ which maps $v \mapsto v\cdot n_{|\partial\Omega}$ can be extended to a continuous, linear mapping

$$\gamma_n\colon H^{\mathrm{div}}(\Omega) \to H^{-1/2}(\partial\Omega), \tag{4.14}$$

where $H^{-1/2}(\partial\Omega)$ is the dual of $H^{1/2}(\partial\Omega)$.

Proof. Let $q \in C^\infty(\overline{\Omega})$ and $v \in C^\infty(\overline{\Omega}; \mathbb{R}^d)$. Then, there holds Green's formula

$$(v, \nabla q)_\Omega + (\nabla\cdot v, q)_\Omega = \langle v\cdot n, q\rangle_{\partial\Omega}.$$

Hence,

$$\left| \int_{\partial\Omega} v\cdot n q \,\mathrm{d}s \right| \le \|v\|_{H^{\mathrm{div}}} \|q\|_{H^1}.$$

Applying the density of $C^\infty(\overline{\Omega})$ in $H^1(\Omega)$ and of $C^\infty(\overline{\Omega};\mathbb{R}^d)$ in $H^{\operatorname{div}}(\Omega)$, we can let q and v pass to a limit, but the inequality holds uniformly.

Now apply that $H^{1/2}(\partial\Omega)$ is the trace space of $H^1(\Omega)$. Therefore, for any $g \in H^{1/2}(\partial\Omega)$, there is a $q \in H^1(\Omega)$ such that $q_{|\partial\Omega} = g$ and $\|q\|_{1;\Omega} \le \|g\|_{1/2;\partial\Omega}$. We obtain

$$\left|\int_{\partial\Omega} v \cdot n g \,\mathrm{d}s\right| \le \|v\|_{H^{\operatorname{div}}(\Omega)} \|g\|_{H^{1/2}(\partial\Omega)} \qquad \forall v \in H^{\operatorname{div}}(\Omega), g \in H^{1/2}(\partial\Omega).$$

Hence,

$$\|v \cdot n\|_{H^{-1/2}(\partial\Omega)} \le \|v\|_{H^{\operatorname{div}}(\Omega)} \qquad \forall v \in H^{\operatorname{div}}(\Omega).$$

Thus, we have proven the continuity of the extension of γ_n to $H^{\operatorname{div}}(\Omega)$. □

Remark 4.1.13. The trace theorem tells us that our interpretation of the spaces $H_0^{\operatorname{div}}(\Omega)$ and $H_\Gamma^{\operatorname{div}}(\Omega)$ as spaces with zero boundary condition of the normal component is justified. This notion will be fortified by the two theorems below. Therefore, we will later avoid the notational overhead of using γ_n and will simply write $v \cdot n_{|\partial\Omega}$.

4.1.14 Problem: Show the following result. Let $p \in H^1(\Omega)$ and $\Delta p \in L^2(\Omega)$. Then, $\partial_n p \in H^{-1/2}(\partial\Omega)$ and

$$(\nabla p, \nabla q) = -(\Delta p, q) + \langle \partial_n p, q\rangle_{\partial\Omega} \quad \forall q \in H^1(\Omega).$$

4.1.15 Theorem: The trace theorem is optimal in the sense that $\gamma_n \colon H^{\operatorname{div}}(\Omega) \to H^{-1/2}(\partial\Omega)$ is surjective.

Proof. Let $\mu \in H^{-1/2}(\partial\Omega)$. We have to show that there exists $v \in H^{\operatorname{div}}(\Omega)$ such that

$$v \cdot n = \mu \quad \text{on } \partial\Omega \qquad \text{and} \qquad \|v\|_{H^{\operatorname{div}}(\Omega)} \le \|\mu\|_{H^{-1/2}(\partial\Omega)}.$$

We know that the problem

$$\begin{aligned} -\Delta\varphi + \varphi &= 0 && \text{in } \Omega,\\ \partial_n\varphi &= \mu && \text{on } \partial\Omega, \end{aligned}$$

has a unique solution $\varphi \in H^1(\Omega)$ with

$$\|\varphi\|_{H^1(\Omega)}^2 = \langle \mu, \varphi\rangle_{\partial\Omega} \le \|\mu\|_{H^{-1/2}(\partial\Omega)} \|\varphi\|_{H^1(\Omega)}.$$

The first equation then implies $\Delta\varphi \in L^2(\Omega)$ and thus $v = \nabla\varphi \in H^{\text{div}}(\Omega)$. Since from this equation there even holds $\nabla\cdot v = \varphi$, we obtain

$$\|v\|_{H^{\text{div}}(\Omega)} \le \|\mu\|_{H^{-1/2}(\partial\Omega)}.$$

□

4.1.16 Theorem: There holds

$$\ker(\gamma_n) = H_0^{\text{div}}(\Omega). \tag{4.15}$$

Proof. The inclusion $H_0^{\text{div}}(\Omega) \subset \ker(\gamma_n)$ follows immediately from the definition and continuity of γ_n. For the opposite inclusion, we have to show that the traces of functions in $C_{00}^\infty(\Omega)$ are dense in $\ker(\gamma_n)$. We do this by using, that a subspace W is dense in a space V if and only if all linear functionals vanishing on W also vanish on V. Choose $u \in \ker(\gamma_n)$ and use the Riesz representation theorem to associate with it $L \in \ker(\gamma_n)^*$ by

$$L(v) = \langle u, v\rangle_{H^{\text{div}}} \qquad \forall v \in \ker(\gamma_n).$$

Assume now that $L(\varphi) = 0$ for all $\varphi \in C_{00}^\infty(\Omega;\mathbb{R}^d)$. This implies by

$$0 = L(\varphi) = (u,\varphi)_{L^2} + (\nabla\cdot u, \nabla\cdot\varphi),$$

that $u = \nabla\nabla\cdot u$ in distributional sense, and by taking limits of φ in H^1 that $\nabla\cdot u \in H^1(\Omega)$. Hence, Green's formula yields

$$L(v) = (\nabla\nabla\cdot u, v) + (\nabla\cdot u, \nabla\cdot v) = \langle v\cdot n, \nabla\cdot u\rangle_{\partial\Omega} = 0 \qquad \forall v \in \ker(\gamma_n).$$

Thus, L vanishes on all elements of $\ker(\gamma_n)$ and the theorem is proven. □

4.1.17 Theorem: Let Ω be connected. Let

$$V_0 = \bigl\{v \in H_0^{\text{div}}(\Omega) \big| \nabla\cdot v = 0\bigr\}. \tag{4.16}$$

Then,

$$L^2(\Omega;\mathbb{R}^d) = V_0 \oplus V^\perp, \tag{4.17}$$

and

$$V^\perp = \bigl\{v = \nabla q \big| q \in H^1(\Omega)\bigr\}. \tag{4.18}$$

Proof. Let $X = \{v = \nabla q | q \in H^1(\Omega)\}$. we have to show $V^\perp = X$. Observe that X is closed in L^2 since H^1 is complete. We show that $V_0 = X^\perp$ and thus

$$V_0^\perp = (V_0^\perp)^\perp = \overline{X} = X.$$

First, let $u \in V_0$. Then, Green's formula reduces to

$$(u, \nabla q) = 0 \qquad \forall q \in H^1(\Omega).$$

Hence, $V_0 \subset X^\perp$. Let now conversely $u \in L^2(\Omega; \mathbb{R}^d)$ such that the previous identity holds. Choosing $q \in C_{00}^\infty(\Omega)$ yields $\nabla \cdot u = 0$, which in turn means $u \in H^{\mathrm{div}}(\Omega)$. Therefore, we can use Green's formula to obtain $u \cdot n = 0$ on $\partial\Omega$. This together implies $u \in V_0$, proving $X^\perp \subset V_0$. □

4.1.2 Well-posedness of the dual mixed formulation

4.1.18. In order to apply the theory from Chapter 2, we have to define the abstract bilinear forms $a(.,.)$ and $b(.,.)$. We read from the dual mixed formulation

$$a(u, v) = \left(K^{-1}u, v\right)$$
$$b(v, q) = (\nabla \cdot v, q).$$

4.1.19 Problem: In both the primal and the dual mixed formulation, we ignored inhomogeneous essential boundary conditions. Show that the usual lifting method applies. Determine the modified equations and the spaces needed for the liftings.

4.1.20 Lemma: Let $V = H^{\mathrm{div}}(\Omega)$ and $Q = L^2(\Omega)$ with their norms. Let

$$V_0 = \ker(B) = \left\{v \in V \middle| (\nabla \cdot v, q) = 0 \; \forall q \in Q\right\}. \tag{4.19}$$

Assume there exist constants γ and $\|a\|$ such that

$$\gamma|\xi|^2 \le \xi^T K^{-1} \xi \le \|a\| |\xi|^2 \qquad \forall \xi \in \mathbb{R}^d. \tag{4.20}$$

Then, there holds

$$a(u, v) \le \|a\| \|u\|_V \|v\|_V \qquad \forall u, v \in V \tag{4.21}$$
$$a(u, u) \ge \gamma \|u\|_V^2 \qquad \forall u \in \ker(B). \tag{4.22}$$

Remark 4.1.21. Differing from the Stokes problem, ellipticity of $a(.,.)$ cannot be extended to the whole space V. This is going to be the major difference between this chapter and the previous.

4.1.22 Lemma: Let $V = H^{\text{div}}(\Omega)$ and $Q = L^2(\Omega)$ with their norms. Then, the inf-sup condition

$$\inf_{q\in Q} \sup_{v\in V} \frac{b(v,q)}{\|v\|_V \|q\|_Q} \geq \beta \tag{4.23}$$

holds with a constant β depending on the domain.

Proof. We can use the construction leading to Corollary 3.1.7. In fact, since the norm of H^{div} is weaker than the one of H^1, the same function v can be chosen in the Stokes inf-sup condition (3.1), yielding a constant β not worse than for Stokes. □

Combining these lemmas yields the assumptions of Theorem 2.3.3. Thus, we have proven:

4.1.23 Theorem: Under the assumptions on Lemma 4.1.20, the dual mixed formulation is well-posed.

4.2 Discretization of dual mixed problems

4.2.1 Conforming subspaces of $H^{\mathbf{div}}(\Omega)$

4.2.1. Our goal in this section is the derivation of general criteria applying to the approximation of $H^{\text{div}}(\Omega)$ by piecewise polynomial functions. This affects in particular continuity conditions and the properties of $\ker(B_h)$.

As be before, we will assume that all families of meshes $\mathbb{T}_h$ for $h \to 0$ are shape-regular. We will also assume that meshes are regular, unless otherwise stated.

4.2.2 Lemma: Let $\mathbb{T}_h$ be a subdivision of the domain Ω. Let the space V_h be cell-wise polynomial. We have $V_h \subset H^{\text{div}}(\Omega)$ if and only if on each interior face F between two cells T_1 and T_2 holds

$$v_1 \cdot n_1 + v_2 \cdot n_2 = 0. \tag{4.24}$$

Here, v_1 and v_2 are the traces of the functions on F from each cell.

Proof. Since V_h is by definition finite dimensional, all norms are bounded. It remains to show that the distributional divergence is in $L^2(\Omega)$, that is, all its contributions which are Borel measures of faces vanish. To this end, let $\varphi \in C^\infty_{00}(\Omega)$ be a test function such that its support does not have a nonempty intersection with any face except F. Then, we have by Green's formula for $u \in V_h$

$$(\nabla\cdot u, \varphi) = -(u, \nabla\varphi) + \langle u_1 \cdot n_1 + u_2 \cdot n_2, \varphi\rangle_F$$

We have $\nabla\cdot u \in L^2(\Omega)$ if and only if the face term vanishes. □

4.2.3 Problem: Show that the corresponding condition for H^1-conforming finite elements is continuity of the function. In particular, this implies continuity at vertices.

Remark 4.2.4. The continuity of the normal component over faces does not imply continuity at vertices, since it is not transferred in tangential direction.

As a consequence of this remark, part of the construction of H^{div}-conforming finite element spaces consists of defining a polynomial trace space on each face, such that continuity of normal traces can be established by this space.

4.2.5. In Lemma 4.1.20, we saw that the bilinear form $a(.,.)$ is elliptic only on the kernel of B. Indeed, for the simplest case with $K \equiv 1$, we conclude that the uniform estimate for $v_h \in \ker(B_h)$

$$\|v_h\|^2_{L^2} \geq \gamma\|v_h\|^2_{H^{\mathrm{div}}} = \gamma\bigl(\|v_h\|^2_{L^2} + \|\nabla\cdot v_h\|^2_{L^2}\bigr),$$

necessary for quasi-bestapproximation requires a constant c independent of h such that

$$\|\nabla\cdot v_h\|^2_{L^2} \leq c\|v_h\|^2_{L^2} \qquad \forall v_h \in \ker(B_h).$$

The inverse estimate is insufficient by two powers of h, such that this is actually a hard condition. Therefore, we focus on methods where

$$\ker(B_h) \subset \ker(B). \tag{4.25}$$

Remark 4.2.6. A particularly elegant way to achieve (4.25) is the choice

$$\nabla\cdot V_h = Q_h. \tag{4.26}$$

We will indeed focus on methods with this property.

4.2.2 Finite elements on simplices

4.2.7. Simplicial elements based on the polynomial spaces $\mathbb{P}_k$ of polynomials of total degree less or equal k can be defined on the actual mesh cell. We present here the two most common families.

4.2.8 Definition: The **Raviart-Thomas element** of degree $k \geq 0$ on simplices consists of the polynomial space

$$RT_k = \mathbb{P}_k^d + x\mathbb{P}_k. \tag{4.27}$$

Its node functionals are

$$\mathcal{N}_{1,i,j}(v) = \int_{F_i} v \cdot n \, q_j \, \mathrm{d}s \qquad q_j \in \mathbb{P}_k(F_i) \qquad F_i \subset \partial T, \tag{4.28}$$

$$\mathcal{N}_{2,i}(v) = \int_T v \cdot w_i \, \mathrm{d}x \qquad w_i \in \mathbb{P}_{k-1}^d(T). \tag{4.29}$$

Here, and in further definitions of this kind, the notation q_j and w_i indicate that we choose a basis for the polynomial spaces.

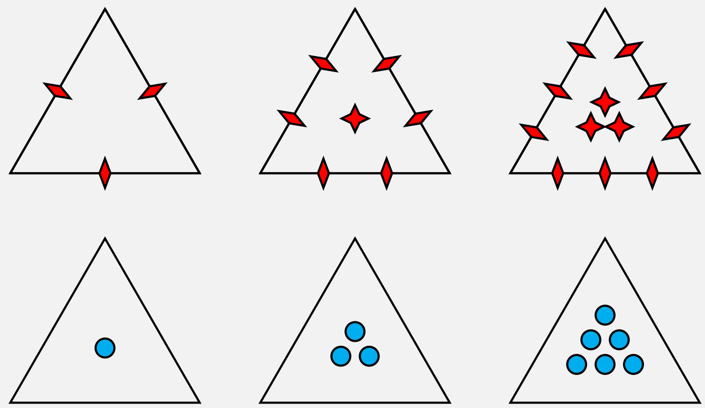
Remark 4.2.9. In equation (4.28), the F_i are all faces of the simplex, that is, three edges in two dimensions and four triangular faces in three dimensions.

If the simplex T is obtained by affine transformation from the reference simplex $\widehat{T}$, the definitions of the space RT_k directly on the cell T and by mapping from $\widehat{T}$ coincide. Therefore, we can use arguments by mapping or without at our convenience.

For $k = 0$ there are no nodal values of type $\mathcal{N}_{2,i}$ since all gradients of functions in $\mathbb{P}_0$ are zero.

The unisolvence will be shown in Lemma 4.2.13. But first, we have to look at some important properties.

4.2.10 Example: The first members of the Raviart-Thomas family on triangles are



4.2.11 Lemma: For any simplex $T \in \mathbb{R}^d$ we have for any $v \in RT_k$ and any $F \subset \partial T$

$$\nabla\cdot v \in \mathbb{P}_k(T), \tag{4.30}$$

$$v \cdot n_{|F} \in \mathbb{P}_k(F). \tag{4.31}$$

The divergence operator is surjective from RT_k to $\mathbb{P}_k$, hence

$$\nabla\cdot RT_k = \mathbb{P}_k. \tag{4.32}$$

For the divergence free functions holds

$$RT_{k,0} = \left\{ v \in RT_k \middle| \nabla\cdot v = 0 \right\} \subset \mathbb{P}_k^d. \tag{4.33}$$

Proof. We write an arbitrary element $v \in RT_k$ as $v = v_0 + xp_k$ where $v_0 \in \mathbb{P}_k^d$ and $p_k \in \mathbb{P}_k$. Clearly, $\nabla\cdot v_0 \in \mathbb{P}_{k-1}$. On the other hand,

$$\nabla\cdot(xp_k) = \nabla\cdot xp_k + x\nabla p_k = dp_k + xq,$$

where $q \in \mathbb{P}_{k-1}^d$. Therefore, $\nabla\cdot(xp_k) \in \mathbb{P}_k$ and so is $\nabla\cdot v$.

For a given face F, choose $x_0 \in F$. Every other point $x \in \mathbb{F}$ can be represented as $x = x_0 + \tau$, where τ is a vector tangential to F. Therefore, using the same splitting of v as above

$$v \cdot n = v_0 \cdot n + p_k(x \cdot n) = v_0 \cdot n + p_k(x_0 \cdot n + \tau \cdot n).$$

The last term vanishes by definition of τ and n and the two other terms are both in $\mathbb{P}_k$.

Finally, we show surjectivity of the divergence operator. We show indeed that the divergence is surjective from $\tilde{V} = (x - x_c)\mathbb{P}_k$ to $\mathbb{P}_k$, where x_c is the center of T. Note that $x_c\mathbb{P}_k \in \mathbb{P}_k^d$ such that $\tilde{V} \subset RT_k$. Furthermore, $\tilde{V}$ and $\mathbb{P}_k$ have the same dimension. Therefore, it is sufficient to show that the divergence is injective. For simplicity, we assume $x_c = 0$. Then, for any $p \in \mathbb{P}_k$

$$\begin{aligned}\int_T \nabla\cdot(xp)p\,\mathrm{d}x &= d\int_T p^2\,\mathrm{d}x + \int_T x\cdot\nabla pp\,\mathrm{d}x \\ &= d\int_T p^2 + \frac{1}{2}\int_T x\cdot\nabla(p^2)\,\mathrm{d}x \\ &= \frac{d}{2}\int_T p^2 + \frac{1}{2}\int_{\partial T}(x\cdot n)p^2\,\mathrm{d}s.\end{aligned}$$

Thus, $\nabla\cdot(xp) = 0$ implies $p = 0$ and thus $xp = 0$, which proves the injectivity. Using the same idea, we see that $\nabla\cdot v = 0$ for $v = v_0 + xp$ implies $xp = 0$ and thus $v = v_0 \in \mathbb{P}_k^d$. □

4.2.12 Lemma: There for the simplicial Raviart-Thomas element in $\mathbb{R}^d$ there holds

$$\begin{aligned}\dim RT_k &= (d+1)\dim\mathbb{P}_k - \dim\mathbb{P}_{k-1} \\ &= (d^2 + kd + d)\frac{(k+d-1)!}{d!k!}.\end{aligned} \tag{4.34}$$

In particular,

$$\dim RT_k = \begin{cases}(k+1)(k+3) & d = 2, \\ \frac{1}{2}(k+1)(k+2)(k+4) & d = 3.\end{cases} \tag{4.35}$$

Proof. First, we observe that

$$RT_k = \mathbb{P}_k^d \oplus x\breve{\mathbb{P}}_k,$$

where $\breve{\mathbb{P}}_k$ is the space of homogeneous polynomials of degree k, that is, those strictly of degree k. There also holds

$$\mathbb{P}_k = \mathbb{P}_{k-1} \oplus \breve{\mathbb{P}}_k.$$

Hence,

$$\dim RT_k = d\dim\mathbb{P}_k + \dim\breve{\mathbb{P}}_k = d\dim\mathbb{P}_k + \dim\mathbb{P}_k - \dim\mathbb{P}_{k-1},$$

which proves the general formula. Using

$$\dim \mathbb{P}_k = \frac{1}{d!}\frac{(k+d)!}{k!} = \begin{cases} \frac{(k+1)(k+2)}{2} & d=2, \\ \frac{(k+1)(k+2)(k+3)}{6} & d=3, \end{cases} \tag{4.36}$$

proves the explicit formulas. □

4.2.13 Lemma: The Raviart-Thomas element with the nodal values in Definition 4.2.8 is unisolvent.

Proof. As usual, the proof consists of two parts: first, we prove that the number of node functionals equals the dimension of RT_k. To this end, we observe that

$$\dim \mathbb{P}_k(\mathbb{R}^d) = \frac{1}{d!}\frac{(k+d)!}{k!} = \frac{k+d}{d} \dim \mathbb{P}_k(\mathbb{R}^{d-1}). \tag{4.37}$$

The number of node functionals is

$$\begin{aligned} N &= (d+1)\dim \mathbb{P}_k(\mathbb{R}^{d-1}) + d \dim \mathbb{P}_{k-1}(\mathbb{R}^d) \\ &= (d+1)\frac{(k+d-1)!}{(d-1)!k!} + d\frac{(k+d-1)!}{d!(k-1)!} \\ &= (d^2+d+kd)\frac{(k+d-1)!}{d!k!} \end{aligned}$$

Thus, the number of node functionals is equal to the dimension of RT_k. Therefore, every element in RT_k is uniquely determined by the node functionals if and only if for $v \in RT_k$

$$\left\{ \begin{array}{cc} \mathcal{N}_{1,i,j}(v) = 0 & \forall i,j \\ \mathcal{N}_{2,i}(v) = 0 & \forall i \end{array} \right\} \quad \Longrightarrow \quad v = 0.$$

To this end, we first observe that due to (4.31) the node functionals $\mathcal{N}_{1,i,j}$ for $j = 1, \dots, \dim \mathbb{P}_k(F_i)$ uniquely determine v on each face F_i. Therefore,

$$\{\mathcal{N}_{1,i,j}(v) = 0 \quad \forall i,j\} \quad \Longrightarrow \quad v \in H_0^{\text{div}}(T).$$

Next, we test (4.29) with $w = \nabla q$ and $q \in \mathbb{P}_k$ arbitrary. After integration by parts, this implies $v \in RT_{k,0} \cap H_0^{\text{div}}(T) \subset \mathbb{P}_k^d$.

For the remaining part of the proof, we need a result which will be presented later in full generality. At this point, we only mention that in two dimensions, the space V_0 of divergence free functions in $L^2(T)$ has the representation

$$V_0 = \left\{ \nabla\times\varphi \middle| \varphi \in H^1(T) \right\}, \qquad \nabla\times\varphi = \begin{pmatrix} \partial_2\varphi \\ -\partial_1\varphi \end{pmatrix} \tag{4.38}$$

We also notice that $v\cdot n = \partial_\tau \varphi$, where τ is the tangential vector with the domain T on the right. Thus, $v \in H_0^{\text{div}}(T)$ implies φ is constant on the boundary. Since moreover we only use derivatives of φ, we can choose $\varphi \in H_0^1(T)$. Finally, since $v \in RT_{k,0}$, we have $\varphi \in \mathbb{P}_{k+1}$. Any function φ with these properties can be expressed by the cubic bubble function as

$$\varphi = b_T \psi \qquad \psi \in \mathbb{P}_{k-2}.$$

We conclude the proof with

$$0 = \int_T v \cdot w \,\mathrm{d}x = \int_T \nabla\times\varphi \cdot w \,\mathrm{d}x = \int_T b_T \psi (\partial_2 w - \partial_1 w) \,\mathrm{d}x.$$

Choose w such that $\partial_2 w - \partial_1 w = \psi$ to obtain $\psi = 0$. □

4.2.14 Definition: The **BDM element** (Brezzi-Douglas-Marini) of degree $k \geq 1$ on simplices consists of the polynomial space

$$BDM_k = \mathbb{P}_k^d. \tag{4.39}$$

Its node functionals are

$$\mathcal{N}_{1,i,j}(v) = \int_{F_i} v \cdot n \, q_j \,\mathrm{d}s \qquad q_j \in \mathbb{P}_k(F_i) \qquad F_i \subset \partial T, \tag{4.40}$$

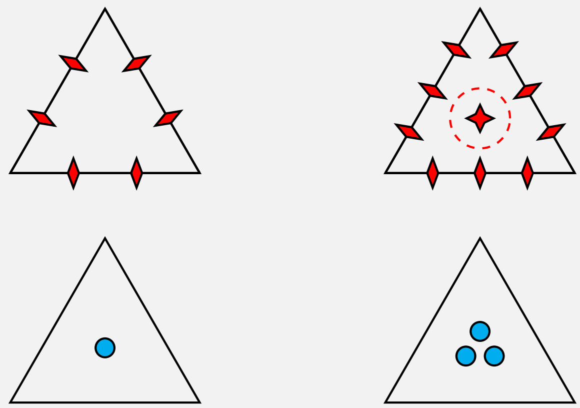
$$\mathcal{N}_{2,i}(v) = \int_T v \cdot \nabla q_i \,\mathrm{d}x \qquad q_i \in \mathbb{P}_{k-1}(T). \tag{4.41}$$

$$\mathcal{N}_{3,i}(v) = \int_T v \cdot w_i \,\mathrm{d}x \qquad w_i \in V_{k,0}(T), \tag{4.42}$$

where

$$V_{k,0}(T) = \left\{ v \in \mathbb{P}_k^d(T) \cap H_0^{\text{div}}(T) \,\middle|\, \nabla\cdot v = 0 \right\}. \tag{4.43}$$

4.2.15 Example: The first members of the BDM family on triangles are



4.2.16 Lemma: The BDM element with the nodal values in Definition 4.2.14 is unisolvent.

Proof. Let $v \in BDM_k(T)$. First, we note that $\nabla \cdot BDM_k \subset \mathbb{P}_{k-1}$. Therefore, setting the node functionals in (4.40) and (4.41) to zero implies $v \in V_{0,k}$. But then, the remaining node functionals are an inner product on $V_{0,k}$ and thus $v = 0$. □

Remark 4.2.17. What is missing here is a characterization of the space $V_{0,k}$. Thus, we cannot really implement the method yet. Furthermore, we cannot verify that the node functionals form a dual basis for $BDM_k(T)$. The answer to these questions will be given in Chapter 6.2.

4.2.3 Stability by commuting diagrams

4.2.18. Again, we show stability by constructing a Fortin projection. To this end, we show that the nodal interpolation of the Raviart-Thomas and BDM families indeed commute with the divergence operator. We first show that this holds for smooth functions and then discuss the extension to functions in $H^{\text{div}}(\Omega)$.

4.2.19 Definition: Let a finite element be defined by its shape function space $\mathcal{P}_T$ and the node functionals $\mathcal{N}_i$ for $i = 1, \dots, n$ where $n = \dim \mathcal{P}_T$. Let $\{\varphi_j\}$ be the basis of $\mathcal{P}_T$ such that

$$\mathcal{N}_i(\varphi_j) = \delta_{ij} \qquad i, j = 1, \dots, n.$$

Then, the operator $I_h \colon C^\infty(T) \to \mathcal{P}_T$ defined for any $f \in C^\infty(T)$ by

$$I_h(f) = \sum_{i=1}^{n} \mathcal{N}_i(f)\varphi_i, \tag{4.44}$$

is called the **canonical interpolation** operator. The definition applies to vector valued elements replacing $C^\infty(T)$ by $C^\infty(T;\mathbb{R}^d)$.

Remark 4.2.20. For a vector polynomial space $\mathcal{V}_T$, we define its divergence space

$$\mathcal{P}_T = \nabla\cdot\mathcal{V}_T.$$

From Lemma 4.2.11, we obtain

$$\nabla\cdot RT_k = \mathbb{P}_k. \tag{4.45}$$

For the BDM family, it is easy to verify

$$\nabla\cdot BDM_k = \mathbb{P}_{k-1}.$$

4.2.21 Lemma: Let $I_h \colon C^\infty(T;\mathbb{R}^d) \to \mathcal{V}_T$ be the canonical interpolation onto the space $\mathcal{V}_T$, which is either $RT_k(T)$ or $BDM_{k+1}(T)$. Let $\mathcal{P}_T = \nabla\cdot\mathcal{V}_T = \mathbb{P}_k$ and $\Pi_h \colon C^\infty(T) \to \mathcal{P}_T$ be the L^2-projection onto $\mathcal{P}_T$. Then, the diagram

$$\begin{array}{ccc} C^\infty(T;\mathbb{R}^d) & \xrightarrow{\nabla\cdot} & C^\infty(T) \\ {\scriptstyle I_h}\downarrow & & \downarrow{\scriptstyle \Pi_h} \\ \mathcal{V}_T & \xrightarrow{\nabla\cdot} & \mathcal{P}_T \end{array} \tag{4.46}$$

commutes, that is, for any $v \in C^\infty(T;\mathbb{R}^d)$, there holds

$$\nabla\cdot(I_h v) = \Pi_h(\nabla\cdot v). \tag{4.47}$$

Proof. Let $v \in C^\infty(T;\mathbb{R}^d)$ and $q \in C^\infty(\Omega)$ be chosen arbitrarily. Then,

$$\int_T q(\nabla\cdot(I_h v) - \nabla\cdot v)\,\mathrm{d}x = \int_T (v - I_h v)\cdot\nabla q\,\mathrm{d}x - \int_{\partial T}(v - I_h v)\cdot nq\,\mathrm{d}s$$

Let now $q \in \mathcal{P}_T$ chosen as

$$q = \Pi_h(\nabla\cdot(I_h v) - \nabla\cdot v).$$

Then, the left hand side of the equation above becomes $\|\nabla\cdot(I_h v) - \Pi_h(\nabla\cdot v)\|_T^2$. The first integral on the right vanishes by testing (4.29) with ∇q and by testing (4.41) with q, respectively. The same holds for the second integral using the node values in equations (4.28) and (4.40), respectively. Thus, we have proven $\nabla\cdot(I_h v) = \Pi_h(\nabla\cdot v)$. □

Remark 4.2.22. The next natural step is the extension of I_h to $H^{\operatorname{div}}(\Omega)$. Then, I_h would be our Fortin projection. Unfortunately, this is not possible, as the following example shows. Meanwhile, we note that the operator I_h is well-defined on the space $\tilde{V} = H^{\operatorname{div}}(\Omega) \cap H^s(\Omega;\mathbb{R}^d)$ for any $s > 0$. Thus, if the domain allows for an inf-sup condition of the form

$$\inf_{q\in Q}\sup_{v\in\tilde{V}} \frac{(\nabla\cdot v, q)}{\|v\|_{\tilde{V}}\|q\|_Q} \geq \beta > 0,$$

then we are done here. The case of minimal regularity will require us to extend the ideas of the Clément interpolant to commuting interpolation operators, which will be done in Chapter 6.2.

Example 4.2.23. The trace theorem involves the space $H^{-1/2}(\partial\Omega)$, which requires a short discussion. On one dimensional boundaries, elements in $H^{1/2}(\partial\Omega)$ have continuous representatives. The situation in three dimensions is similar, where no jumps across a line, for instance between two faces is allowed. Therefore, functions in $H^{-1/2}(\partial\Omega)$ cannot be localized to parts of the boundary, for instance the edge of a cell.

We give an example (modified from [BFB13, Section 2.5.1]) of this phenomenon. On the disc $\mathcal{D}$ around the origin of radius e^{-1} consider the function

$$u(x,y) = \ln\Big(-\ln\big(\sqrt{x^2+y^2}\big)\Big).$$

There holds $u \in H_0^1(\mathcal{D})$. Now, consider the domain Ω consisting only of the upper half circle:

$$\begin{aligned}\Omega &= \big\{(x,y)\in\mathbb{R}^2 \big| x^2+y^2 < e^{-2} \text{ and } y>0\big\}\\ \partial\Omega &= [-e^{-1},e^{-1}]\times\{0\}\cup\big\{(x,\sqrt{e^{-2}-x^2})\big| x\in(-e^{-1},e^{-1})\big\}\end{aligned}$$

Thus, the trace of u on the boundary is in $H^{1/2}(\partial\Omega)$. We now define $\mu \in H^{-1/2}(\partial\Omega)$ as the distributional derivative in tangential direction, say counterclockwise,

$$\langle\mu,\varphi\rangle = -\int_{\partial\Omega} u\partial_\tau\varphi \qquad \forall\varphi\in C^1(\partial\Omega).$$

Computation yields

$$\mu(x) = \frac{1}{x \ln|x|} \times \begin{cases} 1 & x \in (-e^{-1}, 0) \\ -1 & x \in (0, e^{-1}) \\ 0 & x \notin (-e^{-1}, e^{-1}). \end{cases}$$

Both integrals

$$\int_{-e^{-1}}^{0} \mu(x)\,\mathrm{d}x \qquad \int_{0}^{e^{-1}} \mu(x)\,\mathrm{d}x,$$

are not bounded, such that on these parts of the boundary, μ cannot even be tested with a constant function. Now, we consider the integral

$$\langle \mu, \varphi \rangle = \int_{-e^{-1}}^{e^{-1}} \ln(-\ln|x|)\partial_\tau \varphi\,\mathrm{d}x.$$

If we split φ into an odd and an even part, the integral with the even part vanishes, since $\ln(-\ln|x|)$ is even and the derivative is odd. Therefore,

$$\langle \mu, \varphi \rangle = \int_{-e^{-1}}^{e^{-1}} \frac{\varphi_{\mathrm{odd}}(x)}{x \ln|x|}\,\mathrm{d}x$$

Finally, we use the fact that $\varphi_{\mathrm{odd}}(0) = 0$ and that it's growth is limited by its regularity. In order to be square integrable, the growth of φ_{odd} must be limited by a positive, fractional power,

$$|\varphi_{\mathrm{odd}}(x)| \le c|x|^\alpha, \qquad \alpha > 0.$$

Then,

$$\int_{0}^{e^{-1}} \frac{|\varphi_{\mathrm{odd}}(x)|}{x \ln|x|}\,\mathrm{d}x \le \frac{x^{\alpha-1}}{\ln|x|}\,\mathrm{d}x < \infty.$$

4.2.24. Very much like the construction of Clément for $H^1(\Omega)$, we define interpolation operators stable on $H^{\mathrm{div}}(\Omega)$ by replacing the face integrals by volume integrals. We only have to make sure we do not destroy conformity with $H^{\mathrm{div}}(\Omega)$, in particular continuity of normal components. But, this can be achieved by integrating over both cells adjacent to a face and using this integral for interpolation.

4.2.25 Definition: An H^{div}-stable interpolation operator is obtained from Definition 4.2.19 of the canonical interpolation

1. by choosing the standard degrees of freedom for every cell integral as in (4.29), (4.41), and (4.42), and
2. by replacing every face integral as in (4.28) and (4.40) by integrals of the form

$$\int_F f\cdot n\,q\,\mathrm{d}s \to c_q \int_{\Omega_F} f\cdot n\,q\,\mathrm{d}x, \tag{4.48}$$

   where Ω_F consists of the cells sharing F and c_q is a normalization constant.

Summarizing the results from this section and applying the general theory, in particular Corollary 2.4.6 and Theorem 2.4.7, we obtain

4.2.26 Theorem: Let $V_h \subset V \subset H^{\mathrm{div}}(\Omega)$ and $Q_h \subset Q$ be chosen such that an H^{div}-stable commuting interpolation operator exists. Then, the solutions $(u,p) \in V\times Q$ and $(u_h,p_h) \in V_h\times Q_h$ admit the quasi-optimality estimates

$$\|u-u_h\|_{H^{\mathrm{div}}} \le c_1 \inf_{v\in V_h} \|u-v_h\|_{H^{\mathrm{div}}} \tag{4.49}$$

$$\|p-p_h\|_{L^2} \le c_2 \inf_{v\in V_h} \|u-v_h\|_{H^{\mathrm{div}}} + c_3 \inf_{q\in Q_h} \|p-q_h\|_{L^2}. \tag{4.50}$$

4.2.27 Corollary: The elements RT_k and BDM_{k+1} with their matching pressure space $\mathbb{P}_k$ admit the error estimates

$$\|u-u_h\|_{L^2} \le ch^{k+1}|u|_{H^{k+1,\mathrm{div}}} \tag{4.51}$$

$$\|\nabla\cdot u-\nabla\cdot u_h\|_{L^2} \le ch^{k+1}|u|_{H^{k+1,\mathrm{div}}} \tag{4.52}$$

$$\|p-p_h\|_{L^2} \le ch^{k+1}\big(|u|_{H^{k+1,\mathrm{div}}} + |p|_{H^{k+1}}\big), \tag{4.53}$$

where

$$|u|^2_{H^{k+1,\mathrm{div}}} = |u|^2_{H^{k+1}} + |\nabla\cdot u|^2_{H^{k+1}}. \tag{4.54}$$

4.2.4 Finite elements on quadrilaterals and hexahedra

4.2.28. Shape functions for quadrilaterals and hexahedra can only be defined on a reference cell. But, when mapping from a reference cell $\widehat{T}$ to the actual mesh cell T, we have to preserve the information whether a vector field is normal or tangential to a face. To this end, we recapitulate the basic notation and properties of the transformation of scalar fields and then add the definition for vector fields in $H^{\mathrm{div}}(\Omega)$.

4.2.29 Notation: Let $\widehat{T}$ be a reference cell, either the reference simplex spanned by $\{0, e_1, \dots, e_d\}$ or the reference hypercube $[-1,1]^d$. Then, a mesh cell $T \in \mathbb{T}_h$ is defined as the image of $\widehat{T}$ under a mapping Φ (we suppress the index T and understand that Φ is different for every cell). We define the Jacobi matrix, the Jacobi determinant, and the face Jacobian

$$\mathrm{D}\Phi(\widehat{x}) = \Big(\partial_j \Phi_i\Big), \qquad \mathrm{J}(\widehat{x}) = \det \mathrm{D}\Phi(\widehat{x}), \qquad \mathrm{J}_n(\widehat{x}) = J|\mathrm{D}\Phi^{-T}(\widehat{x})\widehat{n}|, \tag{4.55}$$

The basic relations are for $\widehat{x} \in \widehat{T}$ and shape functions $\widehat{p}$:

$$x = \Phi(\widehat{x}), \qquad p(x) = \widehat{p}(\widehat{x}) \qquad \nabla p(x) = \mathrm{D}\Phi^{-T}(\widehat{x})\widehat{\nabla}\widehat{p}(\widehat{x}). \tag{4.56}$$

Integrals transform as

$$\int_T p \,\mathrm{d}x = \int_{\widehat{T}} \widehat{p} J \,\mathrm{d}\widehat{x} \qquad \qquad \int_{\partial T} p \,\mathrm{d}s = \int_{\partial \widehat{T}} \widehat{p} J_n \,\mathrm{d}\widehat{s}.$$

4.2.30 Definition: The **Piola transform** or contravariant transformation of a vector field under the mapping $\Phi\colon \widehat{T} \to T$ is the mapping

$$v(x) = \tfrac{1}{\mathrm{J}} \mathrm{D}\Phi \widehat{v}(\widehat{x}). \tag{4.57}$$

There holds

$$\nabla v(x) = \tfrac{1}{\mathrm{J}} \mathrm{D}\Phi \big[\widehat{\nabla}\widehat{v}(\widehat{x})\big] \mathrm{D}\Phi^{-1}, \qquad \nabla \cdot v(x) = \tfrac{1}{\mathrm{J}} \widehat{\nabla} \cdot \widehat{v}(\widehat{x}). \tag{4.58}$$

4.2.31 Lemma: Let q be a scalar function and v be a contravariant vector field mapped by the Piola transform. Then, cell and surface integrals are transformed according to the rules

$$\begin{aligned} \int_T v \cdot \nabla q \,\mathrm{d}x &= \int_{\widehat{T}} \widehat{v} \cdot \widehat{\nabla}\widehat{q} \,\mathrm{d}\widehat{x}, \\ \int_T q \nabla \cdot v \,\mathrm{d}x &= \int_{\widehat{T}} \widehat{q} \widehat{\nabla} \cdot \widehat{v} \,\mathrm{d}\widehat{x}, \\ \int_{\partial T} q v \cdot n \,\mathrm{d}s &= \int_{\partial \widehat{T}} \widehat{q} \widehat{v} \cdot \widehat{n} \,\mathrm{d}\widehat{s}. \end{aligned} \tag{4.59}$$

4.2.32 Problem: Verify Lemma 4.2.31.

Remark 4.2.33. The last equation of the previous lemma indicates, that the Piola transform preserves normal components of a vector field. Thus, it can be used to define shape functions for normal continuity on a reference cell.

4.2.34 Notation: The space of tensor product polynomials $\mathbb{Q}_k$ in d space dimensions is

$$\begin{aligned} \mathbb{Q}_k(\mathbb{R}^d) &= \underbrace{\mathbb{P}_k(\mathbb{R}) \otimes \cdots \otimes \mathbb{P}_k(\mathbb{R})}_{d \text{ factors}}, \\ q(x_1, \dots, x_d) &= \prod_{i=1}^d p_i(x_i) \qquad p_i \in \mathbb{P}_k(\mathbb{R}). \end{aligned} \tag{4.60}$$

Similarly, we define anisotropic tensor product polynomials

$$\begin{aligned} \mathbb{Q}_{k1,\dots,k_d}(\mathbb{R}^d) &= \underbrace{\mathbb{P}_{k_1}(\mathbb{R}) \otimes \cdots \otimes \mathbb{P}_{k_d}(\mathbb{R})}_{d \text{ factors}}, \\ q(x_1, \dots, x_d) &= \prod_{i=1}^d p_i(x_i) \qquad p_i \in \mathbb{P}_{k_i}(\mathbb{R}). \end{aligned} \tag{4.61}$$

4.2.35 Definition: The **Raviart-Thomas element** of degree $k \ge 0$ on the reference cell $\widehat{T} = [-1,1]^d$ consists of the polynomial space

$$RT_{[k]}(\widehat{T}) = \mathbb{Q}_k^d(\widehat{T}) + x\mathbb{Q}_k(\widehat{T}). \tag{4.62}$$

Its node functionals are

$$\mathcal{N}_{1,i,j}(v) = \int_{F_i} v \cdot n\, q_j \,\mathrm{d}s \qquad q_j \in \mathbb{Q}_k(F_i) \qquad F_i \subset \partial\widehat{T}, \tag{4.63}$$

$$\mathcal{N}_{2,i}(v) = \int_{\widehat{T}} v \cdot w_i \,\mathrm{d}x \qquad w_i \in \mathbb{Q}_{k-1,k\dots k} \times \cdots \times \mathbb{Q}_{k\dots k,k-1}. \tag{4.64}$$

4.2.36 Lemma: There holds

$$\dim RT_{[k]} = d(k+1)^{d-1}(k+2), \tag{4.65}$$

and

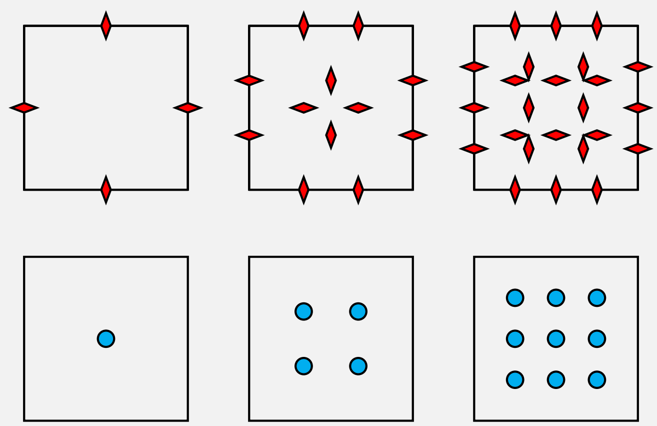
$$\nabla\cdot RT_{[k]} = \mathbb{Q}_k. \tag{4.66}$$

Furthermore, for each $F \subset \widehat{T}$ and each $v \in RT_{[k]}(\widehat{T})$there holds

$$v \cdot n|_F \in \mathbb{Q}_k. \tag{4.67}$$

Proof. The proof of this lemma is exactly the same as the one of Lemma 4.2.11. □

4.2.37 Example: The first members of the Raviart-Thomas family on quadrilaterals are



The construction principle of the Raviart-Thomas element on simplices as well as on rectangles and cubes can be seen as adding vector polynomials to the velocity space until its divergence is equal to $\mathbb{P}_k$ or $\mathbb{Q}_k$. The principle of the BDM elements is the opposite: Starting from the polynomial spaces for triangles, we add divergence free shape functions until we achieve continuity of the normal component over all edges. We give their definitions on squares and cubes.

4.2.38 Definition: The BDM element of degree $k \geq 1$ on the reference cell $\widehat{T} = [-1,1]^2$ consists of the polynomial space

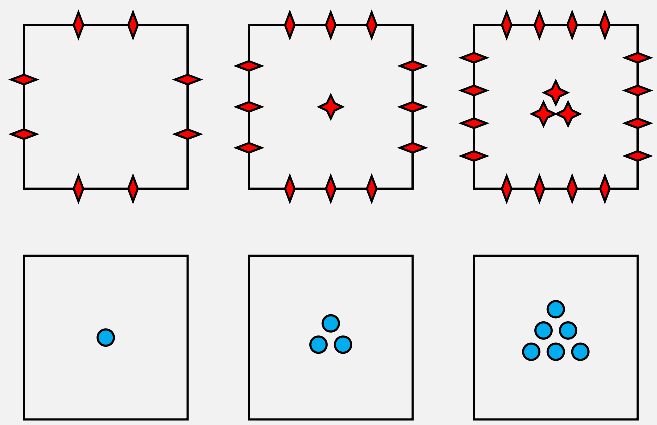
$$BDM_{[k]} = \mathbb{P}_k^2 \oplus \operatorname{span}\bigl\{\nabla\times(x^{k+1}y), \nabla\times(xy^{k+1}\bigr\}. \tag{4.68}$$

Its node functionals are

$$\mathcal{N}_{1,i,j}(v) = \int_{F_i} v\cdot n\, q_j \,\mathrm{d}s \qquad q_j \in \mathbb{P}_k(F_i) \qquad F_i \subset \partial\widehat{T}, \tag{4.69}$$

$$\mathcal{N}_{2,i}(v) = \int_{\widehat{T}} v\cdot w_i \,\mathrm{d}x \qquad w_i \in \mathbb{P}_{k-2}^2. \tag{4.70}$$

4.2.39 Example: The first members of the Brezzi-Douglas-Marini family on quadrilaterals are



4.2.40 Lemma: The dimension of the space $BDM_{[k]}$ is

$$\dim BDM_{[k]} = (k+1)(k+2) + 2. \tag{4.71}$$

The element in Definition 4.2.38 is unisolvent.

4.2.41 Problem: Prove Lemma 4.2.40.

4.2.42 Corollary: Let the mesh be such that each cell is obtained by affine transformation from the reference cell $\widehat{T}$. Then, Theorem 4.2.26 applies and we obtain quasi-bestapproximation.

Remark 4.2.43. If the mapping of the cells is not affine, we do not have

$$\nabla\cdot V_h = Q_h.$$

Indeed, on each cell, we have

$$\nabla\cdot V_h = \tfrac{1}{J} Q_h, \tag{4.72}$$

where the Jacobi determinant J is not constant. As a consequence, Lemma 4.2.21 about the commuting diagram property does not apply directly anymore and indeed, approximation may suffer [ABF05]. In particular, it is shown there that for the $RT_{[k]}$ on general quadrilateral meshes, which do *not* converge to affine meshes as $h \to 0$, there holds

$$\inf_{v_h\in V_h} \|u - v_h\| = \mathcal{O}(h^{k+1}),$$
$$\inf_{v_h\in V_h} \|\nabla\cdot u - \nabla\cdot v_h\| = \mathcal{O}(h^{k}).$$

The optimal approximation of the divergence can be recovered by enriching the space like the Arnold-Boffi-Falk element below.

4.2.44 Definition: The **Arnold-Boffi-Falk element** of degree $k \geq 0$ on the reference cell $\widehat{T} = [-1,1]^2$ consists of the polynomial space

$$ABF_k(\widehat{T}) = \mathbb{Q}_{k+2,k} \times \mathbb{Q}_{k,k+2} \tag{4.73}$$

Its node functionals are

$$\mathcal{N}_{1,i,j}(v) = \int_{F_i} v\cdot n\, q_j \,\mathrm{d}s \qquad q_j \in \mathbb{Q}_k(F_i) \qquad F_i \subset \partial\widehat{T}, \tag{4.74}$$

$$\mathcal{N}_{2,i}(v) = \int_{\widehat{T}} v\cdot w_i \,\mathrm{d}x \qquad w_i \in \mathbb{Q}_{k-1,k} \times \cdots \times \mathbb{Q}_{k,k-1}, \tag{4.75}$$

$$\mathcal{N}_{3,x,i}(v) = \int_{\widehat{T}} \nabla\cdot v(x^i y^{k+1}) \qquad i = 1,\dots,k, \tag{4.76}$$

$$\mathcal{N}_{3,y,i}(v) = \int_{\widehat{T}} \nabla\cdot v(x^{k+1} y^i) \qquad i = 1,\dots,k. \tag{4.77}$$

Remark 4.2.45. Degrees of freedom in this section have been written as moments with respect to polynomials in $\mathbb{P}_k$ or $\mathbb{Q}_k$, which is natural in this context and allows for an easy proof of the commutating diagram property of the Fortin projection. On the other hand, degrees of freedom based on point interpolation are sometimes more natural for the implementation.

In one dimension, for instance for the integration over edges, we realize that

$$\mathcal{N}_{1,i,j}(v) = \int_F v\cdot n\, q_j \,\mathrm{d}s = \sum_{\ell=1}^{k+1} \omega_\ell v(x_\ell)\cdot n\, q(x_\ell),$$

for any Gauss-Legendre or Gauss-Lobatto quadrature rule on F. What was left unspecified in the definition of the element was the choice of a basis $\{q_j\}$ for $\mathbb{P}_k$. From the point of view of moments, it is natural to choose Legendre polynomials as a basis. But, we can also choose the basis which is orthogonal with respect to the quadrature rule, which is up to the weights a Lagrange basis. Thus, we can transform the moment degrees of freedom back to interpolating degrees of freedom easily.

This construction extends automatically to tensor product space $\mathbb{Q}_k$. For polynomials space $\mathbb{P}_k$, suitable quadrature sets on triangles and quadrilaterals must be constructed.

Chapter 5

Divergence conforming discontinuous Galerkin methods

5.0.1. In the previous chapter, we studied discretizations with $\nabla\cdot V_h = Q_h$ with two advantages. First, due to Corollary 2.4.6 the velocity error is independent of the pressure. Second, the divergence converges faster than the gradient. A natural question arising is whether we can do something similar for the Stokes problem. There, the equation

$$(\nabla\cdot v_h, q_h) = 0 \qquad \forall q_h \in Q_h,$$

would immediately imply $\nabla\cdot v_h = 0$, that is, the discrete solution is exactly divergence free.

The answer to this question is a current research topic. So far, beginning with the element by Scott and Vogelius, several methods have been proposed for special mesh geometry or macro meshes. The difficulty is balancing the condition $\nabla\cdot V_h = Q_h$ with the H^1-conformity of the velocity space. All the spaces in the previous chapter were only H^{div}-conforming with discontinuous tangential components.

A fairly simple solution to this question though can be obtained by using discontinuous Galerkin methods. These were introduced to obtain formulations *consistent* with H^1 while not *conforming*. Thus, we can apply them directly to Raviart-Thomas and Brezzi-Douglas-Marini elements to obtain a consistent method with divergence free solutions.

We begin this chapter by a quick review of the interior penalty method before diving into divergence conforming methods.

5.1 The interior penalty method

5.1.1. We review the basic definitions necessary to describe discontinuous Galerkin (DG) methods. In particular, we need the sets of faces $\mathbb{F}_h$ of a mesh, discontinuous piecewise polynomial spaces and broken integrals.

5.1.2 Definition: Let $\mathbb{T}_h$ be a mesh of $\Omega \subset \mathbb{R}^d$ consisting of mesh cells T_i. For every boundary facet $F \subset \partial T_i$, we assume[a] that either $F \subset \partial\Omega$ or F is a boundary facet of another cell T_j. In the second case, we indicate this relation by labeling this facet F_{ij}. The set of all facets F_{ij} is the set of interior faces $\mathbb{F}_h^i$. The set of facets on the boundary is $\mathbb{F}_h^\partial$.

[a]This assumption can indeed be relaxed

5.1.3 Definition: The discontinuous finite element space on $\mathbb{T}_h$ is constructed by concatenation of all shape function spaces P_T for $T \in \mathbb{T}_h$ without additional continuity requirements:

$$V_h = \left\{ v \in L^2(\Omega) \middle| v_{|T} \in P_T \; \forall T \in \mathbb{T}_h \right\}. \tag{5.1}$$

5.1.4 Definition: For any set of cells $\mathbb{T}_h$ or faces $\mathbb{F}_h$, we define the bilinear forms

$$(u, v)_{\mathbb{T}_h} = \sum_{T \in \mathbb{T}_h} (u, v)_T, \tag{5.2}$$

$$\langle u, v \rangle_{\mathbb{F}_h} = \sum_{F \in \mathbb{F}_h} \langle u, v \rangle_F. \tag{5.3}$$

$$\tag{5.4}$$

5.1.5. We start out with the equation

$$-\Delta u = f.$$

Integrating by parts on each mesh cell yields

$$(-\Delta u, v)_T = (\nabla u, \nabla v)_T - \langle \partial_n u, v \rangle_{\partial T} = (f, v)_T.$$

We realize that the choice of discontinuous finite element spaces introduces a consistency term on the interfaces between cells and on the boundary.

On interior faces, there is the issue that u and $\partial_n u$ actually have two values on the interface, one from the left cell and one from the right. Therefore, we have to consolidate these two values into one. To this end, we introduce the concept of a numerical flux, which constructs a single value out of these two. Thus, we introduce on the interface F between two cells T^+ and T^-

$$\mathcal{F}(\nabla u) = \frac{\nabla u^+ + \nabla u^-}{2} =: \{\!\{\nabla u\}\!\}.$$

Using $\langle \partial_n u, v\rangle = \langle \nabla u, vn\rangle$ we change our point of view and instead of integrating over the boundary ∂T, we integrate over a face F between two cells T^+ and T^-. Adding up integrals from both sides, we obtain the term

$$-\big\langle \{\!\{\nabla u\}\!\}, v^+n^+ + v^-n^- \big\rangle_F = -2\langle \{\!\{\nabla u\}\!\}, \{\!\{vn\}\!\}\rangle_F.$$

On boundary faces, we simply get

$$\langle \partial_n u, v\rangle_F.$$

Adding over all cells and faces, we obtain the equation

$$(\nabla u, \nabla v)_{\mathbb{T}_h} - 2\langle \{\!\{\nabla u\}\!\}, \{\!\{vn\}\!\}\rangle_{\mathbb{F}_h^i} - \langle \partial_n u, v\rangle_{\mathbb{F}_h^\partial} = (f, v)_\Omega.$$

Following the idea of Nitsche, we symmetrize this term to obtain

$$\begin{aligned}(\nabla u, \nabla v)_{\mathbb{T}_h} - 2\langle \{\!\{\nabla u\}\!\}, \{\!\{vn\}\!\}\rangle_{\mathbb{F}_h^i} - 2\langle \{\!\{un\}\!\}, \{\!\{\nabla v\}\!\}\rangle_{\mathbb{F}_h^i}& \\ - \langle \partial_n u, v\rangle_{\mathbb{F}_h^\partial} - \langle u, \partial_n v\rangle_{\mathbb{F}_h^\partial} = (f, v)_\Omega - \langle u^o, \partial_n v\rangle_{\mathbb{F}_h^\partial}&.\end{aligned}$$

Here the second term on the right was introduced for consistency. Finally, it turns out that this method is not stable and needs stabilization by a jump term. This will be done in Definition 5.1.8. Before, we introduce the notation for averaging and jump operators.

5.1.6 Notation: Let F be a face between the cells T^+ and T^-. Let n^+ and $n^- = -n^+$ be the outer normal vectors of the cells at a point $x \in F$. For a function $u \in V_h$, the traces u^+ and u_- of u on F taken from the cell T^+ and T^- are defined as:

$$\begin{aligned} u^+(x) &= \lim_{\varepsilon \searrow 0} u(x - \varepsilon n^+), \\ u^-(x) &= \lim_{\varepsilon \searrow 0} u(x - \varepsilon n^-). \end{aligned}$$

We define the **averaging operator** $\{\!\{.\}\!\}$ and the **jump operator** $[\![.]\!]$ as

$$\{\!\{u\}\!\} = \frac{u^+ + u^-}{2}, \qquad [\![u]\!] = u^+ - u^-. \tag{5.5}$$

Not that the sign of the jump of u depends on the choice of the cells T^+ and T^-. It will only be used in quadratic terms.

Remark 5.1.7. The jump can be denoted as the mean value of the product of a function and the normal vector,

$$[\![u]\!] = 2\{\!\{un\}\!\} \cdot n^+ = -2\{\!\{un\}\!\} \cdot n^-. \tag{5.6}$$

5.1.8 Definition: The **interior penalty method**[a] uses the bilinear form

$$\begin{aligned} a_h(u,v) = (\nabla u, \nabla v)_{\mathbb{T}_h} &+ \langle \sigma_h [\![u]\!], [\![v]\!] \rangle_{\mathbb{F}_h^i} + \langle \sigma_h u, v \rangle_{\mathbb{F}_h^\partial} \\ &- 2\langle \{\!\{\nabla u\}\!\}, \{\!\{vn\}\!\} \rangle_{\mathbb{F}_h^i} - 2\langle \{\!\{un\}\!\}, \{\!\{\nabla v\}\!\} \rangle_{\mathbb{F}_h^i} \\ &- \langle \partial_n u, v \rangle_{\mathbb{F}_h^\partial} - \langle u, \partial_n v \rangle_{\mathbb{F}_h^\partial}, \end{aligned} \tag{5.7}$$

and the linear form

$$f_h(v) = (f,v)_\Omega - \left\langle u^D, \partial_n v \right\rangle_{\mathbb{F}_h^\partial} + \left\langle \sigma_h u^D, v \right\rangle_{\mathbb{F}_h^\partial}, \tag{5.8}$$

where f is the right hand side of the equation and u^D the Dirichlet boundary value.

[a] Also known as symmetric interior penalty (SIPG) or IP-DG.

5.1.9 Definition: On the space V_h we define the norm $\|.\|_{1,h}$ by

$$\|v\|_{1,h}^2 = \sum_{T \in \mathbb{T}_h} \|\nabla v\|_T^2 + \sum_{F \in \mathbb{F}_h^i} \|\sqrt{\sigma_h}\, [\![v]\!]\|_F^2 + \sum_{F \in \mathbb{F}_h^\partial} \|\sqrt{\sigma_h} v\|_F^2. \tag{5.9}$$

5.1.10 Problem: Prove that the norm defined in (5.9) is indeed a norm on V_h.

5.1.11 Lemma: Let $\mathbb{T}_h$ be shape-regular and chosen on each face F as $\sigma_h = \sigma_0/h_F$, where h_T is the minimal diameter of a cell adjacent to F. Then, there is a $\sigma_0 > 0$ such that there exists a constant $\gamma > 0$, such that independent of h there holds

$$a_h(u_h, u_h) \geq \gamma \|u_h\|_{1,h}^2 \quad \forall u_h \in V_h. \tag{5.10}$$

5.1.12 Problem: Prove Lemma 5.1.11.

5.1.13 Lemma: Let $f \in L^2(\Omega)$ and let the boundary conditions admit that for the solution to

$$\begin{aligned} -\Delta u &= f && \text{in } \Omega, \\ u &= u^D && \text{on } \partial\Omega, \end{aligned}$$

there holds $u \in H^{1+\varepsilon}(\Omega)$ for a positive ε. Then, the interior penalty method is consistent, that is,

$$a_h(u, v_h) = f_h(v_h) \quad \forall v_h \in V_h. \tag{5.11}$$

Proof. From $f \in L^2(\Omega)$ we deduce that $\nabla u \in H^{\text{div}}(\Omega)$. Thus, with the extra regularity, the traces of $\partial_n u$ on faces are well-defined and coincide from both sides. The remainder is integration by parts. □

5.1.14 Theorem: For $k \geq 1$ let $\mathbb{P}_k \subset P_T$ and $u \in H^{s+1}(\Omega)$ with $1/2 \leq s \leq k$. Then, the interior penalty method admits the error estimate

$$\|u - u_h\|_{1,h} \leq ch^s |u|_{s+1}. \tag{5.12}$$

If furthermore the boundary condition admits elliptic regularity, there holds

$$\|u - u_h\|_0 \leq ch^{s+1} |u|_{s+1}. \tag{5.13}$$

5.1.1 Bounded formulation in H^1

5.1.15. The interior penalty method introduced so far is V_h-elliptic and consistent, but it is not bounded on $H^1(\Omega)$. This was a reason, why we could not use standard techniques for the proof of the convergence result and after applying consistency had to estimate each term separately.

In this section, we will introduce a reformulation of the interior penalty method, which is equivalent to the original method on V_h, but is also bounded in $H^1(\Omega)$. As an unpleasant side effect, it turns out that this method is inconsistent, and we have to estimate the consistency error.

The main technique applied here is the use of lifting operators, such that the traces of derivatives on faces can be replaced by volume terms. Note that the lifting operators, while very useful for the analysis of the method, are not actually used in the implementation of the interior penalty method.

5.1.16 Definition: Define the auxiliary space

$$\Sigma_h = \left\{\tau \in L^2(\Omega;\mathbb{R}^d) \middle| \forall T \in \mathbb{T}_h : \tau_{|T} \in \Sigma_T\right\}, \tag{5.14}$$

where Σ_T is a (possibly mapped) polynomial space chosen such that $\nabla V_T \subset \Sigma_T$. Then, we define the **lifting operator**

$$\mathscr{L}\colon V + V_h \to \Sigma_h \tag{5.15}$$

by

$$(\mathscr{L}v, \tau)_{\mathbb{T}_h} = 2\langle\{\!\{\tau\}\!\}, \{\!\{vn\}\!\}\rangle_{\mathbb{F}_h^i} + \langle \tau \cdot n, v\rangle_{\mathbb{F}_h^\partial}. \tag{5.16}$$

5.1.17 Lemma: The lifting operator is a bounded operator from $L^2(\mathbb{F}_h)$ to Σ_h, such that

$$\|\mathscr{L}v\|_{L^2(\Omega)} \le c\left\|\frac{1}{\sqrt{h}}\left[\!\left[v\right]\!\right]\right\|_{\mathbb{F}_h^i} + \left\|\frac{1}{\sqrt{h}}v\right\|_{\mathbb{F}_h^\partial}. \tag{5.17}$$

In particular, it is bounded on $H^1(\Omega)$.

Proof. It is clear, that the operator is bounded on $L^2(\mathbb{F}_h)$, since its definition involves face integrals weighted with polynomial functions. The dependence on the mesh size is due to the standard scaling argument. □

5.1.18 Definition: The **interior penalty method** with lifting operators uses the bilinear form

$$\begin{aligned} a_h(u,v) = (\nabla u, \nabla v)_{\mathbb{T}_h} - (\mathscr{L}u, \nabla v)_{\mathbb{T}_h} - (\nabla u, \mathscr{L}v)_{\mathbb{T}_h} \\ + \langle \sigma_h \left[\!\left[u\right]\!\right], \left[\!\left[v\right]\!\right]\rangle_{\mathbb{F}_h^-} + \langle \sigma_h u, v\rangle_{\mathbb{F}_h^\partial}. \end{aligned} \tag{5.18}$$

and the linear form (5.8) of the original interior penalty method. Its residual operator is

$$\operatorname{Res}(u,v) = a_h(u,v) - (f,v). \tag{5.19}$$

5.1.19 Lemma: The interior penalty method in flux form (Definition 5.1.8) and in lifting form (Definition 5.1.8) coincide on the discrete space V_h if Σ_h is chosen such that $\nabla V_h \subset \Sigma_h$.

Proof. Since $\nabla V_h \subset \Sigma_h$, ∇u_h and ∇v_h are valid test functions in the definition (5.16) of the lifting operator, and the equality

$$(\mathscr{L} u_h, \nabla v_h)_{\mathbb{T}_h} = 2\langle \{\!\{u_h n\}\!\}, \{\!\{\nabla v_h\}\!\}\rangle_{\mathbb{F}_h^i} + \langle u_h, \partial_n v_h\rangle_{\mathbb{F}_h^\partial}.$$

□

5.1.20 Definition: Let $V \subset H^1(\Omega)$ and let $u, u^* \in V$ solve the primal and dual problems

$$a(u,v) = f(v), \qquad a(v,u^*) = \psi(v), \qquad \forall v \in V, \tag{5.20}$$

with a bounded, V-elliptic bilinear form $a(.,.)$. For a discrete bilinear form $a_h(.,.)$ defined on $V + V_h$, we define the primal and dual **residual operator**s

$$\begin{aligned} \operatorname{Res}(u,v) &= a_h(u,v) - f(v), \\ \operatorname{Res}^*(u^*,v) &= a_h(v,u^*) - \psi(v). \end{aligned} \tag{5.21}$$

5.1.21 Lemma: Let $a_h(.,.)$ be a bounded bilinear form on $V + V_h$ and elliptic on V_h with norm $\|.\|_{V_h}$ and constant γ. Then, the error $u - u_h$ admits the estimate

$$\|u-u_h\|_{V_h} \le \frac{1}{\gamma}\|\operatorname{Res}(u,.)\|_{V_h^*} + \left(1 + \frac{\|a_h\|}{\gamma}\right) \inf_{w_h \in V_h} \|u - w_h\| \tag{5.22}$$

Proof. First, by the definition of the residual, we have the error equation

$$a_h(u-u_h, v_h) = \operatorname{Res}(u, v_h), \qquad \forall v_h \in V_h. \tag{5.23}$$

Inserting $w_h - w_h$ for an arbitrary element $w_h \in V_h$, we obtain

$$a_h(w_h - u_h, v_h) = \operatorname{Res}(u, v_h) - a_h(u - w_h, v_h), \qquad \forall v_h \in V_h.$$

Using $v_h = w_h - u_h$ and ellipticity, we obtain

$$\begin{aligned} \gamma\|w_h - u_h\|_{V_h}^2 &\le a_h(w_h - u_h, w_h - u_h) \\ &= \operatorname{Res}(u, w_h - u_h) - a_h(u - w_h, w_h - u_h) \\ &\le \big(\|\operatorname{Res}(u,.)\|_{V_h^*} + \|a_h\|\|u - w_h\|_{V_h}\big)\|w_h - u_h\|_{V_h}. \end{aligned}$$

Hence, by triangle inequality

$$\|u-u_h\|_{V_h} \le \frac{1}{\gamma}\|\operatorname{Res}(u,.)\|_{V_h^*} + \left(1 + \frac{\|a_h\|}{\gamma}\right) \inf_{w_h \in V_h} \|u - w_h\|_{V_h}$$

□

5.1.22 Lemma: Let $u \in V$ be the solution to the Poisson equation with right hand side $f \in L^2(\Omega)$. Assume $u \in H^s(\Omega)$ with $s > 3/2$. Then, we have for $v \in V + V_h$:

$$(f, v) = (\nabla u, \nabla v)_{\mathbb{T}_h} - 2\langle \nabla u, \{\!\{vn\}\!\}\rangle_{\mathbb{F}_h^i} - \langle \partial_n u, v\rangle_{\mathbb{F}_h^\partial}. \tag{5.24}$$

Proof. We set out from the strong form of the Poisson equation and integrate by parts.

$$(f, v) = (-\Delta u, v) = (\nabla u, \nabla v)_{\mathbb{T}_h} - \sum_{T\in\mathbb{T}_h} \langle \partial_n u, v\rangle_{\partial T}.$$

Under the regularity assumptions of the lemma, all of these integrals make sense at least as duality pairings. In particular, $\partial_n u \in L^2(\partial T)$, and thus we can split ∂T into individual faces. Therefore,

$$\sum_{T\in\mathbb{T}_h} \langle \partial_n u, v\rangle_{\partial T} = 2\langle \nabla u, \{\!\{v \otimes n\}\!\}\rangle_{\mathbb{F}_h^i} + \langle \partial_n u, v\rangle_{\mathbb{F}_h^\partial}.$$

The proof concludes by collecting the results. □

5.1.23 Lemma: Let $k \geq 1$ and let V_h such that $\mathbb{P}_{k-1} \subset \Sigma_T$. Then, if $u \in H^{k+1}(\Omega)$ and $v \in V + V_h$, there holds

$$\begin{aligned} |\mathrm{Res}(u, v)| &\leq ch^k |u|_{k+1} \big(\|\sqrt{\sigma_h}\, [\![v]\!]\|_{\mathbb{F}_h^i} + \|\sqrt{\sigma_h} v\|_{\mathbb{F}_h^\partial} \big) \\ &\leq ch^k |u|_{k+1} \|v\|_{1,h}. \end{aligned} \tag{5.25}$$

Proof. First, we observe that by the regularity assumption, $[\![u]\!] = 0$ and thus, $\mathscr{L}u = 0$. Hence,

$$a_h(u, v) = (\nabla u, \nabla v)_{\mathbb{T}_h} - (\nabla u, \mathscr{L}v)_{\mathbb{T}_h}.$$

By Lemma 5.1.22 and regularity of u,

$$\begin{aligned} \mathrm{Res}(u, v) &= 2\langle \nabla u, \{\!\{vn\}\!\}\rangle_{\mathbb{F}_h^i} + \langle \partial_n u, v\rangle_{\mathbb{F}_h^\partial} - (\nabla u, \mathscr{L}v)_{\mathbb{T}_h} \\ &= 2\langle \{\!\{\nabla u\}\!\}, \{\!\{vn\}\!\}\rangle_{\mathbb{F}_h^i} + \langle \partial_n u, v\rangle_{\mathbb{F}_h^\partial} - (\nabla u, \mathscr{L}v)_{\mathbb{T}_h} \\ &= 2\langle \{\!\{\nabla u\}\!\}, \{\!\{vn\}\!\}\rangle_{\mathbb{F}_h^i} + \langle \partial_n u, v\rangle_{\mathbb{F}_h^\partial} - (\Pi_{\Sigma_h} \nabla u, \mathscr{L}v)_{\mathbb{T}_h}, \end{aligned}$$

where Π_{Σ_h} is the L^2-projection. Now, we can apply the definition of the lifting term to obtain

$$\begin{aligned} \mathrm{Res}(u, v) = 2\Big\langle \tfrac{1}{\sigma_h} \{\!\{\nabla u - \Pi_{\Sigma_h} \nabla u\}\!\}, \sigma_h \{\!\{vn\}\!\} \Big\rangle_{\mathbb{F}_h^i} \\ + \Big\langle \tfrac{1}{\sigma_h} (\nabla u - \Pi_{\Sigma_h} \nabla u) \cdot n, \sigma_h v \Big\rangle_{\mathbb{F}_h^\partial}. \end{aligned}$$

Application of standard approximation and trace estimates yields the result observing that $\sigma_h = \sigma_0/h$. $\square$

5.1.24 Theorem: Let $k \geq 1$ and V_h such that $\mathbb{P}_k \subset V_T$. Let $u \in H^{k+1}(\Omega)$ be the solution to the continuous Poisson problem. Let $a_h(.,.)$ be the interior penalty method with lifting operators such that $\nabla V_h \subset \Sigma_h$. Then, there holds

$$\|u - u_h\|_{1,h} \leq ch^k |u|_{k+1}. \tag{5.26}$$

Proof. Application of Lemma 5.1.21, Lemma 5.1.23, and standard interpolation results. $\square$

5.1.25 Theorem: Let the assumptions of Theorem 5.1.24 hold and in addition assume that the problem

$$a(v, u^*) = \psi(v), \qquad \forall v \in V,$$

admits the elliptic regularity estimate

$$\|u^*\|_{H^2(\Omega)} \leq c\|\psi\|_{L^2(\Omega)}. \tag{5.27}$$

Then, there holds

$$\|u - u_h\|_{L^2(\Omega)} \leq ch^{k+1} |u|_{H^{k+1}(\Omega)}. \tag{5.28}$$

Proof. The proof uses the duality argument by Aubin and Nitsche, which sets out solving the auxiliary problem

$$a(v, u^*) = (u - u_h, v), \qquad \forall v \in V.$$

Using the definition of the dual residual, we obtain the equation

$$(u - u_h, v) = a_h(v, u^*) - \mathrm{Res}^*(u^*, v), \qquad \forall v \in V + V_h.$$

Testing with $v = u - u_h$ yields

$$\|u - u_h\|^2 = a_h(u - u_h, u^*) - \mathrm{Res}^*(u^*, u - u^*).$$

Additionally, we us the error equation

$$a_h(u - u_h, v_h) = \mathrm{Res}(u, v_h),$$

tested with $v_h = I_h u^*$, to obtain

$$\|u - u_h\|^2 = a_h(u - u_h, u^* - I_h u^*) - \operatorname{Res}^*(u^*, u - u_h) + \operatorname{Res}(u, I_h u^*).$$

Using the regularity of u^*, the first term on the right admits the estimate

$$|a_h(u - u_h, u^* - I_h u^*)| \le \|u - u_h\|_{1,h} \|u^* - I_h u^*\|_{1,h} \le ch\|u - u_h\|_{1,h}.$$

For the second term, we use Lemma 5.1.23 to obtain

$$|\operatorname{Res}^*(u^*, u - u_h)| \le ch|u^*|_2 \|u - u_h\|_{1,h}.$$

Finally, using $[\![u^*]\!] = 0$, the same lemma yields

$$\begin{aligned}
|\operatorname{Res}(u, I_h u^*)| &\le ch|u|_2 \big(\|\sqrt{\sigma_h}\,[\![I_h u^*]\!]\|_{\mathbb{F}_h^i} + \|\sqrt{\sigma_h} I_h u^*\|_{\mathbb{F}_h^\partial}\big) \\
&= ch|u|_2 \big(\|\sqrt{\sigma_h}\,[\![u^* - I_h u^*]\!]\|_{\mathbb{F}_h^i} + \|\sqrt{\sigma_h}(u^* - I_h u^*)\|_{\mathbb{F}_h^\partial}\big) \\
&\le ch|u|_2 h^k |u^*|_{k+1}
\end{aligned}$$

Using the energy estimate in Theorem 5.1.24 we can conclude the prove. □

5.2 Divergence conforming IP

Remark 5.2.1. The extension of the interior penalty method to vector-valued problems is obvious. Furthermore, since the method generates an elliptic bilinear form on the discontinuous space V_h, this ellipticity is inherited by any subspace of $V_h \cap H^{\text{div}}(\Omega)$. Thus, we can write down the weak formulation of a divergence conforming DG method for the Stokes equations. In the following definition, we assume slip or no-slip boundary conditions, that is, $v \cdot n = 0$ on the whole boundary.

5.2.2 Definition: A divergence conforming DG method for the Stokes equations consists of a discrete velocity space $V_h \subset H_0^{\text{div}}(\Omega)$ and a pressure space $Q_h \subset L_0^2(\Omega)$ such that

$$\nabla \cdot V_h = Q_h. \tag{5.29}$$

Using the interior penalty bilinear form $a_h(.,.)$, we search for solutions $(u_h, p_h) \in V_h \times Q_h$ such that for all $(v, q) \in V_h \times Q_h$ there holds

$$a_h(u_h, v) + (\nabla \cdot v, p_h) + (\nabla \cdot u_h, q) = f(v). \tag{5.30}$$

Remark 5.2.3. Due to the fact that $V_h \not\subset V$, we have introduced the norm $\|.\|_{1,h}$ on V_h. In particular, the norm $\|.\|_1$ is not defined for all elements of V_h. Therefore, we need a modification of Fortin's lemma (Lemma 2.4.12), where the norm on the left hand side of the stability estimate (2.53) uses the discrete norm, namely,

$$\|\Pi_{V_h} v\|_{V_h} \le c\|v\|_V,$$

5.2.4 Lemma: Let $\{\mathbb{T}_h\}$ be a shape-regular sequence of meshes. Then, the canonical interpolation operators of the Brezzi-Douglas-Marini and Raviart-Thomas elements admit the bound

$$\|I_h v\|_{1,h} \le c|v|_1 \tag{5.31}$$

Proof. First, we note that all degrees of freedom are defined as cell or face integrals with smooth weight functions. Thus, they are bounded on H^1. Thus, since the local polynomial spaces are finite dimensional, there holds on the reference cell $\widehat{T}$ and its faces $\widehat{F}$:

$$\|I_{\widehat{T}} v\|_{1;\widehat{T}} \le c|v|_{1;\widehat{T}},$$
$$\|I_{\widehat{T}} v\|_{0;\widehat{F}} \le c|v|_{1;\widehat{T}}.$$

On shape regular meshes, we have the scaling property

$$|f|_{m;T} \simeq h_T^{\frac{d}{2}-m},$$
$$|f|_{m;F} \simeq h_F^{\frac{d-1}{2}-m},$$

such that for a face F of cell T

$$\|I_T v\|_{1;T} \le c|v|_{1;T},$$
$$\|I_T v\|_{0;\widehat{F}} \le c h^{\frac{1}{2}} |v|_{1;T}.$$

We conclude

$$\|I_h v\|_{1,h}^2 \le \sum_{T\in\mathbb{T}_h} \left[\|I_T v\|_{1;T}^2 + 4 \sum_{F\subset\partial T} \left\| \tfrac{\sigma_0}{h_F} I_T v \right\|_{0;F}^2 \right] \le c|v|_1^2.$$

□

5.2.5 Corollary: Assume that the inf-sup condition (3.1) in Theorem 3.1.8 holds. Then, the method in Definition 5.2.2 admits the inf-sup condition

$$\inf_{q_h \in Q_h} \sup_{v_h \in V_h} \frac{(\nabla \cdot v_h, q_h)}{\|v_h\|_{1,h} \|q_h\|_0} \geq \beta, \tag{5.32}$$

with a constant $\beta > 0$ independent of h.

Proof. First, we make use of the fact that $q_h \in Q_h \subset Q$ to deduce from Theorem 3.1.8 the there is a function $w \in V$ with $\nabla \cdot v = q_h$ and $\|v\|_1 \leq \|q_h\|_0$. To this function, we apply the Fortin operator to define $v_h = I_h v$. By the preceding lemma, we have

$$\|v_h\|_{1,h} \leq c\|v\|_1 \leq \|q_h\|_0,$$

which proves the inf-sup condition. □

5.2.6 Theorem: Assume that $(u_h, p_h) \in V_h \times Q_h$ is the solution to the divergence conforming DG method in Definition 5.2.2 and that the continuous Stokes problem is well-posed as in Theorem 3.1.8. Then, for the Raviart-Thomas pairs $RT_k/\mathbb{P}_k$ and $RT_{[k]}/\mathbb{Q}_k$ with $k \geq 1$ and u sufficiently smooth there holds

$$\|u - u_h\|_{1,h} \leq h^k |u|_{k+1}, \tag{5.33}$$

$$\|p - p_h\|_0 \leq h^k \big(|u|_{k+1} + |p|_k\big). \tag{5.34}$$

Furthermore,

$$\nabla \cdot u_h = 0. \tag{5.35}$$

Proof. The proof follows the lines of the abstract theory of Theorem 2.4.4 and Theorem 2.4.7. But since the setting with $V_h \not\subset V$ exceeds the assumptions of the abstract theory, we adapt the proofs instead of using the results.

Due to consistency of the method, we have

$$a_h(u - u_h, v_h) + (\nabla \cdot v_h, p - p_h) + (\nabla \cdot u - \nabla \cdot u_h, q_h) = 0.$$

Testing with $v_h = 0$ and using $\nabla \cdot V_h = Q_h$ immediately yields $\nabla \cdot u_h = \nabla \cdot u = 0$, or

$$\ker(B_h) \subset \ker(B).$$

In order to use the ellipticity of $a_h(.,.)$, we insert arbitrary functions $w_h \in \ker(B_h)$ and $r_h \in Q_h$. Choosing $q_h = 0$ yields the error equation

$$a_h(u_h - w_h, v_h) + (\nabla\cdot v_h, p_h - r_h) = a_h(u - w_h, v_h) + (\nabla\cdot v_h, p - r_h). \quad (5.36)$$

Testing with $v_h = u_h - w_h$ and employing $\nabla\cdot v_h = 0$, we obtain

$$\gamma\|u_h - w_h\|_{1,h}^2 \le a_h(u_h - w_h, u_h - w_h) = a_h(u - w_h, u_h - w_h).$$

Now, we use the canonical interpolation $w_h = I_h u$ to obtain

$$\gamma\|u_h - w_h\|_{1,h}^2 \le \frac{\gamma}{2}\|u_h - w_h\|_{1,h}^2 + \frac{c}{2\gamma}h^{2k}|u|_{k+1}^2.$$

Finally, we use the inf-sup condition to find a test function $v_h \in V_h$ such that $\nabla\cdot v_h = p_h - r_h$ and $\beta\|v_h\|_{1,h} \le \|p_h - r_h\|$. Then, the error equation (5.36) yields

$$\begin{aligned}\|p_h - r_h\| &= \frac{(\nabla\cdot v_h, p_h - r_h)}{\|p_h - r_h\|} \\ &= \frac{a_h(u - u_h, v_h) + (\nabla\cdot v_h, p - r_h)}{\|p_h - r_h\|} \le \tfrac{\|a_h\|}{\beta}\|u - u_h\|_{1,h} + \|p - r_h\|_0.\end{aligned}$$

Using the previously proven error estimate for u_h and the L^2-projection $r_h = \Pi_h p$ yields the result. □

5.3 Error estimates by duality

5.3.1. So far, we have only considered estimates in the so called energy norm, that is, a norm such that $a_h(.,.)$ is bounded and elliptic[1].

In the context of elliptic equations, we have seen the duality argument of Aubin and Nitsche, which allows us to obtain optimal estimates in weaker norms, for instance in L^2.

A particular difficulty here is the fact, that we have to test the dual solution with the error *and* exploit some kind of Galerkin orthogonality. Thus, we cannot use consistency as before and will introduce residual operators later. The analysis here is a simplified version of the corresponding results in [GKR14].

5.3.2 Definition: The dual problem to the Stokes problem in weak for consists of finding $(u^*, p^*) \in V_h \times Q_h$ such that for all $v \in V$ and $q \in Q$ there holds

$$(\nabla v, \nabla u^*) + (\nabla\cdot u^*, q) + (\nabla\cdot v, p^*) = (\psi, v). \quad (5.37)$$

[1] We use the term energy norm loosely here. Strictly speaking, the energy norm would be $\|v\|_A = \sqrt{a_h(v,v)}$.

5.3.3 Assumption: The dual Stokes problem admits the elliptic regularity estimate

$$\|u^*\|_2 \le c\|f\|_0. \tag{5.38}$$

Remark 5.3.4. Like for scalar elliptic equations, the elliptic regularity assumption holds for domains with smooth boundary or with piecewise smooth boundary where every corner is convex.

5.3.5 Definition: For the solutions $(u,p) \in V \times Q$ and $(u^*,p^*) \in V \times Q$ of the primal and dual Stokes problem, respectively, we define the residual operators

$$\operatorname{Res}(u,p;v) = a_h(u,v) + (\nabla\cdot v, p) - (f,v), \tag{5.39}$$
$$\operatorname{Res}^*(v;u^*,p^*) = a_h(v,u^*) + (\nabla\cdot v, p^*) - (\psi,v), \tag{5.40}$$

for $v \in V + V_h$.

5.3.6 Lemma: Let $(u,p) \in V \times Q$ be the solution to the Stokes problem with right hand side $f \in L^2(\Omega;\mathbb{R}^d)$. Assume $u \in H^s(\Omega;\mathbb{R}^d)$ and $p \in H^{s-1}(\Omega)$ with $s > 3/2$. Then, we have for $v \in V + V_h$:

$$(f,v) = (\nabla u, \nabla v)_{\mathbb{T}_h} - \langle \nabla u, \{\!\{v \otimes n\}\!\}\rangle_{\mathbb{F}_h^i} - \langle \partial_n u, v\rangle_{\mathbb{F}_h^\partial} + (\nabla\cdot v, p). \tag{5.41}$$

Proof. We set out from the strong form of the Stokes equations and integrate by parts.

$$\begin{aligned}(f,v) &= (-\Delta u + \nabla p, v)\\ &= (\nabla u, \nabla v)_{\mathbb{T}_h} - \sum_{T\in\mathbb{T}_h} \langle \partial_n u, v\rangle_{\partial T} - (\nabla\cdot v, p).\end{aligned}$$

Under the regularity assumptions of the lemma, all of these integrals make sense at least as duality pairings. In particular, $\partial_n u \in L^2(\partial T)$, and thus we can split ∂T into individual faces. Therefore,

$$\sum_{T\in\mathbb{T}_h} \langle \partial_n u, v\rangle_{\partial T} = \langle \nabla u, \{\!\{v \otimes n\}\!\}\rangle_{\mathbb{F}_h^i} + \langle \partial_n u, v\rangle_{\mathbb{F}_h^\partial}.$$

The proof concludes by collecting the results. □

5.3.7 Corollary: The residual operators can be expressed as

$$
\begin{aligned}
\operatorname{Res}(u,p;v) &= a_h(u,v) - (\nabla u, \nabla v)_{\mathbb{T}_h} \\
&\qquad + \langle \nabla u, \{\!\{ v \otimes n \}\!\} \rangle_{\mathbb{F}_h^i} + \langle \partial_n u, v \rangle_{\mathbb{F}_h^\partial}. \\
\operatorname{Res}^*(u^*,p^*;v) &= a_h(v,u^*) - (\nabla u, \nabla v)_{\mathbb{T}_h} \\
&\qquad + \langle \nabla u, \{\!\{ v \otimes n \}\!\} \rangle_{\mathbb{F}_h^i} + \langle \partial_n u, v \rangle_{\mathbb{F}_h^\partial}.
\end{aligned}
\tag{5.42}
$$

In particular, the residual operators do not depend on the pressure solutions.

5.3.8 Theorem: Let the assumptions of Theorem 5.2.6 and Assumption 5.3.3 hold. Then,

$$
\|u - u_h\|_0 \le c h^{k+1} |u|_{k+1}. \tag{5.43}
$$

5.3.9 Problem: Adapt the proof of Theorem 5.1.25 to prove Theorem 5.3.8.

Chapter 6

Maxwell's equations and the de Rham complex

6.1 Maxwell's equations

6.1.1 Notation: With $\nabla\times u$ we describe the curl of a vector field u, which in three dimensions is defined as

$$\nabla\times u = \begin{pmatrix} \partial_2 u_3 - \partial_3 u_2 \\ \partial_3 u_1 - \partial_1 u_3 \\ \partial_1 u_2 - \partial_2 u_1 \end{pmatrix}. \tag{6.1}$$

In two dimension, we distinguish between the vector curl of a scalar function and the scalar curl of a vector function

$$\nabla\times u = \partial_1 u_2 - \partial_2 u_1, \qquad \nabla\times\varphi = \begin{pmatrix} \partial_2\varphi \\ -\partial_1\varphi \end{pmatrix}. \tag{6.2}$$

Remark 6.1.2. The scalar curl of a two-dimensional vector field is equal to the third component of the extension of this vector field by zero into $\mathbb{R}^3$, in formulas,

$$\nabla\times\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \nabla\times\begin{pmatrix} u_1 \\ u_2 \\ 0 \end{pmatrix}_3.$$

Similarly, the vector curl of a scalar function φ in two dimensions consistes of the first two components of the curl of a three dimensional function in the last component of t he vector,

$$\nabla\times\varphi = \nabla\times \begin{pmatrix} 0 \\ 0 \\ \varphi \end{pmatrix}_{1,2}.$$

Remark 6.1.3. A popular error in the literature consists of the following argument: since $\nabla\cdot E = 0$, there also holds $\nabla\nabla\cdot E = 0$. Therefore, we can use the formula

$$\Delta u = \nabla\nabla\cdot u - \nabla\times\nabla\times u,$$

and avoid the div-curl-problem alltogether. Unfortunately, this is only true, if solutions of (6.7) are in $H^1(\Omega;\mathbb{R}^d)$, which is not true, depending on the boundary conditions.

6.1.4 Lemma: For vector fields $u, v \in C^1(\overline{\Omega})$, there holds

$$\int_\Omega \nabla\times u\cdot v \,\mathrm{d}x = \int_\Omega u\cdot\nabla\times v\,\mathrm{d}x + \int_{\partial\Omega} (n\times u)\cdot v\,\mathrm{d}s. \tag{6.3}$$

6.1.5. Electromagnetic fields are governed by four laws of nature put together by James Clerk Maxwell to a single system. The laws are

1. Gauss' law for the electric field: the electric flux through a closed surface equals $1/\varepsilon$ times the electric charge enclosed by the surface:

$$\int_{\partial V} E\cdot n\,\mathrm{d}s = \int_V \frac{\varrho}{\varepsilon}\,\mathrm{d}x.$$

2. There are no magnetic monopoles, therefore the magnetic flux through any closed surface vanishes:

$$\int_{\partial V} B\cdot n\,\mathrm{d}s = 0.$$

3. Faraday's law of induction: the voltage induced in a closed loop is proportional to the rate of change of the magnetic field through the surface encloded by the loop:

$$\int_{\partial A} E\cdot\,\mathrm{d}s = -\frac{d}{dt}\int_A B\cdot n\,\mathrm{d}s.$$

4. Ampère's law: the magnetic field induced in a closed loop is proportional to the electric current plus the change of electric field through that loop:

$$\int_{\partial A} B\cdot\,\mathrm{d}s = \mu\int_A J\cdot n\,\mathrm{d}s + \mu\varepsilon\frac{d}{dt}\int_A E\cdot n\,\mathrm{d}s.$$

Using the Gauss theorem for the first two and the Stokes theorem for the remaining two laws, we obtain the **Maxwell equations** of electromagnetics

$$\nabla\cdot E = \frac{\varrho}{\varepsilon} \qquad \nabla\times E = -\partial_t B, \tag{6.4}$$

$$\nabla\cdot B = 0 \qquad \nabla\times B = \mu J + \mu\varepsilon E. \tag{6.5}$$

They are an hyperbolic system of equations and typically have wave solutions. Many simplifications have been developed to suit particular purposes.

6.1.6. An important simplification of the Maxwell equations is obtained by assuming an isolating material, that is, the electric current J vanishes. Additionally, we may assume that there are no electric charges, such that $\nabla\cdot E = 0$. Then, taking the curl of the equation for $\nabla\times E$ and inserting the formula for $\nabla\times B$, we obtain

$$\mu\varepsilon\partial_t^2 E + \nabla\times\nabla\times E = 0 \qquad \nabla\cdot E = 0. \tag{6.6}$$

We can even go further and study the stationary limit

$$\nabla\times\nabla\times E = 0 \qquad \nabla\cdot E = 0. \tag{6.7}$$

This is the equation we are concerned with most, since its solution theory also provides insight into the other forms.

6.1.7 Definition: The Maxwell equation (6.7) is complemented with the following boundary conditions:

- Perfectly conducting:

$$n \times u = 0. \tag{6.8}$$

- Natural:

$$n \times \nabla\times u = 0. \tag{6.9}$$

- Impedance:

$$n \times \nabla\times u - \alpha(n \times u) \times n = 0. \tag{6.10}$$

6.1.8 Definition: For $u \in C^1(\overline{\Omega})$, we define the trace operators

$$\begin{aligned} \gamma_\tau &= n \times u_{|\partial\Omega}, \\ \gamma_T &= n \times u_{|\partial\Omega} \times n. \end{aligned} \tag{6.11}$$

The second of these is the tangential component of u on the boundary. Furthermore, we introduce the space H_0^{curl} as the completion of the space of differentiable functions with compact support under the norm of H^{curl}

$$H_0^{\text{curl}} = \overline{C_{00}^\infty(\Omega; \mathbb{R}^d)}^{H^{\text{curl}}}. \tag{6.12}$$

6.1.9 Theorem: The trace operator γ_τ can be extended to a continuous, surjective operator

$$\gamma_\tau \colon H^{\text{curl}}(\Omega) \to Y(\partial\Omega),$$

where

$$\begin{aligned} Y(\partial\Omega) &= \left\{ u \in H_\tau^{-1/2}(\partial\Omega) \middle| \ \nu \cdot (\nabla\times u) \in H^{-1/2}(\partial\Omega) \right\}, \\ H_\tau^{-1/2}(\partial\Omega) &= \left\{ u \in H^{-1/2}(\partial\Omega; \mathbb{R}^d) \middle| u \cdot n = 0 \text{ a.e.} \right\}. \end{aligned} \tag{6.13}$$

Furthermore, the trace operator γ_T can be extended to a continuous operator

$$\gamma_T \colon H^{\text{curl}}(\Omega) \to Y(\partial\Omega)^*.$$

6.1.10. The trace theorem indicates, that $H_0^{\text{curl}}(\Omega)$ is the correct space to solve the problem with perfectly conducting boundary condition on the whole boundary. It remains now to deal with the divergence constraint. First, we note, that the divergence operator is not well-defined on H^{curl}, and that the subspace of H^{curl} with divergence in L^2 is H^1, which must be avoided. Therefore, we have to resort to a dual formulation of this constraint, which leads to the following weak form of the perfectly conducting Maxwell problem.

6.1.11 Definition: The Maxwell problem for perfectly conducting boundary conditions in weak form reads: find $(u,p) \in V \times Q$, where $V = H_0^{\text{curl}}(\Omega)$ and $Q = H_0^1(\Omega)$ such that there holds

$$\begin{aligned} (\nabla\times u, \nabla\times v) + (v, \nabla p) \quad &= (f, v) \quad && \forall v \in V \\ (u, \nabla q) \quad &= 0 && \forall q \in Q. \end{aligned} \tag{6.14}$$

Remark 6.1.12. At this point, our task is laid out. We have to prove well-posedness of the Maxwell problem in mixed form, then find suitable finite element spaces and commuting interpolation operators. It turns out that this can be done in a more general framework, called the de Rham complex.

6.2 The de Rham complex

6.2.1. We can embed finite element methods for the Darcy problem, also for the Maxwell problem, into a common framework based on the de Rham complex. If we wanted to do this in its full mathematical beauty, we would have to spend some time introducing the concept and notation of differential forms. As an alternative, we can use the concrete vector spaces $H^{\mathrm{div}}(\Omega)$ and $H^{\mathrm{curl}}(\Omega)$. The drawback is, that we have to prove several particular cases, where the abstract theory only knows one common case. Nevertheless, it is worthwhile to begin this way, such that the reader has an easier task reading the full theory in [AFW06; AFW10]. As a byproduct, we will prove in generality some of the properties of polynomial spaces in Chapter 4.

6.2.2. We now know three differential operators, ∇, $\nabla\times$, and $\nabla\cdot$ with the interesting property

$$\nabla\times\nabla\varphi = 0 \qquad \nabla\cdot\nabla\times E = 0. \tag{6.15}$$

As a consequence, for $\varphi \in H^1(\Omega)$ we not only have $\nabla\varphi \in L^2(\Omega;\mathbb{R}^3)$, we also have $\nabla\times\nabla\varphi = 0 \in L^2(\Omega;\mathbb{R}^3)$. This gives rise to the sequence

$$\mathbb{R} \xrightarrow{\subset} H^1(\Omega) \xrightarrow{\nabla} H^{\mathrm{curl}}(\Omega) \xrightarrow{\nabla\times} H^{\mathrm{div}}(\Omega) \xrightarrow{\nabla\cdot} L^2(\Omega) \longrightarrow 0, \tag{6.16}$$

such that the range of an operator is always in the kernel of the operator to its right.

6.2.3 Notation: The notation of exterior calculus of differential forms allows us to write this sequence elegantly as

$$\begin{array}{ccccccccccc}
\mathbb{R} & \xrightarrow{d} & H\Lambda^0(\Omega) & \xrightarrow{d} & H\Lambda^1(\Omega) & \xrightarrow{d} & H\Lambda^2(\Omega) & \xrightarrow{d} & H\Lambda^3(\Omega) & \longrightarrow & 0 \\
 & & \cong\downarrow & & \cong\downarrow & & \cong\downarrow & & \cong\downarrow & & \\
\mathbb{R} & \xrightarrow{\subset} & H^1(\Omega) & \xrightarrow{\nabla} & H^{\mathrm{curl}}(\Omega) & \xrightarrow{\nabla\times} & H^{\mathrm{div}}(\Omega) & \xrightarrow{\nabla\cdot} & L^2(\Omega) & \longrightarrow & 0,
\end{array} \tag{6.17}$$

such that $d = d_k\colon H\Lambda^k(\Omega) \to H\Lambda^{k+1}(\Omega)$ and

$$d^2 = d\circ d = d_{k+1}\circ d_k = 0. \tag{6.18}$$

Remark 6.2.4. The spaces $H\Lambda^k(\Omega)$ are Hilbert spaces with values in the spaces of alternating k-forms on $\mathbb{R}^d$. From linear algebra, we know that all alternating k-forms are zero if k exceeds the dimension of the vector space. Therefore, the sequence above is only valid in three dimensions, and it must be shorter by one member in two dimensions. Changing our view back to differential operators, we realize that there are two relevant sequences in two dimensions. In the following diagram, the sequence on top can be used to formulate Maxwell problems in H^{curl} in two dimensions, while the sequence on the bottom relates to the mixed form of the Laplacian.

We introduce the sequences in two dimensions and afterwards will focus our arguments on the more general case of three dimensions again. Specialization to two dimensions are straight forward.

6.2.5 Notation: In two dimensions, we consider the de Rham sequences

$$\begin{array}{ccccccccc}
\mathbb{R} & \xrightarrow{\subset} & H^1(\Omega) & \xrightarrow{\nabla} & H^{\text{curl}}(\Omega) & \xrightarrow{\nabla\times} & L^2(\Omega) & \longrightarrow & 0 \\
 & & \cong\big\uparrow & & \cong\big\uparrow & & \cong\big\uparrow & & \\
\mathbb{R} & \xrightarrow{d} & H\Lambda^0(\Omega) & \xrightarrow{d} & H\Lambda^1(\Omega) & \xrightarrow{d} & H\Lambda^2(\Omega) & \longrightarrow & 0 \\
 & & \cong\big\downarrow & & \cong\big\downarrow & & \cong\big\downarrow & & \\
\mathbb{R} & \xrightarrow{\subset} & H^1(\Omega) & \xrightarrow{\nabla\times} & H^{\text{div}}(\Omega) & \xrightarrow{\nabla\cdot} & L^2(\Omega) & \longrightarrow & 0,
\end{array} \tag{6.19}$$

6.2.6 Notation: The spaces $H\Lambda^k(\Omega)$ are Hilbert spaces with the inner product

$$\langle u, v\rangle_{H\Lambda^k} = \langle u, v\rangle_{L^2} + \langle du, dv\rangle_{L^2}. \tag{6.20}$$

The value of this notation lies in the following theorem by de Rham, which describes the relation between the elements of the sequence. It is cited here without proof.

6.2.7 Theorem: Assume the domain Ω is Lipschitz. If Ω is simply connected, the sequences in equations (6.17) and (6.19) are exact, that is, there holds

$$\ker(d_{k+1}) = \operatorname{im}(d_k). \tag{6.21}$$

If it is not simply connected, the codimension of $\operatorname{im}(d_k)$ in $\ker(d_{k+1})$ is finite. In particular, in both cases, $\operatorname{im}(d_k)$ is closed in $H\Lambda^{k+1}(\Omega)$.

So far, we have not considered boundary conditions. The next lemma, which is again stated without proof, indicates that the properties of the de Rham complex are inherited, if the appropriate boundary conditions are applied to each space, namely, function values in H^1, tangential traces in H^{curl}, and normal traces in H^{div}. The last restriction from L^2 to L^2_0 is not a boundary condition, but it is the compatibility condition implied by the Gauss theorem on H^{div}.

6.2.8 Lemma: The bounded Hilbert cochain complex

$$\begin{array}{ccccccccccc}
0 & \xrightarrow{d} & H\Lambda^0_0(\Omega) & \xrightarrow{d} & H\Lambda^1_0(\Omega) & \xrightarrow{d} & H\Lambda^2_0(\Omega) & \xrightarrow{d} & H\Lambda^3_0(\Omega) & \longrightarrow & 0 \\
 & & \cong\big\downarrow & & \cong\big\downarrow & & \cong\big\downarrow & & \cong\big\downarrow & & \\
0 & \longrightarrow & H^1_0(\Omega) & \xrightarrow{\nabla} & H^{\mathrm{curl}}_0(\Omega) & \xrightarrow{\nabla\times} & H^{\mathrm{div}}_0(\Omega) & \xrightarrow{\nabla\cdot} & L^2_0(\Omega) & \longrightarrow & 0,
\end{array} \tag{6.22}$$

has the same properties as stated for the Hilbert complex without boundary conditions.

Remark 6.2.9. The complex does not start with $\mathbb{R}$ on the left, but with zero, since the constant functions are not members of $H^1_0(\Omega)$.

On the other hand, we could have replaced the right end of the complex by

$$L^2(\Omega) \xrightarrow{\frac{1}{|\Omega|}\int} \mathbb{R},$$

where the arror is the mean value operator.

6.2.10 Theorem: The Maxwell problem in Definition 6.1.11 is well posed.

Proof. We have to show the inf-sup condition and the ellipticity of the curl-curl bilinear form. Let us introduce

$$a(u,v) = (\nabla\times u, \nabla\times v), \qquad b(v,q) = (v, \nabla q).$$

From the fact that the de Rham complex starts with zero, we obtain that the kernel of the gradient is zero. Thus, for any $q \in H^1_0(\Omega)\backslash\{0\}$, we have $v = \nabla q \neq 0$ and $\|v\|_{H^{\mathrm{curl}}} = \|v\|_{L^2} \le \|q\|_{H^1}$. Thus, the inf-sup condition holds.

We show now that $a(.,.)$ is elliptic on $\ker(B)$. From the definition of $b(.,.)$, we deduce that $\ker(B) \perp \nabla H^1_0(\Omega) = \ker(A)$. Thus, A is an isomorphism between $\ker(B)$ and its dual, and consequently elliptic. □

6.2.11 Problem: Prove well-posedness for the Darcy problem using the de Rham complex for proving Lemma 4.1.20 and Lemma 4.1.22.

6.3 Polynomial complexes for simplicial meshes

6.3.1. We have already seen that adding $x\mathbb{P}_k$ to the space $\mathbb{P}_k^d$, we obtain a surjective divergence operator from the Raviart-Thomas element to the pressure space $\mathbb{P}_k$. In this section, we see that there is a general principle behind this concept and it can be extended to the curl and gradient operators.

6.3.2 Notation: The homogeneous polynomial spaces $\breve{\mathbb{P}}_k$ form the cochain complex

$$\begin{array}{ccccccccccc}
\mathbb{R} & \xrightarrow{d} & \breve{\mathbb{P}}_r\Lambda^0 & \xrightarrow{d} & \breve{\mathbb{P}}_{r-1}\Lambda^1 & \xrightarrow{d} & \breve{\mathbb{P}}_{r-2}\Lambda^2 & \xrightarrow{d} & \breve{\mathbb{P}}_{r-3}\Lambda^3 & \xrightarrow{d} & 0 \\
 & & \cong\downarrow & & \cong\downarrow & & \cong\downarrow & & \cong\downarrow & & \\
\mathbb{R} & \xrightarrow{\subset} & \breve{\mathbb{P}}_r & \xrightarrow{\nabla} & \breve{\mathbb{P}}^3_{r-1} & \xrightarrow{\nabla\times} & \breve{\mathbb{P}}^3_{r-2} & \xrightarrow{\nabla\cdot} & \breve{\mathbb{P}}_{r-3} & \longrightarrow & 0,
\end{array} \tag{6.23}$$

and $d_{k+1} \circ d_k = 0$.

Remark 6.3.3. Since the polynomial space $\mathbb{P}_r$ is the direct sum

$$\mathbb{P}_r = \bigoplus_{s=0}^{r} \breve{\mathbb{P}}_s,$$

the homogeneous polynomial complex above can be extended to a general polynomial complex in a straightforward way.

6.3.1 The Koszul complex

6.3.4 Definition: The homogeneous **Koszul complex** is a polynomial complex of the form

$$0 \longleftarrow \breve{\mathbb{P}}_r\Lambda^0 \xleftarrow{\kappa_1} \breve{\mathbb{P}}_{r-1}\Lambda^1 \xleftarrow{\kappa_2} \breve{\mathbb{P}}_{r-2}\Lambda^2 \xleftarrow{\kappa_3} \breve{\mathbb{P}}_{r-3}\Lambda^3 \longleftarrow 0. \tag{6.24}$$

The **Koszul differential** is defined such that

$$\begin{aligned}
\kappa_1\omega &= x \cdot \omega & \omega &\in \mathbb{P}_s\Lambda^1, \\
\kappa_2\omega &= -x \times \omega & \omega &\in \mathbb{P}_s\Lambda^2, \\
\kappa_3\omega &= x\omega & \omega &\in \mathbb{P}_s\Lambda^3,
\end{aligned} \tag{6.25}$$

and there holds

$$\kappa \circ \kappa = \kappa_{k+1} \circ \kappa_k = 0. \tag{6.26}$$

Note that the "Koszul differential" increases the polynomial order and lowers the order of the form, thus acts in the opposite way of the usual differential d.

6.3.5 Lemma: For $\omega \in \breve{\mathbb{P}}_r\Lambda^k$ there holds

$$\big(d\kappa + \kappa d\big)\omega = (r+k)\omega. \tag{6.27}$$

Proof. Since we are not using differential form technology, we prove this for each k directly. For $k=0$, we have $\kappa\omega = 0$, thus we have to show

$$\kappa d\omega = r\omega.$$

Due to linearity of κ and d, it suffices to prove the result for $\omega = p = x_1^a x_2^b x_3^c$. We note that $dp/d_{x_1} = a/x_1 p$ and $d(x_1 p)/d_{x_1} = (a+1)p$ and analogue for the other coordinates.

$$\kappa_1 d_0 \omega = x \cdot \nabla p = x \cdot \begin{pmatrix} a/x_1 \\ b/x_2 \\ c/x_3 \end{pmatrix} p = (a+b+c)p.$$

The second easy case is $k=3$ such that $d\omega = 0$. Let again $\omega = p$ to obtain

$$d_2\kappa_3\omega = \nabla\cdot(xp) = \nabla\cdot \begin{pmatrix} x_1 p \\ x_2 p \\ x_3 p \end{pmatrix} = (a+1+b+1+c+1)p = (r+3)\omega.$$

For the two vector valued cases, we note that it suffices to prove the result for $\omega = (p,0,0)^T$ and to note that the results for nonzero second and third component follow suite. Thus, for $k=1$

$$\begin{aligned}
\nabla(x\cdot\omega) - x\times\nabla\times\omega &= \nabla(x_1 p) - x \times \begin{pmatrix} 0 \\ c/x_3 \\ -b/x_2 \end{pmatrix} p \\
&= \begin{pmatrix} a+1 \\ bx_1/x_2 \\ cx_1/x_3 \end{pmatrix} p + \begin{pmatrix} b+c \\ -bx_1/x_2 \\ cx_1/x_3 \end{pmatrix} p = \begin{pmatrix} a+b+c+1 \\ 0 \\ 0 \end{pmatrix} p = (r+1)\omega.
\end{aligned}$$

Finally, for $k=2$

$$\begin{aligned}
\nabla\times(-x\times\omega) + x\nabla\cdot\omega &= \nabla\times \begin{pmatrix} 0 \\ -x_3 \\ x_2 \end{pmatrix} p + \begin{pmatrix} x_1 a/x_1 \\ x_2 a/x_1 \\ x_3 a/x_1 \end{pmatrix} p \\
&= \begin{pmatrix} b+1+c+1 \\ -ax_2/x_1 \\ -ax_3/x_1 \end{pmatrix} p + \begin{pmatrix} a \\ ax_2/x_1 \\ ax_3/x_1 \end{pmatrix} p = \begin{pmatrix} a+b+c+2 \\ 0 \\ 0 \end{pmatrix} p = (r+2)\omega.
\end{aligned}$$

□

6.3.6 Lemma: The restriction of operator d to $\operatorname{im}(\kappa)$ is injective and vice versa, or equivalently for any polynomial form $\omega \in \breve{\mathbb{P}}_r\Lambda^k$ there holds

$$\begin{aligned} d\kappa\omega = 0 \quad &\Longrightarrow \quad \kappa\omega = 0, \\ \kappa d\omega = 0 \quad &\Longrightarrow \quad d\omega = 0. \end{aligned} \tag{6.28}$$

Proof. If $r = k = 0$, then $\kappa\omega = d\omega = 0$, such that the lemma holds trivially. For $r + k \neq 0$, we apply κ to equation (6.27) to obtain

$$\kappa\omega = \frac{1}{r+k}\big(\kappa d\kappa\omega + \kappa^2 d\omega\big) = \frac{1}{r+k}\kappa d\kappa\omega.$$

Thus, we have proven $d\kappa\omega = 0$ implies $\kappa\omega = 0$. The second implication is proven by applying d to (6.27). □

6.3.7 Theorem: The polynomial de Rham complex and the Koszul complex are exact for $r > 3$. Furthermore for $r + k > 0$, there holds

$$\breve{\mathbb{P}}_r\Lambda^k = \kappa\breve{\mathbb{P}}_{r-1}\Lambda^{k+1} \oplus d\breve{\mathbb{P}}_{r+1}\Lambda^{k-1}. \tag{6.29}$$

Proof. We already know $\operatorname{im}(\kappa_{k-1}) \subset \ker(\kappa_k)$. Thus, it remains to show the opposite inclusion. Let therefore $\omega \in \breve{\mathbb{P}}_r\Lambda^k$ such that $\kappa\omega = 0$. Then,

$$\omega = \frac{1}{r+k}(d\kappa\omega + \kappa d\omega) = \frac{1}{r+k}\kappa d\omega =: \kappa\eta$$

with $\eta \in \breve{\mathbb{P}}_{r-1}\Lambda^{k+1}$. Thus, $\omega \in \operatorname{im}(\kappa_{k-1})$. Again, the proof for the de Rham complex is obtained by replacing κ by d.

In order to see that $\breve{\mathbb{P}}_r\Lambda^k$ is the sum of the two spaces, we let for arbitrary $\omega \in \breve{\mathbb{P}}_r\Lambda^k$

$$\eta = \frac{1}{r+k}d\omega \in \breve{\mathbb{P}}_{r-1}\Lambda^{k+1}, \qquad \mu = \frac{1}{r+k}\kappa\omega \in \breve{\mathbb{P}}_{k+1}\Lambda^{k-1}.$$

By equation (6.27), we have $\omega = \kappa\eta + d\mu$. It remains to show that the intersection of the spaces is zero. Therefore, let ω be chosen from the intersection. Then, $\omega = \kappa\eta = d\mu$ and

$$(r+k)\omega = d\kappa\omega + \kappa d\omega = d\kappa^2\eta + \kappa d^2\mu = 0.$$

□

6.3.8 Corollary: Theorem 6.3.7 holds as well for the polynomial complexes

$$0 \longrightarrow \mathbb{P}_r\Lambda^0 \xrightarrow{d} \mathbb{P}_{r-1}\Lambda^1 \xrightarrow{d} \mathbb{P}_{r-2}\Lambda^2 \xrightarrow{d} \mathbb{P}_{r-3}\Lambda^3 \longrightarrow 0, \tag{6.30}$$

and

$$\mathbb{R} \longleftarrow \mathbb{P}_r\Lambda^0 \xleftarrow{\kappa} \mathbb{P}_{r-1}\Lambda^1 \xleftarrow{\kappa} \mathbb{P}_{r-2}\Lambda^2 \xleftarrow{\kappa} \mathbb{P}_{r-3}\Lambda^3 \longleftarrow 0. \tag{6.31}$$

Proof. This is due to the fact that the polynomial spaces $\mathbb{P}_r$ are the direct sums of homogeneous polynomial space $\breve{\mathbb{P}}_s$. □

6.3.9 Definition: The polynomial space of k-forms $\mathbb{P}_r^+\Lambda^k$ is defined as

$$\mathbb{P}_r^+\Lambda^k = \mathbb{P}_r\Lambda^k \oplus \kappa\mathbb{P}_r\Lambda^{k+1}. \tag{6.32}$$

It is also referred to as $\mathbb{P}_{r+1}^-\Lambda^k$. Furthermore,

$$\mathbb{P}_r^+\Lambda^0 = \mathbb{P}_{r+1}\Lambda^0, \qquad \mathbb{P}_r^+\Lambda^d = \mathbb{P}_r\Lambda^d.$$

Remark 6.3.10. We have used the construction principle

$$\mathbb{P}_r\Lambda^k = \mathbb{P}_{r-1}\Lambda^k \oplus \breve{\mathbb{P}}_r\Lambda^k.$$

Using its decomposition, we obtain

$$\mathbb{P}_r\Lambda^k = \bigoplus_{s=1}^{r-1} \kappa\breve{\mathbb{P}}_{s-1}\Lambda^{k+1} \bigoplus_{s=1}^{r-1} d\breve{\mathbb{P}}_{s+1}\Lambda^{k-1} \oplus \kappa\breve{\mathbb{P}}_{r-1}\Lambda^{k+1} \oplus d\breve{\mathbb{P}}_{r+1}\Lambda^{k-1}.$$

If we leave out the last factor, we get the new space $\mathbb{P}_{r-1}^+\Lambda^k$.

6.3.11 Lemma: If $\omega \in \mathbb{P}_r^+\Lambda^k$ and $d\omega = 0$, then $\omega \in \mathbb{P}_r\Lambda^k$.

Proof. Let $\omega = \omega_1 + \kappa\eta$ with $\omega_1 \in \mathbb{P}_r\Lambda^k$ and $\eta \in \breve{\mathbb{P}}_r\Lambda^{k+1}$. Then, $d\omega_1 \in \mathbb{P}_{r-1}\Lambda^{k+1}$ and $d\kappa\eta \in \breve{\mathbb{P}}_r\Lambda^{k+1}$. Therefore, $d\omega_1 = d\kappa\eta = 0$. By Lemma 6.3.6, $\kappa\eta = 0$, such that $\omega = \omega_1$. □

6.3.12 Lemma: For $r \geq 1$ and $0 \leq k < d$ there holds

$$d\mathbb{P}_r^+\Lambda^k \subset d\mathbb{P}_{r+1}\Lambda^k \subset \mathbb{P}_r\Lambda^{k+1} \subset \mathbb{P}_r^+\Lambda^{k+1}. \tag{6.33}$$

The following four mappings d have the same kernel:

$$\begin{aligned} &d\colon \mathbb{P}_r\Lambda_k \to \mathbb{P}_{r-1}\Lambda^{k+1} \quad && d\colon \mathbb{P}_r^+\Lambda_k \to \mathbb{P}_r\Lambda^{k+1} \\ &d\colon \mathbb{P}_r\Lambda_k \to \mathbb{P}_{r-1}^+\Lambda^{k+1} \quad && d\colon \mathbb{P}_r^+\Lambda_k \to \mathbb{P}_r^+\Lambda^{k+1} \end{aligned} \tag{6.34}$$

The following four mappings d have the same range:

$$\begin{aligned} &d\colon \mathbb{P}_r\Lambda_k \to \mathbb{P}_{r-1}\Lambda^{k+1} \quad && d\colon \mathbb{P}_{r-1}^+\Lambda_k \to \mathbb{P}_{r-1}\Lambda^{k+1} \\ &d\colon \mathbb{P}_r\Lambda_k \to \mathbb{P}_{r-1}^+\Lambda^{k+1} \quad && d\colon \mathbb{P}_{r-1}^+\Lambda_k \to \mathbb{P}_{r-1}^+\Lambda^{k+1} \end{aligned} \tag{6.35}$$

Proof. The first statement follows from the inclusions of $\mathbb{P}_r$ and $\mathbb{P}_r^+$. The horizontal equality of the second statement follows from Lemma 6.3.11. The vertical identities from the decomposition (6.29). For the last set of identities, we observe that by construction

$$\mathbb{P}_r\Lambda^k = \mathbb{P}_{r-1}^+\Lambda^k \oplus d\mathbb{P}_{r+1}\Lambda^{k-1}.$$

Thus, $d\mathbb{P}_r\Lambda^k = d\mathbb{P}_{r-1}^+\Lambda^k \subset \mathbb{P}_{r-1}\Lambda^{k+1}$. □

6.3.13 Theorem: Let $r \geq 0$ and $1 \leq k \leq d$. Then,

$$\begin{aligned} \dim \kappa\breve{P}_r\Lambda^k(\mathbb{R}^d) &= \dim d\breve{P}_{r+1}\Lambda^{k-1}(\mathbb{R}^d) \\ &= \binom{d+r}{d-k}\binom{r+k-1}{k-1}. \end{aligned} \tag{6.36}$$

Proof. First, we prove the equality of the two dimensions by applying κ to equation (6.29), yielding

$$\kappa\breve{P}_r\Lambda^k(\mathbb{R}^d) = \kappa d\breve{P}_{r+1}\Lambda^{k-1}(\mathbb{R}^d).$$

By Lemma 6.3.6, the two spaces are isomorphic and the equality holds.

The dimension formula is proven first for $r = 0$ and $k \geq 1$. The Koszul operator is injective on $\mathbb{P}_0\Lambda^K(\mathbb{R}^d)$ since the first factor in equation (6.29) vanishes. It is also injective on $\breve{P}_r\Lambda^d(\mathbb{R}^d)$ for $r \geq 0$.

For all other combinations of r and k it is proven by induction over k. For $k = d$,

$$\dim \breve{\mathbb{P}}_r\Lambda^d(\mathbb{R}^d) = \dim \breve{\mathbb{P}}_r(\mathbb{R}^d) = \binom{d+r-1}{d-1}.$$

For $k < d$, we assume the formula proven for $k+1$. We have

$$\breve{P}_r\Lambda^k(\mathbb{R}^d) = \binom{d+r-1}{d-1}\binom{d}{k}.$$

Now, the dimension formula

$$\dim \operatorname{im}(\varphi) = \dim V - \dim ker\varphi,$$

yields

$$\dim \kappa\breve{\mathbb{P}}_r\Lambda^k(\mathbb{R}^d) = \dim \breve{\mathbb{P}}_r\Lambda^k(\mathbb{R}^d) - \dim \kappa\breve{\mathbb{P}}_{r-1}\Lambda^{k+1}(\mathbb{R}^d),$$

where we have used the exactness of the Koszul complex. Using the induction hypothesis yields by the binomial identity

$$\begin{aligned}
\dim \kappa\breve{\mathbb{P}}_r\Lambda^k(\mathbb{R}^d) &= \binom{d+r-1}{d-1}\binom{d}{k} - \binom{d+r-1}{d-k-1}\binom{r+k-1}{k} \\
&= \frac{(d+r-1)!d!}{(d-1)!r!k!(d-k)!} - \frac{(d+r-1)!(r+k-1)!}{(d-k-1)!(r+k)!k!(r-1)!} \\
&= \frac{(d+r-1)!d(r+k) - (d-k)(d+r-1)!r}{r!k!(d-k)!(r+k)} \\
&= \frac{(d+r-1)!\big(d(r+k)-(d-k)r\big)}{r!k!(d-k)!(r+k)} \\
&= \frac{(d+r)!}{r!(k-1)!(d-k)!(r+k)} \\
&= \binom{d+r}{d-k}\binom{r+k-1}{k-1}.
\end{aligned}$$

□

6.3.2 Degrees of freedom and bases for simplicial meshes

6.3.14. After having studied the properties of the de Rham complex and the Koszul complex of polynomial spaces, we continue like with standard finite elements and define a basis of shape functions and sets of degrees of freedom dual to this basis. Note that the following definition subsumes the definitions of conforming finite elements for H^1, H^{curl} and H^{div} in a single statement.

6.3.15 Definition: Given a space of polynomial forms $\mathbb{P}_r\Lambda^k = \mathbb{P}_r\Lambda^k(\mathbb{R}^d)$, we define the space of finite element polynomial forms on a mesh $\mathbb{T}_h$ covering the domain $\Omega \subset \mathbb{R}^d$ as

$$\mathbb{P}_r\Lambda^k(\mathbb{T}) = \left\{\omega \in H\Lambda^k \,\middle|\, \forall T \in \mathbb{T}\colon \omega_{|T} \in \mathbb{P}_r\Lambda^k\right\}. \tag{6.37}$$

6.3.16. The degrees of freedom have to be designed such that they guarantee the necessary continuity between cells. To this end, we have to study the traces of polynomial forms on the boundaries (called subsimplices below) of the simplex T. Then, we can start decomposing degrees of freedom and node values such that they can be allocated to these subsimplices.

Geometric structure of simplices

6.3.17 Definition: Let $x_0, \dots, x_k$ for $k \le d$ be a set of $k+1$ points in $\mathbb{R}^d$. Then, we call the set of convex combinations of these points the k-simplex f spanned by $\{x_0, \dots, x_k\}$.

6.3.18 Definition: Let T be the simplex in $\mathbb{R}^d$ spanned by the points $x_0, \dots, x_d$. Then, every ascending subset $\sigma \subset \{0, \dots, d\}$ of length $k+1$ defines a k-dimensional **subsimplex** of T denoted as f_σ.
The set of all subsimplices of T, including T itself is called $\Delta(T)$. The set of all k-dimensional subsimplices is $\Delta_k(T)$.

Example 6.3.19. A d-dimensional simplex T has $\binom{d+1}{k+1}$ subsimplices of dimension k.

The one-dimensional simplex $[x_0, x_1]$ has two subsimplices of dimension zero, namely the two points x_0 and x_1.

The triangle spanned by the points x_0, x_1, x_3 has three one-dimensional subsimplices (edges) and three zero-dimensional subsimplices (vertices).

The tetrahedron spanned by the points $x_0, \dots, x_3$ has

- 4 triangular faces,
- 6 edges,
- 4 vertices.

Geometric decomposition of $\mathbb{P}_r(T)$

Remark 6.3.20. For a simplex $T \subset \mathbb{R}^d$, the barycentric coordinates are the uniquely determined linear interpolating polynomials such that

$$\lambda_i(x_j) = \delta_{ij}.$$

Then,

$$T = \left\{ x = \sum_{i=0}^{d} \lambda_i \middle| \lambda_i \geq 0, \quad \sum_{i=0}^{d} \lambda_i = 1 \right\}.$$

Let $f = f_\sigma$ be the k-dimensional subsimplex spanned by the points $x_{\sigma_0}, \dots, x_{\sigma_k}$. Then, the linear polynomials $\lambda_{\sigma_1}, \dots, \lambda_{\sigma_k}$ form a set of barycentric coordinates for f_σ, that is,

$$\begin{aligned} f_\sigma &= \left\{ x \in T \middle| \lambda_j = 0 \text{ for } j \notin \sigma \right\}. \\ &= \left\{ x = \sum_{i\in\sigma} \lambda_i \middle| \lambda_i \geq 0, \quad \sum_{i\in\sigma} \lambda_i = 1 \right\}. \end{aligned} \tag{6.38}$$

Remark 6.3.21. When we introduced barycentric coordinates in order to define standard shape functions on simplices, we generated a basis for $\mathbb{P}_r(\mathbb{R}^d)$ by selecting polynomials of the λ_i. Closer inspection reveals that these polynomials were homogeneous. Therefore, we defined an isomorphism

$$\breve{\mathbb{P}}_k(\mathbb{R}^{d+1}) \equiv \mathbb{P}_k(\mathbb{R}^d), \tag{6.39}$$

which reads: for every $p \in \mathbb{P}_k(\mathbb{R}^d)$ there is $q \in \breve{\mathbb{P}}_k(\mathbb{R}^{d+1})$ such that

$$p(x_1, \dots, x_d) = q(\lambda_0, \dots, \lambda_d).$$

6.3.22 Definition: For each k-dimensional subsimplex f_σ of T with $\sigma = \sigma_0, \dots, \sigma_k$, the space $\mathbb{P}_r(f_\sigma) \equiv \breve{\mathbb{P}}_r(\mathbb{R}^{k+1})$ is defined as

$$\mathbb{P}_r(f_\sigma) = \left\{ q(\lambda_{\sigma_0}, \dots, \lambda_{\sigma_k}) \middle| q \in \breve{\mathbb{P}}_r(\mathbb{R}^{k+1}) \right\}. \tag{6.40}$$

The bubble function associated with f_σ is

$$b_{f\sigma} = \lambda_{\sigma_0} \cdots \lambda_{\sigma_k}. \tag{6.41}$$

The **extension operator** $E_{f_\sigma \to T}$ is defined as

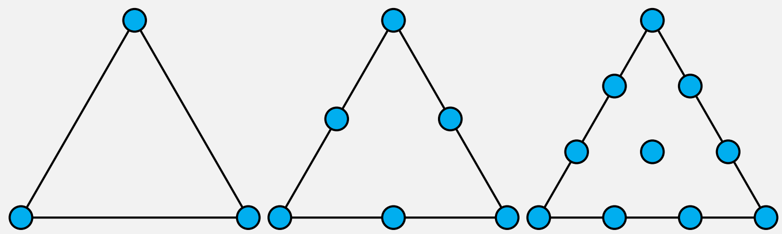
$$\begin{aligned} E_{f_\sigma \to T} &\colon \mathbb{P}_r(f_\sigma) \to \mathbb{P}_r(\mathbb{R}^d), \\ p(\lambda_0, \dots, \lambda_d) &= q(\lambda_{\sigma_0}, \dots, \lambda_{\sigma_k}), \end{aligned} \tag{6.42}$$

where q is chosen as in the definition of $\mathbb{P}_r(f_\sigma)$.

6.3.23 Lemma: Every function in $\mathbb{P}_r(f)$ vanishes on every subsimplex g which is disjoint from f.
The bubble function b_f vanishes on every subsimplex not containing f.

6.3.24 Problem: Show: $\mathbb{P}_r(f_\sigma)$ is isomorphic to $\mathbb{P}_r(\mathbb{R}^k)$. Prove Lemma 6.3.23.

6.3.25 Example:



Unisolvent interpolation conditions for $\mathbb{P}_r(T)$

$u(f) = 0$		$\dim f = 0$
$(u, q)_f = 0$	$q \in \mathbb{P}_{r-2}(f)$	$\dim f = 1$
$(u, q)_f = 0$	$q \in \mathbb{P}_{r-3}(f)$	$\dim f = 2$
$\vdots$	$\vdots$	$\vdots$

We are now generalizing and formalizing this example in order to derive a geometric decomposition of $\mathbb{P}_r(T)$ and its dual.

6.3.26 Definition: For every $f \in \Delta(T)$, we define $V(f) \subset \mathbb{P}_r(T)$ for $\dim f > 0$ as

$$V(f) = \left\{p = E_{f\to T} b_f q \middle| \, q \in \mathbb{P}_{r-\dim f-1}(f)\right\}, \tag{6.43}$$

and for $\dim f = 0$

$$V(f) = \left\{\lambda_i^r \middle| \, f = \{x_i\}\right\}. \tag{6.44}$$

6.3.27 Definition: For every $f \in \Delta(T)$, we define $W(f) \subset \mathbb{P}_r(T)^*$ for $\dim f > 0$ as

$$W(f) = \left\{\varphi(p) = (p, q)_f \middle| \, q \in \mathbb{P}_{r-\dim f-1}(f)\right\}, \tag{6.45}$$

and for $\dim f = 0$

$$W(f) = \left\{\varphi(p) = p(x_i) \middle| \, f = \{x_i\}\right\}. \tag{6.46}$$

6.3.28 Lemma: There holds

$$\mathbb{P}_r(T) = \bigoplus_{f\in\Delta(T)} V(f), \qquad \mathbb{P}_r(T)^* = \bigoplus_{f\in\Delta(T)} W(f). \tag{6.47}$$

Proof. We begin to show that

$$\mathbb{P}_r(T) = \bigoplus_{f\in\Delta(T)} V(f).$$

First, we note that for any $f \in \Delta(T)$ every function in $V(f)$ is also in $\mathbb{P}_r(T)$. For $\dim f = 0$, that is, $f = \{x_i\}$ for some vertex x_i, the only homogeneous polynomial of order r is λ_i^r.

For $\dim f > 0$ we have by the first statement of Lemma 6.3.23, that the spaces $V(f)$ where f is a vertex are disjoint. By the second statement of the same lemma, the spaces $V(f)$ for all f with equal dimension are disjoint. Therefore, the sum

$$V_d(T) = \sum_{k=0}^{d-1} \sum_{f\in\Delta_k(T)} V(f),$$

is direct. But, $V(T) \cap V_d(T) = \{0\}$, since all elements in $V(T)$ contain a bubble function factor. Therefore,

$$\bigoplus_{f\in\Delta(T)} V(f) \subset \mathbb{P}_r(T).$$

We conclude by showing that dimensions on both sides are equal. On the right, we use

$$\dim \mathbb{P}_r(\mathbb{R}^d) = \binom{r+d}{r} = \frac{(r+d)!}{d!r!}$$

On the left, we have

$$\begin{aligned}\dim \bigoplus_{f\in\Delta(T)} V(f) &= \sum_{k=0}^{d} \binom{d+1}{k+1}\binom{r+k}{k} \\ &= \frac{(d+1)!}{r!} \sum_{k=0}^{d} \frac{1}{(k+1)!(d-k)!} \frac{(r+k)!}{k!}.\end{aligned}$$

It remains to show the decomposition for $\mathbb{P}_r(T)^*$. To this end, we first notice that for any $f \in \Delta(T)$ there holds $\dim W(f) = \dim V(f)$ by their definition. Furthermore, for $p \in V(f)$ there holds

$$\Bigl(\varphi(p) = 0 \quad \forall \varphi \in W(f)\Bigr) \quad \Rightarrow \quad p = 0.$$

Thus, for $p \in \mathbb{P}_r(T)$ there holds

$$\Big(\varphi(p) = 0 \quad \forall \varphi \in \sum W(f)\Big) \quad \Rightarrow \quad p = 0.$$

Consequently,

$$\mathbb{P}_r(T)^* = \sum W(f).$$

since we have already proven that

$$\dim \mathbb{P}_r(T)^* = \sum \dim W(f),$$

the sum on the right must be direct. □

Results for $\mathbb{P}_r\Lambda^k$ and $\mathbb{P}_r^+\Lambda^k$ and applications

For polynomial differential forms on simplices, we cite the main results without proof and then discuss their application to H^{div} and H^{curl}.

6.3.29 Theorem: Let $k, r \geq 1$. Then, $\mathbb{P}_r^+(T)$ admits a geometric decomposition

$$\mathbb{P}_r^+\Lambda^k(T) = \bigoplus_{f\in\Delta(T)} V(f), \qquad \mathbb{P}_r^+\Lambda^k(T)^* = \bigoplus_{f\in\Delta(T)} W(f), \tag{6.48}$$

where

$$V(f) \equiv \begin{cases} 0 & \dim f < k \\ \mathbb{P}_{r+k-\dim f}\Lambda^{\dim f - k}(f) & \text{else} \\ 0 & \dim f > r+k. \end{cases} \tag{6.49}$$

$$W(f) \equiv \begin{cases} 0 & \dim f < k \\ \mathbb{P}_{r+k-\dim f}\Lambda^{\dim f - k}(f) & \text{else} \\ 0 & \dim f > r+k. \end{cases} \tag{6.50}$$

6.3.30 Example:

$\dim f$	$\mathbb{P}_r^+\Lambda^0$	$\mathbb{P}_{r+1}$	$\mathbb{P}_r^+\Lambda^1$	N_r^{1e}	$\mathbb{P}_r^+\Lambda^2$	RT_r
3	$\mathbb{P}_{r-3}\Lambda^3$	$\mathbb{P}_{r-3}$	$\mathbb{P}_{r-2}\Lambda^2$	BDM_{r-2}	$\mathbb{P}_{r-1}\Lambda^1$	N_{r-1}^{2e}
2	$\mathbb{P}_{r-2}\Lambda^2$	$\mathbb{P}_{r-2}$	$\mathbb{P}_{r-1}\Lambda^1$	BDM_{r-1}	$\mathbb{P}_r\Lambda^0$	$\mathbb{P}_r$
1	$\mathbb{P}_{r-1}\Lambda^1$	$\mathbb{P}_{r-1}$	$\mathbb{P}_r\Lambda^0$	$\mathbb{P}_r$	$--$	$--$
0	$\mathbb{R}$	$\mathbb{R}$	$--$	$--$	$--$	$--$

(6.51)

Geometric decomposition of $\mathbb{P}_r^+\Lambda^k$ and their spaces of degrees of freedom.

N^{1e} (H^{curl}) Nedelec 1st family edge element

RT (H^{div}) Raviart-Thomas (also Nedelec 1st face in 3D)

N^{2e} (H^{curl}) Nedelec 2nd family edge element

6.3.31 Theorem: Let $k, r \geq 1$. Then, $\mathbb{P}_r(T)$ admits a geometric decomposition

$$\mathbb{P}_r\Lambda^k(T) = \bigoplus_{f\in\Delta(T)} V(f), \tag{6.52}$$

where

$$V(f) \equiv \begin{cases} 0 & \dim f < k \\ \mathbb{P}^+_{r+k-\dim f-1}\Lambda^{\dim f-k}(f) & \text{else} \\ 0 & \dim f \geq r+k. \end{cases} \tag{6.53}$$

6.3.32 Example:

$\dim f$	$\mathbb{P}_r\Lambda^1$	N_r^{2e}	$\mathbb{P}_r\Lambda^2$	BDM_r
3	$\mathbb{P}^+_{r-3}\Lambda^2$	RT_{r-2}	$\mathbb{P}^+_{r-1}\Lambda^1$	N_{r-1}^{1e}
2	$\mathbb{P}^+_{r-2}\Lambda^1$	RT_{r-1}	$\mathbb{P}^+_r\Lambda^0$	$\mathbb{P}_r$
1	$\mathbb{P}^+_{r-1}\Lambda^0$	$\mathbb{P}_r$	$--$	$--$

(6.54)

Geometric decomposition of $\mathbb{P}_r\Lambda^k$ and their spaces of degrees of freedom.

N^{2e} (H^{curl}) Nedelec 2nd family edge element

BDM (H^{div}) Brezzi-Douglas-Marini (also Nedelec 2nd face in 3D)

N^{1e} (H^{curl}) Nedelec 1st family edge element

6.4 The complex of tensor product polynomials

6.4.1 Lemma: Tensor product polynomials form the exact sequence

$$\mathbb{R} \xrightarrow{\subset} \mathbb{Q}_r \xrightarrow{\nabla} \begin{pmatrix} \mathbb{Q}_{r,r+1,r+1} \\ \mathbb{Q}_{r+1,r,r+1} \\ \mathbb{Q}_{r+1,r+1,r} \end{pmatrix} \xrightarrow{\nabla\times} \begin{pmatrix} \mathbb{Q}_{r+1,r,r} \\ \mathbb{Q}_{r,r+1,r} \\ \mathbb{Q}_{r,r,r+1} \end{pmatrix} \xrightarrow{\nabla\cdot} \mathbb{Q}_r \longrightarrow 0,$$

Proof. First, we show that the differential operators map into the right spaces. Let $q \in \mathbb{Q}_{r+1}$ such that $q(x) = q_1(x_1)q_2(x_2)q_3(x_3)$ with $q_i \in \mathbb{Q}_{r+1}$. Then

$$\nabla q = \begin{pmatrix} q_1' q_2 q_3 \\ q_1 q_2' q_3 \\ q_1 q_2 q_3' \end{pmatrix} \in \begin{pmatrix} \mathbb{Q}_{r,r+1,r+1} \\ \mathbb{Q}_{r+1,r,r+1} \\ \mathbb{Q}_{r+1,r+1,r} \end{pmatrix}.$$

Similarly, we can compute this directly for $\nabla\times$ and $\nabla\cdot$. Since polynomials are differentiable, we have $d^2 = 0$. It remains to show that the sequence is exact, which we will prove at the example of H^{curl}. Let $u \in \ker(\nabla\times)$,

$$u = \begin{pmatrix} \varphi_1\varphi_2\varphi_3 \\ \psi_1\psi_2\psi_3 \\ \pi_1\pi_2\pi_3 \end{pmatrix}, \qquad 0 = \nabla\times u = \begin{pmatrix} \pi_1\pi_2'\pi_3 - \psi_1\psi_2\psi_3' \\ \varphi_1\varphi_2\varphi_3' - \pi_1'\pi_2\pi_3 \\ \psi_1'\psi_2\psi_3 - \varphi_1\varphi_2'\varphi_3 \end{pmatrix},$$

where each polynomial with index i only depends on x_i. Furthermore, $\varphi_1, \psi_2, \pi_3 \in \mathbb{P}_r$ and all other in $\mathbb{P}_{r+1}$. From the continuous de Rham complex, we know that there is a function p such that $u = \nabla p$. It remains to show that $p \in \mathbb{Q}_{r+1}$. There holds

$$\partial_1 p = \varphi_1\varphi_2\varphi_3.$$

Thus, we make the ansatz

$$p = \Phi_1\varphi_2\varphi_3,$$

where $\Phi_1 \in \mathbb{P}_{r+1}$ is the antiderivative of φ_1. Thus, $p \in \mathbb{Q}_{r+1}$ It remains to show that this ansatz is consistent with the other two derivatives, thus,

$$p = \Phi_1\varphi_2\varphi_3 = \psi_1\Psi_2\psi_3 = \pi_1\pi_2\Pi_3.$$

Integrating the first component of $\nabla\times u$ with respect to x_2 and x_3, we obtain

$$\pi_1\pi_2\Pi_3 = \psi_1\Psi_2\psi_3.$$

Doing the same with the second component, we see indeed that the function p is consistently defined and thus $u = \nabla p$. □

6.4.2 Definition: The one-dimensional de Rham complex on the interval $I = [\xi_0, \xi_1]$ and its polynomial subcomplex are

$$
\begin{array}{ccccccc}
\mathbb{R} & \xrightarrow{\subset} & H^1(I) = H\Lambda^0(I) & \xrightarrow{\frac{d}{dx}} & L^2(I) = H\Lambda^1(I) & \longrightarrow & 0 \\
& & \subset\Big\uparrow & & \subset\Big\uparrow & & \\
\mathbb{R} & \xrightarrow{\subset} & \mathbb{P}_{r+1} = \mathbb{P}_{r+1}\Lambda^0 & \xrightarrow{\frac{d}{dx}} & \mathbb{P}_r = \mathbb{P}_r\Lambda^1 & \longrightarrow & 0
\end{array}
\tag{6.55}
$$

The degrees of freedom for $\mathbb{P}_{r+1}\Lambda^0$ are

$$
\begin{aligned}
\mathcal{N}_{0,0}(p) &= p(\chi_0), \\
\mathcal{N}_{0,1}(p) &= p(\chi_1),
\end{aligned}
\qquad \mathcal{N}_{1.q}(p) = \int_I pq\,\mathrm{d}x, \quad \forall q \in \mathbb{P}_{r-1}. \tag{6.56}
$$

The degrees of freedom for $\mathbb{P}_r\Lambda^1$ are

$$
\mathcal{N}_{1,q}(p) = \int_I pq\,\mathrm{d}x, \quad \forall q \in \mathbb{P}_r. \tag{6.57}
$$

Remark 6.4.3. The degrees of freedom for $\mathbb{P}_{r+1}\Lambda^0$ are chosen such that the finite element function on a subdivision of I is continuous, thus ih H^1. For $\mathbb{P}_r\Lambda^1$, we do not require continuity and thus only need interior degrees of freedom.

6.4.4 Definition: Let

$$
\mathcal{N}_{0,0}p = p(chi_0), \qquad \mathcal{N}_{0,1}p = p(\chi_1), \qquad \mathcal{N}_{1,i} = \int_I pq_i\,\mathrm{d}x, \tag{6.58}
$$

be the degrees of freedom of a one-dimensional element where the q_i are a basis for $\mathbb{P}_{r-1}$ and $\mathbb{P}_r$ in case of $\mathbb{P}_{r+1}\Lambda^0$ and $\mathbb{P}_r\Lambda^1$, respectively. Then, the tensor product of these degrees of freedom applied to the function $p(x_1, x_2) = p_1(x_1)p_2(x_2)$ is defined as

$$
\begin{aligned}
\mathcal{N}_{0,i} \otimes \mathcal{N}_{0,j}(p) &= p_1(\chi_i)p_2(\chi_j), \\
\mathcal{N}_{0,i} \otimes \mathcal{N}_{1,j}(p) &= \int_I p_1(\chi_i)p_2(x)q_j(x)\,\mathrm{d}x, \\
\mathcal{N}_{1,i} \otimes \mathcal{N}_{0,j}(p) &= \int_I p_1(x)q_j(x)p_2(\chi_j)\,\mathrm{d}x, \\
\mathcal{N}_{1,i} \otimes \mathcal{N}_{1,j}(p) &= \iint_I p_1(x_1)p_2(x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2.
\end{aligned}
\tag{6.59}
$$

6.4.5 Lemma: For the two elements

$$RT_r = \begin{pmatrix} \mathbb{P}_{r+1}\Lambda^0 \otimes \mathbb{P}_r\Lambda^1 \\ \mathbb{P}_r\Lambda^1 \otimes \mathbb{P}_{r+1}\Lambda^0 \end{pmatrix}, \qquad N_r^{1e} = \begin{pmatrix} \mathbb{P}_r\Lambda^1 \otimes \mathbb{P}_{r+1}\Lambda^0 \\ \mathbb{P}_{r+1}\Lambda^0 \otimes \mathbb{P}_r\Lambda^1 \end{pmatrix}, \tag{6.60}$$

the tensorized degrees of freedom uniquely determine the normal and tangential components, respectively, on the face of a square.
Both elements with their tensor degrees of freedom are unisolvent.

Proof. The elements are unisolvent by the following argument: Let $\{\varphi_i\}$ and $\{\psi_j\}$ be the basis dual to the degrees of freedom of $\mathbb{P}_{r+1}\Lambda^0$ or $\mathbb{P}_r\Lambda^1$, respectively. Then, (renumbering the degrees of freedom)

$$\mathcal{N}_k \otimes \mathcal{N}_l(\varphi_i \otimes \psi_j) = \delta_{ik}\delta_{jl}.$$

Thus, the mapping between tensorized degrees of freedom and tensorized basis functions is one-to-one.

For the Raviart-Thomas element RT_r, the normal component is $\mathbb{P}_r\Lambda^1$. Take for instance the face $x_1 = \chi_0$. Then, the degrees of freedom associated to this face are

$$\mathcal{N}_{0,0} \otimes \mathcal{N}1, qp = \int_I p_1(\chi_1)p_2(x_2)q(x_2)\,\mathrm{d}x_2 \quad \forall q \in \mathbb{P}_r.$$

Since the trace of p on this face is $p_2 \in \mathbb{P}_r$, this polynomial is uniquely determined by the degree of freedom.

The argument for the Nedelec edge element N_r^{1e} follows by exchanging tangential and normal component. □

Remark 6.4.6. The construction extends to the three-dimensional products.

Bibliography

[ABF02] D. N. Arnold, D. Boffi, and R. S. Falk. "Approximation by quadrilateral finite elements." In: *Math. Comput.* 71 (2002), pp. 909–922.

[ABF05] D. N. Arnold, D. Boffi, and R. S. Falk. "Quadrilateral $H(div)$ finite elements." In: *SIAM J. Numer. Anal.* 42.6 (2005), pp. 2429–2451.

[AF03] R. A. Adams and J. J. F. Fournier. *Sobolev Spaces*. 2nd edition. Vol. 140. Pure and Applied Mathematics. Academic Press, 2003.

[AFW06] Douglas N. Arnold, Richard S. Falk, and Ragnar Winther. "Finite element exterior calculus, homological techniques, and applications." In: *Acta Numerica* 15 (2006), pp. 1–155. DOI: 10.1017/S0962492906210018.

[AFW10] Douglas N. Arnold, Richard S. Falk, and Ragnar Winther. "Finite element exterior calculus: from Hodge theory to numerical stability." In: *Bull. Amer. Math. Soc. (N.S.)* 47.2 (2010), pp. 281–354. ISSN: 0273-0979. DOI: 10.1090/S0273-0979-10-01278-4. URL: http://dx.doi.org/10.1090/S0273-0979-10-01278-4.

[BFB13] D. Boffi, M. Fortin, and F. Brezzi. *Mixed finite element methods and applications*. Springer series in computational mathematics. Berlin, Heidelberg: Springer, 2013.

[Bra13] D. Braess. *Finite Elemente*. 5. Aufl. Springer, 2013.

[Bra97] D. Braess. *Finite elements*. Cambridge University Press, 1997.

[Cia88] Ph. G. Ciarlet. *Mathematical Elasticity. Vol. I: Three-dimensional Elasticity*. Vol. 20. Studies in mathematics and its applications. North-Holland, 1988.

[DL76] G. Duvaut and J. L. Lions. *Inequalities in Mechanics and Physics*. Springer, 1976.

[EG04] A. Ern and J.-L. Guermond. *Theory and practice of finite elements*. Vol. 159. Applied Mathematical Sciences. Springer-Verlag, New York, 2004. DOI: 10.1007/978-1-4757-4355-5.

[GKR14] V. Girault, G. Kanschat, and B. Rivière. "Error analysis for a monolithic discretization of coupled Darcy and Stokes problems." In: *J. Numer. Math.* 22.2 (2014), pp. 109–142. DOI: 10.1515/jnma-2014-0005.

[GR86] V. Girault and P.-A. Raviart. *Finite element approximations of the Navier-Stokes equations*. Springer, 1986.

[Mon03] P. Monk. *Finite Element Methods for Maxwell's Equations*. Oxford, 2003.

[SS96] R. Stenberg and M. Suri. "Mixed *hp* finite element methods for problems in elasticity and Stokes flow." In: *Numer. Math.* 72.3 (1996), pp. 367–389. DOI: 10.1007/s002110050174.

[Ste90] R. Stenberg. "A technique for analysing finite element methods for viscous incompressible flow." In: *Int. J. Numer. Methods Fluids* 11.6 (1990), pp. 935–948. DOI: 10.1002/fld.1650110615.

[Ver84] R. Verfürth. "Error estimates for a mixed finite element approximation of Stokes problem." In: *RAIRO Anal. Numer.* 18.2 (1984), pp. 175–182. DOI: 10.1051/m2an/1984180201751.

[Yos80] K. Yosida. *Functional Analysis*. Springer, 1980.

Index