An Abstract Framework of Pressure Robustness for Saddle Point Problems in Hilbert Spaces

by

Yuan (Greg) Wang

A research paper
presented to the University of Waterloo
in partial fulfillment of the
requirement for the degree of
Master of Mathematics
in
Computational Mathematics

Supervisor:
Dr. Sander Rhebergen

Waterloo, Ontario, Canada, 2019

Abstract

In recent years pressure-robust finite element methods have become popular for incompressible flows as they are generally more accurate than finite element methods that are not pressure-robust. The main concept behind pressure-robust discretizations, in which the velocity error is independent of the best approximation error of the pressure, was introduced in [A. Linke and C. Merdon, Comput. Methods Appl. Mech. Engrg. 2016].

However, pressure-robustness is not a new idea. For a general saddle point problem, a priori error estimates are studied in [Brezzi and Fortin, Mixed and Hybrid Finite Element Method Springer-Verlag 1991] in which one of the unknowns admits estimates that are independent of the other unknown. For the generalized case, we call this 'Q-robustness'.

Unfortunately, the Q-robustness analysis in this book by Brezzi and Fortin applies only to conforming finite element methods and needs to be extended for the nonconforming case. This is the topic of this research paper; we extend the analysis of Q-robust error estimates presented in the book by Brezzi and Fortin to nonconforming finite element methods.

We give sufficient conditions under which Q-robustness is achieved. We then apply the extended theory to an $H(\mathrm{div})$-conforming discontinuous Galerkin method for Stokes equations. We provide a proof of Q-robustness for the Stokes problem.

Table of Contents

Chapter 1

Introduction

1.1 The Saddle Point Problem

A classical quadratic energy functional corresponding to a **minimization principle** is given by

$$J(v) = \frac{1}{2}a(v, v) - l(v), \tag{1.1a}$$

$$u = \arg\min_v J(v). \tag{1.1b}$$

The associated variational formulation is given by

$$a(u, v) = l(v) \quad \forall v \in V, \tag{1.1c}$$

with V a suitable function space and the solution u minimizes the energy functional $J(\cdot)$. The Galerkin method seeks a solution $u_h \in V_h \subset V$, with V_h a suitable finite dimensional subspace of V, such that

$$a(u_h, v_h) = l(v_h) \quad \forall v_h \in V_h, \tag{1.1d}$$

A finite element is a Galerkin method in which V_h is the space of piecewise polynomials.

Alternatively, a variational principle can also express a **saddle point equilibrium**. Let V and Q be suitable function spaces. A pair $(u, p) \in V \times Q$ is said to be a saddle-point of a Lagrangian functional $\mathcal{L} : V \times Q \to \mathbb{R}$ if

$$\mathcal{L}(u, q) \leq \mathcal{L}(u, p) \leq \mathcal{L}(v, p) \quad \forall (v, q) \in V \times Q. \tag{1.2a}$$

The motivation behind the name "saddle-point" can be seen by the following example where we have a Lagrangian functional $z = \mathcal{L}(x, y) = x^2 - y^2$ and for $x = 0$ and $y = 0$ it holds that (see

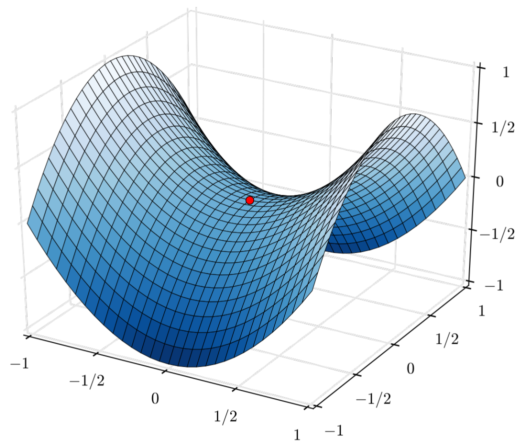


Figure 1.1: $z = \mathcal{L}(x, y) = x^2 - y^2$, $(0, 0)$ is the saddle point which is marked red.

fig. 1.1):

$$\mathcal{L}(0, y) \leq \mathcal{L}(0, 0) \leq \mathcal{L}(x, 0) \quad \forall (x, y) \in \mathbb{R}^2. \tag{1.2b}$$

If the Lagrangian functional $\mathcal{L} : V \times Q \to \mathbb{R}$ is defined as [8, Proposition 2.39]

$$\mathcal{L}(v, q) = \frac{1}{2} a(v, v) + b(v, q) - f(v) - g(q), \tag{1.2c}$$

the variational formulation of the saddle point problem is given by

$$a(u, v) + b(v, p) = f(v) \quad \forall v \in V, \tag{1.2d}$$
$$b(u, q) = g(q) \quad \forall q \in Q, \tag{1.2e}$$

and the discrete problem is given by

$$a(u_h, v_h) + b(v_h, p_h) = f(v_h) \quad \forall v_h \in V_h \subset V, \tag{1.2f}$$
$$b(u_h, q_h) = g(q_h) \quad \forall q_h \in Q_h \subset Q. \tag{1.2g}$$

Using piecewise polynomial function spaces, we obtain the **mixed finite element method**.

1.2 Pressure-robustness for the Stokes Problem

The Stokes problem is given by

$$-\nu\nabla^2\boldsymbol{u} + \nabla p = \boldsymbol{f}, \tag{1.3a}$$

$$-\nabla\cdot\boldsymbol{u} = g, \tag{1.3b}$$

with suitable Dirichlet and/or Neumann boundary conditions. In variational form,

$$a(\boldsymbol{u},\boldsymbol{v}) + b(\boldsymbol{v},p) = (\boldsymbol{f},\boldsymbol{v}), \quad \forall \boldsymbol{v}\in\boldsymbol{V}, \tag{1.4a}$$

$$b(\boldsymbol{u},q) = (g,q), \quad \forall q\in Q, \tag{1.4b}$$

where the biliear forms are given by

$$a(\boldsymbol{u},\boldsymbol{v}) := \nu(\nabla\boldsymbol{u},\nabla\boldsymbol{v}) = \nu\int_\Omega \nabla\boldsymbol{u} : \nabla\boldsymbol{v}, \tag{1.4c}$$

$$b(\boldsymbol{v},q) := -(\nabla\cdot\boldsymbol{v},q) = -\int_\Omega q\nabla\cdot\boldsymbol{v}. \tag{1.4d}$$

The corresponding discrete problem is given by

$$a(\boldsymbol{u}_h,\boldsymbol{v}_h) + b(\boldsymbol{v}_h,p_h) = (\boldsymbol{f},\boldsymbol{v}_h), \quad \forall \boldsymbol{v}_h\in\boldsymbol{V}_h\subset\boldsymbol{V}, \tag{1.4e}$$

$$b(\boldsymbol{u}_h,q_h) = (g,q_h), \quad \forall q_h\in Q_h\subset Q. \tag{1.4f}$$

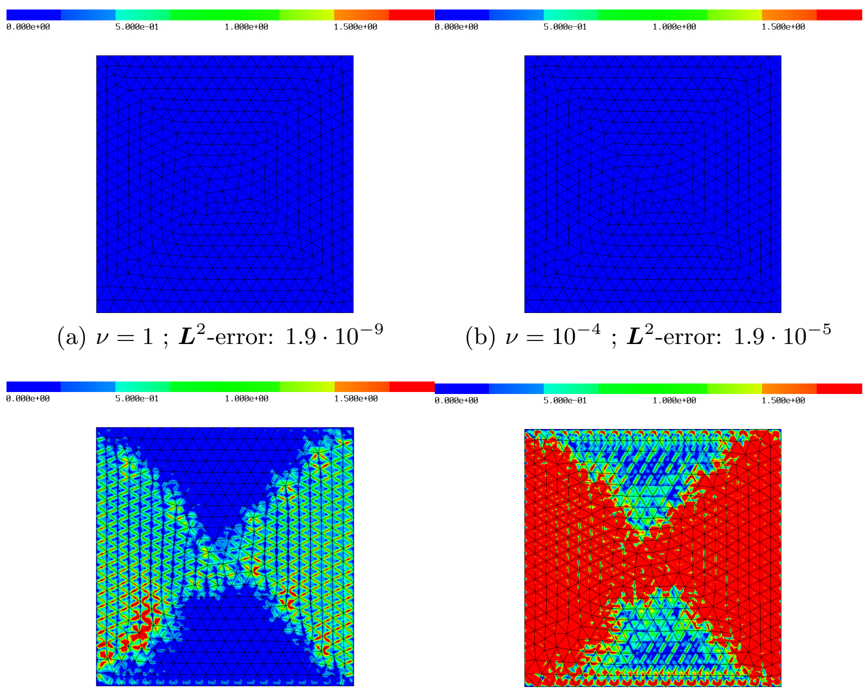
An *a priori* error analysis of many mixed finite element methods results in the following velocity error estimate:

$$\|\boldsymbol{u}-\boldsymbol{u}_h\| \lesssim \inf_{\boldsymbol{v}_h\in\boldsymbol{V}_h}\|\boldsymbol{u}-\boldsymbol{v}_h\| + \nu^{-1}\inf_{q_h\in Q_h}\|p-q_h\|. \tag{1.5}$$

It is clear that the estimate of the velocity error depends on the best approximation error of the pressure scaled by the inverse of the viscosity ν. This is undesirable when ν is small, as we show in the following example.

Example 1.1 (Lack of pressure-robustness for Taylor–Hood)**.** *Consider the Stokes equations in the unit square. Set* $\boldsymbol{u}=\boldsymbol{0}$ *on the boundary and* $\boldsymbol{f}$ *and* g *such that the exact solution is given by* $(\boldsymbol{u},p)=(\boldsymbol{0},cos(y))$*. To discretize the Stokes problem we use the* $\boldsymbol{\mathcal{P}}_2-\mathcal{P}_1$ *Taylor–Hood finite element spaces. Here* $\boldsymbol{\mathcal{P}}_2$ *is the space of continuous piecewise polynomials of degree two. Similarly,* $\mathcal{P}_1$ *is the space of continuous piecewise polynomials of degree one. We plot the solution in fig. 1.2 for different values of* ν*. We observe that as* ν *decreases, the error in* $\boldsymbol{u}$ *increases by* ν^{-1}*.*

Under certain conditions on the function spaces, it is possible to obtain an *a priori* error



(a) $\nu = 1$; $\boldsymbol{L}^2$-error: $1.9 \cdot 10^{-9}$

(b) $\nu = 10^{-4}$; $\boldsymbol{L}^2$-error: $1.9 \cdot 10^{-5}$

(c) $\nu = 10^{-9}$; $\boldsymbol{L}^2$-error: $1.9 \cdot 10^{0}$

(d) $\nu = 10^{-10}$; $\boldsymbol{L}^2$-error: $1.9 \cdot 10^{+1}$

Figure 1.2: $\boldsymbol{\mathcal{P}}_2 - \mathcal{P}_1$ Taylor–Hood elements to solve the Stokes problem, see example 1.1.

estimate of the velocity that satisfies

$$\|\boldsymbol{u} - \boldsymbol{u}_h\| \lesssim \inf_{\boldsymbol{v}_h \in \boldsymbol{V}_h} \|\boldsymbol{u} - \boldsymbol{v}_h\| \,. \tag{1.6}$$

Such an error estimate is said to be *pressure-robust*; the error in the velocity does not depend on the best approximation error of the pressure scaled by the inverse of the viscosity. In example 1.2 we demonstrate the advantages of the pressure-robust scheme.

Example 1.2 (A pressure-robust finite element method)**.** *Consider the same setup as in example 1.1. We now use the Scott–Vogelius finite element to solve the Stokes problem on a mesh created as a barycentric refinement of a regular triangular mesh. For the Scott–Vogelius element we consider a piecewise continuous polynomial space of degree two for the velocity. For the pressure we use a piecewise discontinuous polynomial space of degree one. Note that the difference between the Scott–Vogelius and the Taylor–Hood element is that the Scott–Vogelius element approximates the pressure by discontinuous polynomials. The Taylor–Hood element uses continuous polynomials. We plot the solution in fig. 1.3 for different values of* ν*. We now observe that* ν *has no effect on the error in* $\boldsymbol{u}$*. The Scott–Vogelius element is pressure-robust. We remark that the increase in the* $\boldsymbol{L}^2$*-error of the velocity comes from a machine precision term which dominates the velocity error in this test case.*

The comparison between Taylor–Hood and Scott–Vogelius shows that the Scott–Vogelius element is significantly more accurate in approximating the velocity than the Taylor–Hood element. This is due to the special choice of the pair of velocity and pressure spaces. As we will see, for the Stokes problem it is important the divergence of the approximate velocity lies in the pressure space.

Remark 1.1. *Both example 1.1 and example 1.2 were implemented using NGSolve [20].*

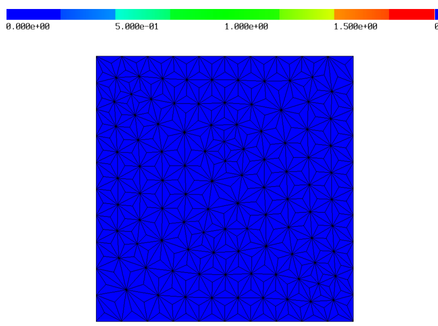
1.3 Goal of This Research Paper

As we saw in the previous section, the finite element used to discretize the Stokes problem can have a significant impact on the accuracy of the velocity solution $\boldsymbol{u}_h$. In particular, a pressure-robust finite element scheme results in a velocity solution of which its accuracy does not depend on the pressure solution and the viscosity.

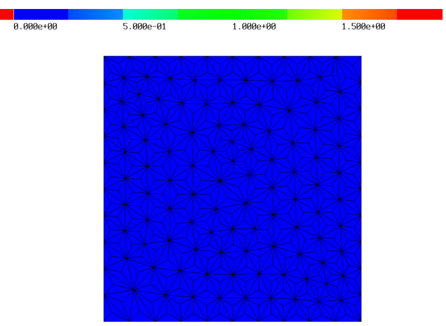
This property can be generalized to a general saddle point problem. Consider a saddle point problem

$$a(u, v) + b(v, p) = f(v) \quad \forall v \in V, \tag{1.7a}$$

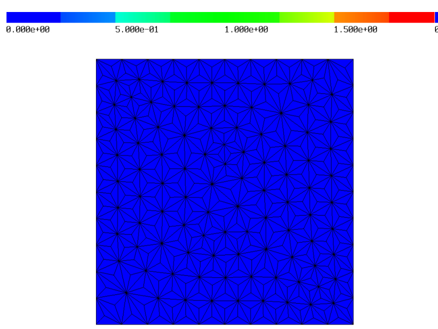
$$b(u, q) = g(q) \quad \forall q \in Q, \tag{1.7b}$$



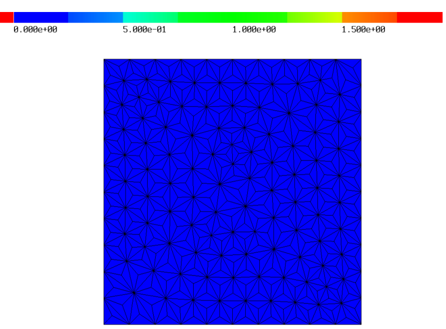
(a) $\nu = 1$; $\boldsymbol{L}^2$-error: $1.5 \cdot 10^{-12}$



(b) $\nu = 10^{-4}$; $\boldsymbol{L}^2$-error: $1.5 \cdot 10^{-8}$



(c) $\nu = 10^{-6}$; $\boldsymbol{L}^2$-error: $1.5 \cdot 10^{-6}$



(d) $\nu = 10^{-8}$; $\boldsymbol{L}^2$-error: $1.5 \cdot 10^{-4}$

Figure 1.3: Scott–Vogelius finite element, on a mesh with barycentric refinement, to solve the Stokes problem, see example 1.2.

and its finite element discretization,

$$a_h(u_h, v_h) + b_h(v_h, p_h) = f(v_h) \quad \forall v_h \in V_h, \tag{1.8a}$$

$$b_h(u_h, q_h) = g(q_h) \quad \forall q_h \in Q_h. \tag{1.8b}$$

where $a_h(\cdot,\cdot)$ and $b_h(\cdot,\cdot)$ may differ from, but are consistent with, the bilinear forms $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$, respectively.

In this research paper we investigate under what condition(s) the discrete system eq. (1.8) admits an *a priori* estimate for $\|u - u_h\|$ that is independent of $\|p - p_h\|$, i.e.,

$$\|u - u_h\| \lesssim \inf_{v_h \in V_h} \|u - v_h\|. \tag{1.9}$$

In the literature the most commonly used notation for the function spaces of a general variation saddle point problem are V and Q. Our goal is to find a bound for $u \in V$ which is independent of $p \in Q$. For this reason we refer to the property eq. (1.9) as **Q-robustness**. We note that in the Stokes case, Q-robustness reduces to pressure-robustness.

1.4 Outline of This Research Paper

Conforming finite element methods for the saddle point problem eq. (1.7) have been studied in [4]. They provide sufficient conditions for conforming finite element methods to be Q-robust. We study this theoretical framework in more detail in chapter 2.

Although non-conforming and pressure-robust finite element methods are well understood for the Stokes problem, c.f., [11, 13, 14, 16, 17], a theoretical framework of Q-robustness for non-conforming methods for a general saddle point problem is missing. In chapter 3 we will provide this theoretical framework.

In chapter 4 we apply the generalized Q-robustness theory to non-conforming finite element approximations of the Stokes problem. We will show that the analysis reduces to existing pressure-robustness analysis. We conclude this research paper in chapter 5.

Chapter 2

Conforming FEM for Saddle Point Problems

In this chapter, we introduce the basic saddle point problem and cast it in operator form (section 2.1). We then give equivalent statements of the different properties of the operator. This will enable us to make equivalent statements about the well-posedness of our problem, see (section 2.2). These preliminary results will aid in proving well-posedness in section 2.3, obtaining a basic error estimate in section 2.4, and a Q-robust error estimate in section 2.5.

2.1 The Basic Saddle Point Problem

Consider two Hilbert spaces, V and Q, and two continuous bilinear forms: $a(\cdot,\cdot)$ on $V \times V$ and $b(\cdot,\cdot)$ on $V \times Q$. The basic saddle point problem reads: Given $f \in V'$and $g \in Q'$, find $(u,p) \in V \times Q$ such that,

$$a(u,v) + b(v,p) = \langle f, v\rangle_{V' \times V} \quad \forall v \in V, \tag{2.1a}$$

$$b(u,q) = \langle g, q\rangle_{Q' \times Q} \quad \forall q \in Q. \tag{2.1b}$$

The continuous bilinear forms are related to bounded linear operators. We can associate with $b(\cdot,\cdot)$ on $V \times Q$ the bounded linear operator B from V to Q' defined as

$$\langle Bv, q\rangle_{Q',Q} := b(v,q) \quad \forall v \in V,\ \forall q \in Q. \tag{2.2}$$

We may similarly associate with $a(\cdot,\cdot)$ on $V \times V$ a bounded linear operator A from V to V'.

We note that a bilinear form is continuous if and only if its associated linear operator is continuous.

We define the **transposed operator** B^t from Q to V' by

$$\langle v, B^t q \rangle_{V,V'} := b(v,q) \quad \forall v \in V,\ \forall q \in Q. \tag{2.3}$$

Note that a transposed operator is a generalization of a transposed matrix, as shown in the next remark.

Remark 2.1. *We note that if $V := \mathbb{R}^n$ and $Q := \mathbb{R}^m$, by the Riesz Representation Theorem (see theorem 2.4), dual operations can be identified with scalar products, and so the linear operator $B : V \to Q'$ (which is always bounded when V is finite dimensional) can be identified as an $m \times n$ matrix,*

$$\langle B\vec{v}, \vec{q} \rangle = \vec{q}^T B\vec{v} = \vec{v}^T B^T \vec{q} = \langle \vec{v}, B^T \vec{q} \rangle \quad \forall \vec{v} \in \mathbb{R}^n,\ \forall \vec{q} \in \mathbb{R}^m.$$

In operator form, we may write eq. (2.1) as

$$Au + B^t p = f \quad \text{in } V', \tag{2.4a}$$

$$Bu = g \quad \text{in } Q'. \tag{2.4b}$$

Let V_h and Q_h be finite-dimensional subspaces of V and Q, respectively. The discrete problem approximating eq. (2.1) is given by: Find $(u_h, p_h) \in V_h \times Q_h$ such that:

$$a(u_h, v_h) + b(v_h, p_h) = \langle f, v_h \rangle_{V_h' \times V_h} \quad \forall v_h \in V_h, \tag{2.5a}$$

$$b(u_h, q_h) = \langle g, q_h \rangle_{Q_h' \times Q_h} \quad \forall q_h \in Q_h. \tag{2.5b}$$

Accordingly, the discrete problem in operator form is given by:

$$A_h u_h + B_h^t p_h = f \quad \text{in } V_h', \tag{2.6a}$$

$$B_h u_h = g \quad \text{in } Q_h', \tag{2.6b}$$

where $B_h : V_h \to Q_h'$ is defined as

$$\langle B_h v_h, q_h \rangle_{Q_h', Q_h} := b(v_h, q_h) \quad \forall v_h \in V_h,\ \forall q_h \in Q_h. \tag{2.7}$$

The operator $A_h : V_h \to V_h'$ is defined similarly.

We next list a few useful definitions.

Definition 2.1 (The kernel and image of a linear operator)**.** *Let D and R be Hilbert spaces (complete normed vector spaces equipped with inner products). Let $M : D \to R$ be a linear operator. The kernel and image of M, denoted respectively by $Ker(M)$ and $Im(M)$, are defined*

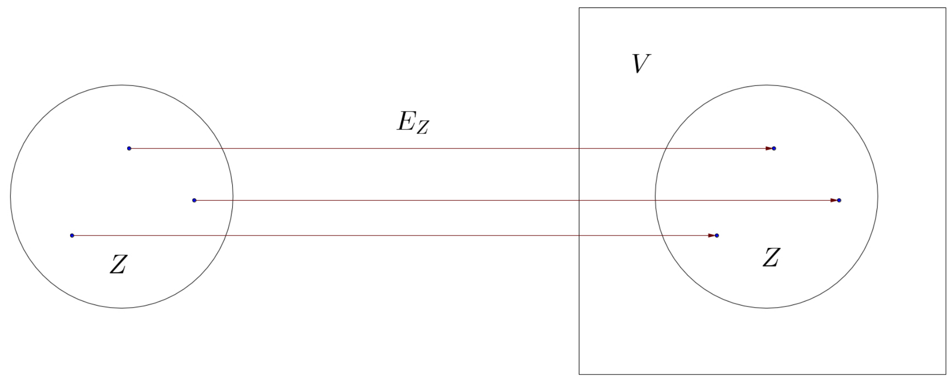


Figure 2.1: An example of the extension operator $E_{Z\to V}$.

as

$$Ker(M) := \{d \in D \text{ such that } Md = 0\},$$
$$Im(M) := \{r \in R \text{ such that } \exists\, d \in D \text{ with } Md = r\}.$$

The following definitions are required to establish the relation between B and B_h in a compact form (see [4, Chapters 4 and 5]).

Definition 2.2 (Extension Operator)**.** *Given a subspace $Z \subseteq V$, the extension operator $E_{Z\to V} : Z \to V$ associates every $z \in Z$ with the same z, thought as an element of V. Where no confusion may occur we will simply denote the extension operator by E_Z. See fig. 2.1.*

Definition 2.3 (Orthogonal Complement)**.** *Given a Hilbert space H and a linear subspace $Z \subseteq H$. The orthogonal complement of Z, denoted by $Z^\perp$, is defined as*

$$Z^\perp := \left\{v \in H \text{ such that } (v, z)_H = 0\ \forall z \in Z\right\}, \tag{2.8}$$

where $(\cdot,\cdot)_H$ is the inner product on H.

Definition 2.4 (Projection Operator)**.** *Let $Z \subseteq H$ be a closed subspace. The projection operator $\Pi_{H\to Z} : H \to Z$ is defined for every $v \in H$ by*

$$\Pi_{H\to Z} v \in Z \text{ and } (\Pi_{H\to Z} v - v) \in Z^\perp. \tag{2.9}$$

Where no confusion may occur we will simply denote the projection operator by Π_Z. See fig. 2.2.

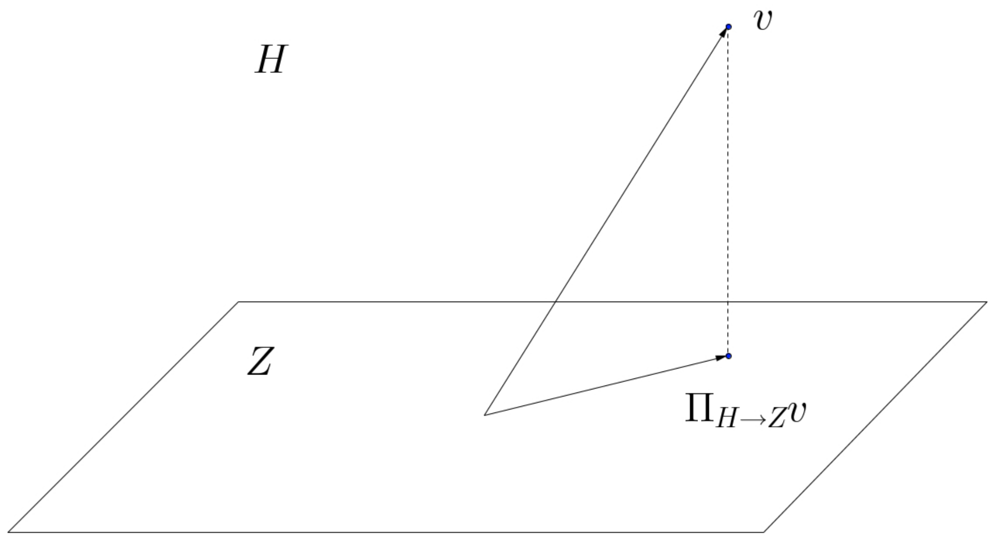


Figure 2.2: An example of a projection operator $\Pi_{H\to Z}$.

The next lemma now establishes the relation between B and B_h.

Lemma 2.1. *The relation between B and B_h can be written as,*

$$B_h v_h = \Pi_{Q'_h} B E_{V_h} v_h \quad \forall v_h \in V_h. \tag{2.10}$$

We end this section by defining conforming methods and consistency.

Definition 2.5 (Conforming method)**.** *A finite element method is called conforming if $V_h \subset V$ and $Q_h \subset Q$.*

Definition 2.6 (Consistency)**.** *If the solution $(u,p) \in V \times Q$ of the continuous problem eq. (2.1) satisfies the discrete problem eq. (2.5), i.e.,*

$$a(u, v_h) + b(v_h, p) = \langle f, v_h \rangle_{V'_h \times V_h} \quad \forall v_h \in V_h \subset V, \tag{2.11a}$$

$$b(u, q_h) = \langle g, q_h \rangle_{Q'_h \times Q_h} \quad \forall q_h \in Q_h \subset Q, \tag{2.11b}$$

then the discrete problem is said to be consistent.

2.2 Equivalent Statements of the Properties of Operators

Before we state theorems for well-posedness of the saddle point problem in section 2.3, we give equivalent statements of surjectivivity and bijectivity of operators. We consider the general setting of Banach operators, i.e., continuous/bounded linear operators between Banach spaces, see [8, Appendix A]. This framework exploits fundamental theorems in Banach theory which we discuss next.

Definition 2.7. *Let U and V be two normed vector spaces. $\mathcal{L}(U;V)$ is the vector space of continuous/bounded linear operators from U to V. A norm equipped on $\mathcal{L}(U;V)$ is given by:*

$$\|l\|_{\mathcal{L}(U;V)} := \sup_{u\in U} \frac{\|l(u)\|_V}{\|u\|_U}.$$

Therefore, $\mathcal{L}(U;V)$ is a normed vector space.

The next proposition gives a condition for $\mathcal{L}(U;V)$ to be complete, i.e., for $\mathcal{L}(U;V)$ to be a Banach space.

Proposition 2.1. *Let U be a normed vector space and let V be a Banach space. Then, $\mathcal{L}(U;V)$ is a Banach space.*

Remark 2.2. *$\mathcal{L}(U;V)$ is a normed vector space*

Remark 2.3. *$\mathcal{L}(U;V) = U'$ when $V = \mathbb{R}$, thus the dual space of a Banach space is a Banach space. The dual space of a dual space is also a Banach space.*

Theorem 2.1 (Hahn–Banach Theorem (simplified version))**.** *Let V be a normed vector space and let G be a subspace of V equipped with the same norm. Let $B_G \in G' = \mathcal{L}(G;\mathbb{R})$ be a linear continuous mapping with norm*

$$\|B_G\|_{G'} = \sup_{g\in G} \frac{\langle B_G, g\rangle_{G'\times G}}{\|g\|_V}.$$

Then there exists $B_V \in V'$ with the following properties:

(i) B_V *is an extension of* B_G*, i.e.,* $B_V g = B_G g$ *for all* $g \in G$.

(ii) $\|B_V\|_{V'} = \|B_G\|_{G'}$.

Definition 2.8 (Dual operator)**.** *Let U and V be two normed vector spaces and let $A \in \mathcal{L}(U;V)$. The dual operator $A^T : V' \to U'$ is defined by*

$$\forall u \in U,\ \forall v' \in V', \quad \langle A^T v', u\rangle_{U'\times U} = \langle v', Au\rangle_{V'\times V}.$$

Theorem 2.2 (Closed Range Theorem)**.** *Let V and W be real Banach spaces. Let $A \in \mathcal{L}(V; W)$. The following statements are equivalent:*

(i) $Im(A)$ *is closed.*

(ii) $Im(A^T)$ *is closed.*

(iii) $Im(A) = (Ker(A^T))^{\perp}$.

(iv) $Im(A^T) = (Ker(A))^{\perp}$.

Theorem 2.3 (Open Mapping Theorem)**.** *Let V and W be real Banach spaces. Let $A \in \mathcal{L}(V; W)$. If A is surjective and U is an open set in V, then $A(U)$ is open in W.*

It may be difficult to check that $Im(A)$ is closed in theorem 2.2. We therefore list equivalent statements, which are a consequence of the open mapping theorem 2.3, in the following lemma.

Lemma 2.2 (Operators with closed range)**.** *[8, Lemma A.36] Let $M \in \mathcal{L}(D; R)$. The following statements are equivalent:*

(i) Im(M) is closed.

(ii) There exists $\beta >$ 0 such that

$$\forall r \in Im(M),\ \exists d \in D,\ Md = r \text{ and } \beta \|d\|_D \leq \|r\|_R .$$

(iii) $\exists L_M \in \mathcal{L}(Im(M), D)$ *and* $\alpha > 0$ *such that:*

$$\forall r \in Im(M),\ L_M r = d,\ M L_M r = r \text{ and } L_M \text{ is bounded as } \|L_M r\|_D \leq \beta^{-1} \|r\|_R .$$

Remark 2.4. *Statement (iii) in lemma 2.2 is not given in [8, Lemma A.36]. Instead, it is a rephrasing of statement (ii): If M is surjective, then $Im(M) = R$ and so the domain of L_M (**a lifting operator**) is R. Hence L_M is the **right inverse** of M. See [4, Eq. (4.2.23)] and [15, Def. 6.6].*

Remark 2.5. *Statement lemma 2.2(ii) can be rewritten as*

$$\beta \|d\|_D \leq \|Md\|_R ,\ \forall d \in D, \tag{2.12}$$

provided that M is also injective. Equation (2.12) is commonly called the bounded below property ('bounding' in [4]) of the operator M.

Remark 2.6. *To summarize remark 2.4 and remark 2.5, if we enhance a bounded linear operator M with surjectivity, then M has a bounded right inverse; if we enhance it with injectivity, then M is bounded below.*

Lemma 2.3 (Surjective operators)**.** *([8, Lemma A.39] adding statement about right inverse.) Let $M \in \mathcal{L}(D; R)$. The following statements are equivalent:*

(i) $M^T : R' \to D'$ *is surjective.*

(ii) $M : D \to R$ *is injective and* $Im(M)$ *is closed in* R.

(iii) The right inverse of M^T *exists, i.e., the lifting operator* L_{M^T}*, such that* $\|L_{M^T} d'\|_{R'} \leq \beta \|d'\|_{D'}$ *with* $\beta > 0$ *a constant.*

(iv) M *is bounded below: there exists* $\beta > 0$ *such that*

$$\forall d \in D, \ \|Md\|_R \geq \beta \|d\|_D \, .$$

(v) There exists $\beta > 0$ *such that*

$$\inf_{d \in D} \sup_{r' \in R'} \frac{\langle r', Md \rangle_{R' \times R}}{\|r'\|_{R'} \|d\|_D} \geq \beta. \tag{2.13}$$

Proof. (i) ⇔ (ii) is a consequence of the closed range theorem 2.2. (i) ⇔ (iii) is discussed already in remark 2.4. So is (iii) ⇔ (iv) in remark 2.5. (iv) ⇔ (v) is a consequence of the Hahn-Banach theorem 2.1. (See [8, Corollary A.17]). □

Alternatively, exchanging M and M^T,

Lemma 2.4 (Surjective operators)**.** *([8, Lemma A.40] adding statement about right inverse) Let $M \in \mathcal{L}(D; R)$. The following statements are equivalent:*

(i) $M : D \to R$ *is surjective.*

(ii) $M^T : R' \to D'$ *is injective and* $Im(M^T)$ *is closed in* R.

(iii) The right inverse of M *exists, i.e., the lifting operator* L_M*, such that* $\|L_M d\|_R \leq \beta \|d\|_D$ *with* $\beta > 0$ *a constant.*

(iv) M^T *is bounded below: there exists* $\beta > 0$ *such that*

$$\forall r' \in R', \ \|M^T r'\|_D' \geq \beta \, \|r'\|_{R'} \, .$$

(v) There exists $\beta > 0$ such that

$$\inf_{r' \in R'} \sup_{d \in D} \frac{\langle M^T r', d \rangle_{D' \times D}}{\|r'\|_{R'} \|d\|_D} \geq \beta. \tag{2.14}$$

Remark 2.7. *Equation (2.14) is the famous **inf-sup condition** of M.*

Lemma 2.5 (Bijective operators)**.** *Let $M \in \mathcal{L}(D; R)$. The following statements are equivalent:*

(i) M is bijective (i.e. isomorphism).

(ii) There exists $\beta > 0$ such that

$$\begin{aligned} \inf_{d \in D} \sup_{r' \in R'} \frac{\langle r', M d \rangle_{R' \times R}}{\|r'\|_{R'} \|d\|_D} &\geq \beta, \\ \inf_{r' \in R'} \sup_{d \in D} \frac{\langle M^T r', d \rangle_{D' \times D}}{\|r'\|_{R'} \|d\|_D} &\geq \beta. \end{aligned} \tag{2.15}$$

Remark 2.8. *Equation (2.15) can be called the **double inf-sup condition** of M (and M^T). It is probably the most compact statement equivalent to bijectivity.*

Given all the above definitions, lemmas and remarks, we now consider again the saddle point problem eq. (2.1) and eq. (2.4).

Consider a bilinear form induced operator B, i.e.,

$$\langle Bv, q \rangle_{Q',Q} := b(v, q) \quad \forall v \in V, \ \forall q \in Q. \tag{2.16}$$

The **dual operator** and the **transposed operator** of $B : V \to Q'$ are defined as, respectively,

$$B^T : Q'' \to V' \qquad \langle v, B^T q'' \rangle_{V,V'} := \langle Bv, q'' \rangle_{Q',Q''} \qquad \forall v \in V, \ \forall q'' \in Q''. \qquad \text{(dual operator)}$$

$$B^t : Q \to V' \qquad \langle v, B^t q \rangle_{V,V'} := b(v, q) \qquad \forall v \in V, \ \forall q \in Q. \qquad \text{(transposed operator)}$$

We want to specify conditions for Banach spaces under which these two operators are identical.

Lemma 2.6 (Isometry with the double dual)**.** *[8, Proposition A.24] Let Q be a Banach space and let $J_Q : Q \to Q''$ be the linear mapping defined by*

$$\forall q \in Q, \ \forall q' \in Q', \quad \langle J_Q q, q' \rangle_{Q'' \times Q'} = \langle q', q \rangle_{Q' \times Q}.$$

Then, J_Q is an isometry.

Definition 2.9 (Reflexive Banach spaces)**.** *Let Q be a Banach space. Q is said to be reflexive if J_Q is an isomorphism.*

Theorem 2.4 (Riesz Representation Theorem)**.** *Let Q be a Hilbert space. For each $q' \in Q'$, there exists a unique $\mathcal{R}q' = q \in Q$, with $\mathcal{R} : Q' \mapsto Q$ the Riesz map, such that*

$$\forall w \in Q, \quad \langle q', w \rangle_{Q' \times Q} = (q, w)_Q.$$

Moreover, the Riesz map $\mathcal{R}$ is an isometric isomorphism.

Remark 2.9. *The Riesz map $\mathcal{R} : Q' \mapsto Q$ introduces an inner product (Hilbertian structure) to the dual space by*

$$(p', q')_{Q'} = (\mathcal{R}p', \mathcal{R}q')_Q.$$

It can easily be seen that this inner product is consistent with the dual norm imposed on the dual space Q' that makes Q' a Banach space. Equipped with this inner product, Q' becomes a Hilbert space. By the same arguments, we can also equip the double dual of Q with Hilbertian structure.

Another interesting fact about the Riesz map is that, in this Hilbertian setting, the isometry with the double dual J_Q can be easily shown to be exactly the composite of Riesz maps of Q and Q'. And since both Riesz maps are isomorphic, we deduce the following lemma:

Lemma 2.7. *Hilbert spaces are reflexive.*

Proof. For $\forall q \in Q,\ \forall q' \in Q'$, we have,

$$\langle (\mathcal{R}')^{-1} \circ \mathcal{R}^{-1} q, q' \rangle_{Q'' \times Q'} = (\mathcal{R}^{-1} q, q')_{Q'} = (q, \mathcal{R}q')_Q = \langle q, q' \rangle_{Q \times Q'}.$$

Therefore, $J_Q = (\mathcal{R}')^{-1} \circ \mathcal{R}^{-1}$, and so J_Q is an isomorphism. □

The following diagram shows the hierarchy of the spaces we consider.

$$\text{Normed vector space} \overset{\text{completion}}{\Longrightarrow} \text{Banach space} \overset{\text{reflexivity}}{\Longrightarrow} \text{Reflexive Banach space} \overset{\text{inner product}}{\Longrightarrow} \text{Hilbert space}$$

Provided that the domain Banach space is reflexive we may replace the dual operator by the transpose operator for our bilinear form induced operators. Furthermore, in the Hilbertian setting we have reflexivity.

2.3 Well-posedness

We are now ready to state the following theorem which gives *necessary* and *sufficient* conditions for the well-posedness of our problem, see also [4, Theorem 4.2.2].

Theorem 2.5 (Well-posedness)**.** *Problem eq. (2.4) has a unique solution if and only if:*

(i) A is an isomorphism (bijective) from $Ker(B)$ to $(Ker(B))'$.

(ii) B is surjective from V to Q'.

Furthermore, the solution is stable, i.e., the solution is bounded by the data:

$$\begin{aligned} \|u\|_V &\le C_1\|f\|_{V'} + C_2\|g\|_{Q'}\,, \\ \|p\|_Q &\le C_3\|f\|_{V'} + C_4\|g\|_{Q'}\,, \end{aligned} \tag{2.17}$$

where C_1, C_2, C_3 and C_4 are positive constants.

Remark 2.10. *The constants in eq. (2.17) vary in size according to different conditions imposed on $a(\cdot,\cdot)$ such as: ellipticity on the whole space or ellipticity just on $Ker(B)$; symmetry of $a(\cdot,\cdot)$, etc.. However, the size of the constants do not play an important role in this research paper.*

The two conditions in theorem 2.5 may be difficult to verify. The following two lemmas give alternative statements that may be easier to verify.

Lemma 2.8 (B is surjective)**.** *Let B be defined as in eq. (2.2). The transposed operator, or equally, the dual operator of B is defined as in eq. (2.3). The following statements are equivalent:*

(i) $B : V \to Q'$ is surjective.

(ii) $B^t : Q \to V'$ is injective and $Im(B^t)$ is closed in V'.

(iii) The right inverse of B, i.e., the lifting operator L_B, exists and is such that $\|L_B\| \le \frac{1}{\beta}$, where $\beta > 0$ is a constant.

(iv) B^t is bounded below, i.e., there exists a constant $\beta > 0$ such that

$$\forall q \in Q,\ \ \|B^t q\|_{V'} \ge \beta \|q\|_Q\,.$$

(v) There exists a constant $\beta > 0$ such that

$$\inf_{q\in Q}\sup_{v\in V}\frac{\langle B^t q, v\rangle_{V'\times V}}{\|q\|_Q\|v\|_V} \ge \beta \tag{2.18}$$

or, equivalently,

$$\inf_{q \in Q} \sup_{v \in V} \frac{b(v,q)}{\|q\|_Q \|v\|_V} \geq \beta. \tag{2.19}$$

*Equation (2.19) is the famous **inf-sup condition** of B.*

Lemma 2.9 (A is bijective on $Ker(B)$)**.** *Let $M \in \mathcal{L}(D; R)$. The following statements are equivalent:*

(i) A is bijective on $Ker(B)$, i.e., A is an isomorphism.

(ii) There exists a constant $\alpha > 0$ such that

$$\begin{aligned} \inf_{v \in Ker(B)} \sup_{w \in Ker(B)} \frac{a(v,w)}{\|v\|_V \|w\|_V} \geq \alpha, \\ \inf_{w \in Ker(B)} \sup_{v \in Ker(B)} \frac{a(v,w)}{\|v\|_V \|w\|_V} \geq \alpha. \end{aligned} \tag{2.20}$$

The difference between eq. (2.4) and its discrete approximation eq. (2.6) is that the spaces in eq. (2.6) are finite dimensional approximations of those in eq. (2.4). The well-posedness for the discrete problem eq. (2.6) is therefore a special case of theorem 2.5 as discussed next.

Theorem 2.6 (Discrete Well-posedness)**.** *Equation (2.6) has a unique solution if and only if:*

(i) A_h is an isomorphism (bijective) from $Ker(B_h)$ to $(Ker(B_h))'$.

(ii) B_h is surjective from V_h to Q'_h.

Furthermore, the solution is stable, i.e., the solution is bounded by the data:

$$\begin{aligned} \|u_h\|_V \leq C_5 \|f\|_{V'} + C_6 \|g\|_{Q'}, \\ \|p_h\|_Q \leq C_7 \|f\|_{V'} + C_8 \|g\|_{Q'}, \end{aligned} \tag{2.21}$$

where C_5, C_6, C_7 and C_8 are positive constants.

Unfortunately, A being bijective from $Ker(B)$ to $(Ker(B))'$ does not imply A_h being bijective from $Ker(B_h)$ to $(Ker(B_h))'$. Similarly, B being surjective from V to Q' does not imply that B_h is surjective from V_h to Q'_h. This is unlike the 'standard' case where well-posedness of eq. (1.1d) follows from the well-posedness of eq. (1.1c) by the Lax–Milgram theorem, e.g., [5, Cor. 2.7.13].

We note, however, that $Ker(B_h) \subseteq Ker(B)$ is sufficient to guarantee the bijectivity of A_h. Unfortunately, we often do not have this kernel inclusion (as we will see in section 2.5). Fortunately, $a(\cdot,\cdot)$ being coercive on $V \times V$ implies that A and A_h are bijective on $Ker(B)$ and

$Ker(B_h)$ respectively, as we show next. If $a(\cdot,\cdot)$ is coercive on $V \times V$, there exists $\alpha > 0$ such that

$$a(u,u) \geq \alpha \|u\|_V^2 \,, \ \forall u \in V.$$

Since $Ker(B) \subset V$, we have

$$a(u,u) \geq \alpha \|u\|_V^2 \,, \ \forall u \in Ker(B).$$

Equivalently,

$$\frac{a(u,u)}{\|u\|_V} \geq \alpha \|u\|_V \,, \ \forall u \in Ker(B).$$

This implies that

$$\sup_{v \in Ker(B)} \frac{a(u,v)}{\|v\|_V} \geq \alpha \|u\|_V \,, \ \forall u \in Ker(B) \quad \textbf{and} \quad \sup_{v \in Ker(B)} \frac{a(v,u)}{\|v\|_V} \geq \alpha \|u\|_V \,, \ \forall u \in Ker(B),$$

and so

$$\inf_{u \in Ker(B)} \sup_{v \in Ker(B)} \frac{a(u,v)}{\|v\|_V} \geq \alpha \quad \textbf{and} \quad \inf_{u \in Ker(B)} \sup_{v \in Ker(B)} \frac{a(v,u)}{\|v\|_V} \geq \alpha.$$

From lemma 2.9(ii) it therefore follows that A is bijective on $Ker(B)$. Similarly, A_h is bijective on $Ker(B_h)$ since $Ker(B_h) \subset V_h \subset V$.

The following proposition [4, Remark 5.1.4] presents an equivalent statement between the surjectivity of B_h (theorem 2.6(ii)), assuming B is surjective, and the kernel inclusion $Ker(B_h^t) \subseteq Ker(B^t)$. This proposition will play an important role in section 2.5.

Proposition 2.2. *Given B is surjective, the following two statements are equivalent:*

(i) $KerB_h^t \subseteq KerB^t$.

(ii) B_h *is surjective.*

Proof. We prove first $(i) \Rightarrow (ii)$. From lemma 2.8(ii) we know that B^t is injective, i.e, $Ker(B^t) = \{0\}$. Since $Ker(B_h^t) \subseteq Ker(B^t)$ we immediately deduce that $Ker(B_h^t) = \{0\}$, i.e., B_h is surjective. We next prove $(ii) \Rightarrow (i)$. We prove the contrapositive. If $Ker(B_h^t) \nsubseteq Ker(B^t)$, then $Ker(B_h^t)$ cannot be $\{0\}$, and so B_h^t cannot be injective. As a result, B_h cannot be surjective (see lemma 2.8). □

2.4 Basic Error Estimate

Let $(u,p) \in V \times Q$ be the solution to eq. (2.1) and let $(u_h, p_h) \in V_h \times Q_h$ be the solution to eq. (2.5). In this section we determine bounds for the errors $\|u - u_h\|_V$ and $\|p - p_h\|_Q$. For this, let $(u_I, p_I) \in V_h \times Q_h$. By the triangle inequality it is clear that

$$\begin{aligned} \|u - u_h\|_V &\leq \|u - u_I\|_V + \|u_h - u_I\|_V \,, \\ \|p - p_h\|_Q &\leq \|p - p_I\|_Q + \|p_h - p_I\|_Q \,. \end{aligned} \tag{2.22}$$

We will show that

$$\begin{aligned} \|u_h - u_I\|_V &\lesssim \|u - u_I\|_V + \|p - p_I\|_Q \,, \\ \|p_h - p_I\|_Q &\lesssim \|u - u_I\|_V + \|p - p_I\|_Q \,, \end{aligned} \tag{2.23}$$

where we used the notation $\|\cdot\| \lesssim \|\cdot\|$ to denote $\|\cdot\| \leq C\|\cdot\|$, where $C > 0$ a constant independent of the mesh size h. Combining eq. (2.22) and eq. (2.23) then results in the main result of this section:

Theorem 2.7 (Basic Error Estimate). *Let $(u, p) \in V \times Q$ be the solution to eq. (2.1) and let $(u_h, p_h) \in V_h \times Q_h$ be the solution to eq. (2.5). We find that*

$$\begin{aligned} \|u - u_h\|_V &\lesssim \inf_{u_I \in V_h} \|u - u_I\|_V + \inf_{p_I \in Q_h} \|p - p_I\|_Q \,, \\ \|p - p_h\|_Q &\lesssim \inf_{u_I \in V_h} \|u - u_I\|_V + \inf_{p_I \in Q_h} \|p - p_I\|_Q \,. \end{aligned} \tag{2.24}$$

Proof. To prove this result we therefore need to prove eq. (2.23). To start, by consistency eq. (2.11) of the conforming case,

$$a(u_h, v_h) + b(v_h, p_h) = a(u, v_h) + b(v_h, p) \qquad \forall v_h \in V_h, \tag{2.25a}$$

$$b(u_h, q_h) = b(u, q_h) \qquad \forall q_h \in Q_h. \tag{2.25b}$$

Subtracting u_I and p_I from both hand sides of eq. (2.25) we obtain:

$$a(u_h - u_I, v_h) + b(v_h, p_h - p_I) = a(u - u_I, v_h) + b(v_h, p - p_I) \qquad \forall v_h \in V_h, \tag{2.26a}$$

$$b(u_h - u_I, q_h) = b(u - u_I, q_h) \qquad \forall q_h \in Q_h. \tag{2.26b}$$

The right hand sides can be treated as linear forms on V_h and Q_h.

$$\langle \mathcal{F}, v_h \rangle := a(u - u_I, v_h) + b(v_h, p - p_I) \qquad \forall v_h \in V_h, \tag{2.27a}$$

$$\langle \mathcal{G}, q_h \rangle := b(u - u_I, q_h) \qquad \forall q_h \in Q_h. \tag{2.27b}$$

Hence,

$$a(u_h - u_I, v_h) + b(v_h, p_h - p_I) = \langle \mathcal{F}, v_h \rangle \qquad \forall v_h \in V_h, \tag{2.28a}$$

$$b(u_h - u_I, q_h) = \langle \mathcal{G}, q_h \rangle \qquad \forall q_h \in Q_h. \tag{2.28b}$$

Note that $(\tilde{u}_h, \tilde{p}_h) := (u_h - u_I, p_h - p_I) \in V_h \times Q_h$ is the solution to:

$$a(\tilde{u}_h, v_h) + b(v_h, \tilde{p}_h) = \langle \mathcal{F}, v_h \rangle \qquad \forall v_h \in V_h, \tag{2.29a}$$
$$b(\tilde{u}_h, q_h) = \langle \mathcal{G}, q_h \rangle \qquad \forall q_h \in Q_h. \tag{2.29b}$$

Assuming well-posedness of the discrete problem, we have the following stability property eq. (2.17):

$$\begin{aligned} \|u_h - u_I\|_V &\lesssim \|\mathcal{F}\|_{V_h'} + \|\mathcal{G}\|_{Q_h'}, \\ \|p_h - p_I\|_Q &\lesssim \|\mathcal{F}\|_{V_h'} + \|\mathcal{G}\|_{Q_h'}. \end{aligned} \tag{2.30}$$

We now need to bound $\|\mathcal{F}\|_{V_h'}$ and $\|\mathcal{G}\|_{Q_h'}$. We may bound these as follows:

$$\begin{aligned} \|\mathcal{F}\|_{V_h'} &\leq \sup_{v_h \in V_h} \frac{a(u - u_I, v_h)}{\|v_h\|_V} + \sup_{v_h \in V_h} \frac{b(v_h, p - p_I)}{\|v_h\|_V}, \\ \|\mathcal{G}\|_{Q_h'} &\leq \sup_{q_h \in Q_h} \frac{b(u - u_I, q_h)}{\|q_h\|_Q}. \end{aligned} \tag{2.31}$$

By boundedness of $a(\cdot, \cdot)$ and $b(\cdot, \cdot)$ it follows that

$$\begin{aligned} \|\mathcal{F}\|_{V_h'} &\lesssim \|u - u_I\|_V + \|p - p_I\|_V, \\ \|\mathcal{G}\|_{Q_h'} &\lesssim \|u - u_I\|_V. \end{aligned} \tag{2.32}$$

Equation (2.30)–eq. (2.32) now imply eq. (2.23). □

2.5 Q-robust Error Estimate

Let $(u, p) \in V \times Q$ be the solution to eq. (2.1) and let $(u_h, p_h) \in V_h \times Q_h$ be the solution to eq. (2.5). Under certain conditions on the function spaces it is possible to improve the error estimate for $\|u - u_h\|_V$ given in theorem 2.7. Determining these improved bounds is the topic of this section.

We define the affine manifolds $(Z_h(Bu), Z_h^*(B^t p))$ as

$$\begin{aligned} Z_h(g) &:= \left\{ v_h \in V_h,\ b(v_h, q_h) = \langle g, q_h \rangle_{Q_h' \times Q_h},\ \forall q_h \in Q_h \right\}, \\ Z_h^*(f) &:= \left\{ q_h \in Q_h,\ b(v_h, q_h) = \langle f, v_h \rangle_{V_h' \times V_h},\ \forall v_h \in V_h \right\}, \end{aligned}$$

and

$$Z_h(Bu) := \left\{ v_h \in V_h,\ b(v_h, q_h) = \langle Bu, q_h \rangle_{Q_h' \times Q_h} = b(u, q_h),\ \forall q_h \in Q_h \right\},$$
$$Z_h^*(B^t p) := \left\{ q_h \in Q_h,\ b(v_h, q_h) = \langle B^t p, v_h \rangle_{V_h' \times V_h} = b(v_h, p),\ \forall v_h \in V_h \right\}.$$

By definition of $\mathcal{F}$ and $\mathcal{G}$ (see eq. (2.26) and eq. (2.28)), we note that

$$u_I \in Z_h(Bu) \Rightarrow \qquad b(u - u_I, q_h) = 0,\ \forall q_h \in Q_h \Rightarrow \qquad \|\mathcal{G}\|_{Q_h'} = 0,$$
$$p_I \in Z_h^*(B^t p) \Rightarrow \qquad b(v_h, p - p_I) = 0,\ \forall v_h \in V_h \Rightarrow \qquad \|\mathcal{F}\|_{V_h'} \lesssim \|u - u_I\|_V .$$

We also require the following proposition.

Proposition 2.3. *[4, Prop. 5.1.3] Suppose that the affine manifold $Z_h(Bu)$ is nonempty, then for any $u \in V$,*

$$\inf_{w_h \in Z_h(Bu)} \|u - w_h\|_V \lesssim \inf_{v_h \in V_h} \|u - v_h\|_V .$$

The Q-robust error estimate is given by the following theorem.

Theorem 2.8 (Conforming Q-robustness)**.** *Suppose that both the affine manifolds $Z_h(Bu)$ and $Z_h^*(B^t p)$ are nonempty. Then,*

$$\|u - u_h\|_V \lesssim \inf_{v_h \in V_h} \|u - v_h\|_V .$$

Proof. As in section 2.4, let $(u_I, p_I) \in V_h \times Q_h$. By the triangle inequality,

$$\|u - u_h\|_V \leq \|u - u_I\|_V + \|u_h - u_I\|_V . \tag{2.33}$$

If we choose $u_I \in Z_h(Bu)$ and $p_I \in Z_h^*(B^t p)$ then

$$\|u_h - u_I\|_V \lesssim \|u - u_I\|_V . \tag{2.34}$$

Combined with eq. (2.33),

$$\|u - u_h\|_V \lesssim \|u - u_I\|_V , \tag{2.35}$$

or, equivalently,

$$\|u - u_h\|_V \lesssim \inf_{u_I \in Z_h(Bu)} \|u - u_I\|_V . \tag{2.36}$$

The result follows using proposition 2.3. ☐

Remark 2.11. *Note that the estimate in theorem 2.8 is* independent *of $\|q - q_h\|_Q$, hence the term 'Q-robustness'. This is an improved bound for the error $\|u - u_h\|_V$ compared to that given in eq. (2.24).*

The Q-robustness result of theorem 2.8 hinges on the nonemptiness of the affine manifold pair $(Z_h(Bu), Z_h^*(B^t p))$. It may sometimes be difficult to verify this property. Fortunately, the following proposition gives sufficient conditions, see [4, Proposition 4.1.6; 5.1.1; 5.1.2].

Proposition 2.4. *The following statements are equivalent:*

(i) There exists a linear mapping $\Pi_h : V \to V_h$ *such that:*

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

(ii) $Ker(B_h^t) \subseteq Ker(B^t)$*.*

From this proposition we note that $\Pi_h v \in Z_h(Bu)$, hence $Z_h(Bu)$ is not empty.

Alternatively, exchanging B and B^t,

Proposition 2.5. *The following statements are equivalent:*

(i) There exists a linear mapping $\Phi_h : Q \to Q_h$ *such that:*

$$b(v_h, q - \Phi_h q) = 0, \ \forall q \in Q, \ \forall v_h \in V_h.$$

(ii) $Ker(B_h) \subseteq Ker(B)$*.*

From this proposition we note that $\Phi_h q \in Z_h^*(B^t p)$, hence $Z_h^*(B^t p)$ is not empty.

We end this section with the following remarks. In the general case we found in section 2.3 the following.

Remark 2.12. *Proposition 2.2 tells us that the kernel inclusion* $Ker(B_h^t) \subseteq Ker(B^t)$ *is equivalent to* B_h *being surjective. Then by lemma 2.8, this is equivalent to the discrete inf-sup condition (assuming the continuous inf-sup condition holds).*

For Q-robustness, the following is important.

Remark 2.13. *For theorem 2.8 to hold we require that* $Z_h(Bu)$ *and* $Z_h^*(B^t p)$ *are nonempty.*

- *Proposition 2.4 gives a sufficient condition for* $Z_h(Bu)$ *to be nonempty, namely* $Ker(B_h^t) \subseteq Ker(B^t)$*. This condition is equivalent to the discrete inf-sup condition (see remark 2.12).*
- *A sufficient condition for* $Z_h^*(B^t p)$ *to be nonempty is given proposition 2.5, namely* $Ker(B_h) \subseteq Ker(B)$*. This condition is* not *possessed by most finite element pairs and explains why most finite element pairs are not Q-robust.*

Note that we have only discussed sufficient conditions for $Z_h(Bu)$ and $Z_h^*(B^t p)$ to be nonempty. There may be finite element methods that do not satisfy these sufficient conditions but that are Q-robust.

Chapter 3

Nonconforming FEM for Saddle Point Problems

The results presented in chapter 2 are known from the literature. This chapter presents a generalization of these results to nonconforming finite element methods and forms the main part of new contributions of this research paper. In particular, we will obtain error estimates for nonconforming finite element methods by generalizing the approach in chapter 2.

3.1 The Nonconforming Discrete Problem

Given two finite dimensional discrete spaces V_h (with norm $|||\cdot|||_{V_h}$) and Q_h (with norm $|||\cdot|||_{Q_h}$). Let $a_h(\cdot,\cdot) : V_h \times V_h \to \mathbb{R}$ and $b_h(\cdot,\cdot) : V_h \times Q_h \to \mathbb{R}$ be bilinear forms that approximate $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ as given in eq. (2.1). We consider the following discrete problem: Given $f \in V_h'$ and $g \in Q_h'$, we want to find $(u_h, p_h) \in V_h \times Q_h$ that solves:

$$a_h(u_h, v_h) + b_h(v_h, p_h) = \langle f, v_h \rangle, \quad \forall v_h \in V_h, \tag{3.1a}$$

$$b_h(u_h, q_h) = \langle g, q_h \rangle, \quad \forall q_h \in Q_h. \tag{3.1b}$$

In the nonconforming setting at least one of the following two inclusions does not hold: $V_h \subset V$ or $Q_h \subset Q$.

3.2 Well-posedness

For the well-posedness of problem eq. (3.1), we first note that bilinear forms $a_h(\cdot,\cdot)$ and $b_h(\cdot,\cdot)$ are bounded on $V_h \times V_h$ and $V_h \times Q_h$ due to their finite dimensionality. We denote by $A_h^{nc} : V_h \to V_h'$

the operator associated with the bilinear form a_h and $B_h^{nc} : V_h \to Q_h'$ the operator associated with the bilinear form b_h. These are defined as:

$$\langle A_h^{nc} u_h, v_h \rangle_{V_h', V_h} := a_h(u_h, v_h) \quad \forall u_h \in V_h,\ \forall v_h \in V_h,$$
$$\langle B_h^{nc} v_h, q_h \rangle_{Q_h', Q_h} := b_h(v_h, q_h) \quad \forall v_h \in V_h,\ \forall q_h \in Q_h.$$

The kernel of B_h^{nc} is defined as

$$Ker(B_h^{nc}) := \left\{ v_h \in V_h, b_h(v_h, q_h) = 0,\ \forall q_h \in Q_h \right\}.$$

We now state the well-posedness result for nonconforming finite element methods.

Theorem 3.1 (Well-posedness nonconforming FEM)**.** *Problem eq. (3.1) has a unique solution if and only if:*

(i) A_h^{nc} is an isomorphism (bijective) from $Ker(B_h^{nc})$ to $(Ker(B_h^{nc}))'$.

(ii) B_h^{nc} is surjective from V_h to Q_h'.

Furthermore, the solution is stable, i.e., the solution is bounded by the data:

$$\begin{aligned} |||u_h|||_{V_h} &\leq C_1 |||f|||_{V_h'} + C_2 |||g|||_{Q_h'}, \\ |||p_h|||_{Q_h} &\leq C_3 |||f|||_{V_h'} + C_4 |||g|||_{Q_h'}. \end{aligned} \tag{3.2}$$

Proof. The proof is identical to the proof of well-posedness for the conforming case, see [4, Theorem 4.2.2]. □

3.3 Basic Error Estimate

In this research paper we assume that the exact solution $(u, p) \in V \times Q$ to eq. (2.1) also satisfies the discrete problem eq. (3.1). However, $V_h \not\subset V$ and/or $Q_h \not\subset Q$. This presents a problem since the bilinear forms $a_h(\cdot,\cdot)$ and $b_h(\cdot,\cdot)$ are not defined on, respectively, $V \times V$ and $V \times Q$. We therefore assume (see for example [15, Section 1.3.3]) that there are subspaces $V_* \subset V$ and $Q_* \subset Q$ to which the exact solutions belong, and such that the bilinear forms $a_h(\cdot,\cdot)$ and $b_h(\cdot,\cdot)$ may be extended to, respectively, $V_{*h} \times V_{*h}$ and $V_{*h} \times Q_{*h}$, where (see fig. 3.1)

$$V_{*h} = V_h + V_*, \qquad \text{and} \qquad Q_{*h} = Q_h + Q_*.$$

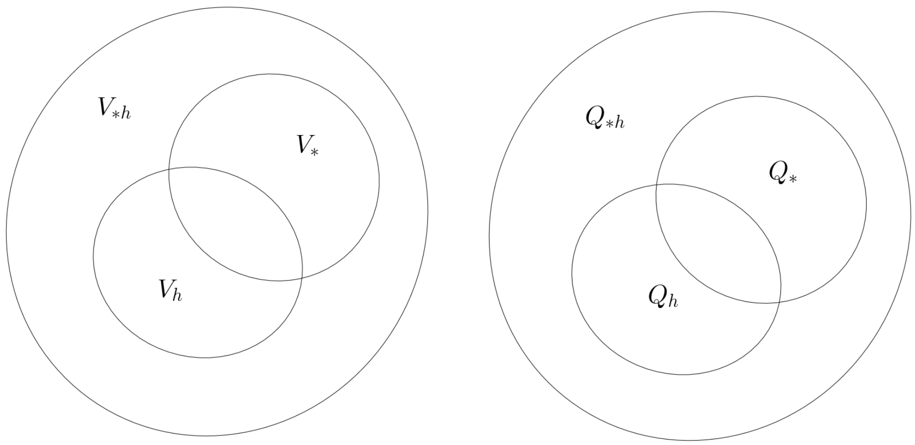


Figure 3.1: Depiction of an example of V_*, V_h and V_{*h} (left) and Q_*, Q_h and Q_{*h} (right).

We also define the operators $A^{nc} : V_{*h} \to V'_{*h}$ and $B^{nc} : V_{*h} \to Q'_{*h}$ by

$$\langle A^{nc}u, v\rangle_{V'_{*h},V_{*h}} := a_h(u, v) \quad \forall u \in V_{*h},\ \forall v \in V_{*h},$$
$$\langle B^{nc}v, q\rangle_{Q'_{*h},Q_{*h}} := b_h(v, q) \quad \forall v \in V_{*h},\ \forall q \in Q_{*h}.$$

The kernel of B^{nc} is defined as

$$Ker(B^{nc}) := \left\{v \in V_{*h},\ b_h(v, q) = 0,\ \forall q \in Q_{*h}\right\}. \tag{3.3}$$

Consistency is now defined as follows.

Definition 3.1 (Consistency)**.** *We say that the discrete problem eq.* (3.1) *is* consistent *if for the exact solution pair* $(u, p) \in V_* \times Q_*$*,*

$$a_h(u, v_h) + b_h(v_h, p) = \langle f, v_h\rangle, \quad \forall v_h \in V_h, \tag{3.4a}$$
$$b_h(u, q_h) = \langle g, q_h\rangle, \quad \forall q_h \in Q_h. \tag{3.4b}$$

It is often not possible to show boundedness of $a_h(\cdot, \cdot)$ and $b_h(\cdot, \cdot)$ on $V_{*h} \times V_{*h}$ and $V_{*h} \times Q_{*h}$ in the norms $|||\cdot|||_{V_h}$ and $|||\cdot|||_{Q_h}$. For this reason we assume norms $|||\cdot|||_{V_*}$ and $|||\cdot|||_{Q_*}$ for which it is possible to show boundedness of $a_h(\cdot, \cdot)$ on $V_{*h} \times V_h$, $b_h(\cdot, \cdot)$ on $V_{*h} \times Q_h$, and $b_h(\cdot, \cdot)$ on $V_h \times Q_{*h}$, see [15, Section 1.3.4]. The norm $|||\cdot|||_{V_*}$ defined on V_{*h} is assumed to be such that for all $v_h \in V_h$, $|||v_h|||_{V_h} \le |||v_h|||_{V_*} \le C|||v_h|||_{V_h}$ with C independent of h. Similarly, $|||\cdot|||_{Q_*}$ is defined on Q_{*h} such that for all $q_h \in Q_h$, $|||q_h|||_{Q_h} \le |||q_h|||_{Q_*} \le C|||q_h|||_{Q_h}$ with C independent of h. We summarize these results in the following two definitions.

Definition 3.2 (Boundedness of a_h)**.** *We assume that the bilinear form $a_h(\cdot,\cdot)$ is bounded on $V_{*h} \times V_{*h}$ if there exists a $C^a_{bnd} > 0$, independent of h, such that,*

$$\left|a_h(v, w_h)\right| \leq C^a_{bnd} |||v|||_{V_*} |||w_h|||_{V_h} \quad \forall (v, w_h) \in V_{*h} \times V_h.$$

Definition 3.3 (Boundedness of b_h)**.** *We assume that the bilinear form $b_h(\cdot,\cdot)$ is bounded in $V_{*h} \times Q_h$ if there exists a $C^{b'}_{bnd} > 0$, independent of h, such that,*

$$\left|b_h(v, q_h)\right| \leq C^{b'}_{bnd} |||v|||_{V_*} |||q_h|||_{Q_h} \quad \forall (v, q_h) \in V_{*h} \times Q_h.$$

We assume that the bilinear form $b_h(\cdot,\cdot)$ is bounded in $V_h \times Q_{*h}$ if there exists a $C^{b''}_{bnd} > 0$, independent of h, such that,

$$\left|b_h(v_h, q)\right| \leq C^{b''}_{bnd} |||v_h|||_{V_h} |||q|||_{Q_*} \quad \forall (v_h, q) \in V_h \times Q_{*h}.$$

We may now prove an error estimate for nonconforming FEM.

Theorem 3.2 (Nonconforming Basic Error Estimate)**.** *Let $(u,p) \in V_* \times Q_*$ be the unique solution of problem eq. (2.1). Let $(u_h, p_h) \in V_h \times Q_h$ solve problem eq. (3.1). Assume consistency and boundedness of the bilinear forms $a_h(\cdot,\cdot)$ and $b_h(\cdot,\cdot)$. Then there exists $C_{err} > 0$, independent of h, such that*

$$|||(u - u_h, p - p_h)|||_{V_h,Q_h} \leq C_{err} \inf_{(v_h,q_h)\in V_h\times Q_h} |||(u - v_h, p - q_h)|||_{V_*,Q_*}, \tag{3.5}$$

where $|||(v,q)|||^2_{V_h,Q_h} := |||v|||^2_{V_h} + |||q|||^2_{Q_h}$ for all $(v,q) \in V_{*h} \times Q_{*h}$.

Proof. By the virtue of consistency,

$$a_h(u_h - u_I, v_h) + b_h(v_h, p_h - p_I) = a_h(u - u_I, v_h) + b_h(v_h, p - p_I) \qquad \forall v_h \in V_h, \tag{3.6a}$$
$$b_h(u_h - u_I, q_h) = b_h(u - u_I, q_h) \qquad \forall q_h \in Q_h, \tag{3.6b}$$

which is similar to eq. (2.26) for conforming FEM. Likewise, considering the right hand sides in eq. (3.6) as linear forms $\mathcal{F}^{nc}$ and $\mathcal{G}^{nc}$ acting on V_h and Q_h,

$$a_h(u_h - u_I, v_h) + b_h(v_h, p_h - p_I) = \langle \mathcal{F}^{nc}, v_h\rangle_h \quad \forall v_h \in V_h, \tag{3.7a}$$
$$b_h(u_h - u_I, q_h) = \langle \mathcal{G}^{nc}, q_h\rangle_h \quad \forall q_h \in Q_h, \tag{3.7b}$$

which is similar to eq. (2.28) for conforming FEM. By the stability property eq. (3.2) we observe that

$$|||u_h - u_I|||_{V_h} \leq C_1 |||\mathcal{F}^{nc}||| + C_2 |||\mathcal{G}^{nc}|||. \tag{3.8}$$

We may bound $|||\mathcal{F}^{nc}|||$ and $|||\mathcal{G}^{nc}|||$ as follows:

$$|||\mathcal{F}^{nc}||| \leq \sup_{v_h \in V_h} \frac{a_h(u - u_I, v_h)}{|||v_h|||_{V_h}} + \sup_{v_h \in V_h} \frac{b_h(v_h, p - p_I)}{|||v_h|||_{V_h}}, \tag{3.9a}$$

$$|||\mathcal{G}^{nc}||| \leq \sup_{v_h \in V_h} \frac{b_h(u - u_I, q_h)}{|||q_h|||_{Q_h}}. \tag{3.9b}$$

By boundedness of a_h and b_h eq. (3.9) becomes

$$|||\mathcal{F}^{nc}||| \lesssim |||u - u_I|||_{V_*} + |||p - p_I|||_{Q_*}, \tag{3.10a}$$

$$|||\mathcal{G}^{nc}||| \lesssim |||u - u_I|||_{V_*}. \tag{3.10b}$$

Combined with eq. (3.8) we then obtain

$$|||u_h - u_I|||_{V_h} \lesssim |||u - u_I|||_{V_*} + |||p - p_I|||_{Q_*}. \tag{3.11}$$

By the triangle inequality,

$$|||u - u_h|||_{V_h} \lesssim \inf_{u_I \in V_h} |||u - u_I|||_{V_*} + \inf_{p_I \in Q_h} |||p - p_I|||_{Q_*},$$

$$|||p - p_h|||_{Q_h} \lesssim \inf_{u_I \in V_h} |||u - u_I|||_{V_*} + \inf_{p_I \in Q_h} |||p - p_I|||_{V_*},$$

which may be combined to eq. (3.5). □

3.4 Q-robust Error Estimate

As we saw previously in section 2.5, it was possible to obtain improved error estimates (Q-robust error estimates) under certain conditions on the function spaces. In this section we generalize these results to the nonconforming setting and improve the error estimates given by theorem 3.2.

We define the affine manifolds $(Z_h(B^{nc}u), Z_h^*(B^{nc^t}p))$ as

$$Z_h(B^{nc}u) := \left\{ v_h \in V_h,\ b_h(u - v_h, q_h) = 0,\ \forall q_h \in Q_h \right\},$$

$$Z_h^*(B^{nc^t}p) := \left\{ q_h \in Q_h,\ b_h(v_h, p - q_h) = 0,\ \forall v_h \in V_h \right\},$$

We require also the following propostion.

Proposition 3.1. *For any $u \in V_{*h}$, there exists C, independent of h, such that*

$$\inf_{w_h \in Z_h(B^{nc}u)} |||u - w_h|||_{V_*} \leq C \inf_{v_h \in V_h} |||u - v_h|||_{V_*}$$

Proof. We show that for arbitrary $v_h \in V_h$ we can always find $w_h \in Z_h(B^{nc}u)$ such that

$$|||u - w_h|||_{V_*} \leq C|||u - v_h|||_{V_*}.$$

We first note that the inf-sup condition

$$\beta_*^{nc}|||q_h|||_{Q_h} \leq \sup_{v_h \in V_h} \frac{b_h(v_h, q_h)}{|||v_h|||_{V_*}}, \ \forall q_h \in Q_h, \tag{3.12}$$

follows directly from

$$\beta^{nc}|||q_h|||_{Q_h} \leq \sup_{v_h \in V_h} \frac{b_h(v_h, q_h)}{|||v_h|||_{V_h}}, \ \forall q_h \in Q_h, \tag{3.13}$$

since $|||v_h|||_{V_*}$ and $|||v_h|||_{V_h}$ are equivalent for $v_h \in V_h$.

Let $L_{B_h^{nc}}$ be the right inverse of B_h^{nc} associated with the inf-sup condition eq. (3.12). (Recall the equivalence between B_h^{nc} being surjective and the fact that the right inverse of B_h^{nc} exists and that is bounded with constant $1/\beta_*^{nc}$, see lemma 2.8.) We then set for all $v_h \in V_h$,

$$d_h := L_{B_h^{nc}}(\Pi_{Q'_h} B^{nc}(u - v_h)),$$

where $\Pi_{Q'_h} : Q'_{*h} \to Q'_h$ is the projection operator as defined in definition 2.4. We then find

$$\begin{aligned} B_h^{nc}(d_h + v_h) &= \Pi_{Q'_h} B^{nc}(u - v_h) + B_h^{nc} v_h \\ &= \Pi_{Q'_h} B^{nc} u. \end{aligned}$$

This shows that $w_h := d_h + v_h \in Z_h(B^{nc}u)$ (which follows by definition of B_h^{nc}, B^{nc} and $\Pi_{Q'_h}$). We also find that

$$|||d_h|||_{V_*} \leq \frac{1}{\beta_*^{nc}} |||\Pi_{Q'_h} B^{nc}(u - v_h)|||_{Q'_h} \leq \frac{C_{bnd}^{b'}}{\beta_*^{nc}} |||u - v_h|||_{V_*},$$

where the first inequality is by boundedness of $L_{B_h^{nc}}$ and the second inequality is by

$$\langle \Pi_{Q'_h} B^{nc}(u - v_h), q_h \rangle_{Q'_h \times Q_h} = b_h(u - v_h, q_h),$$

and boundedness of b_h.

The result then follows by a triangle inequality:

$$|||u - w_h|||_{V_*} \leq |||u - v_h|||_{V_*} + |||d_h|||_{V_*} \lesssim |||u - v_h|||_{V_*}.$$

□

The Q-robust error estimate is given by the following theorem.

Theorem 3.3 (Nonconforming Q-robust Error Estimate)**.** *Let* $(u,p) \in V_* \times Q_*$ *be the unique solution of problem eq.* (2.1)*. Let* $(u_h, p_h) \in V_h \times Q_h$ *solve problem eq.* (3.1)*. Assume the affine manifolds* $Z_h(B^{nc}u)$ *and* $Z_h^*(B^{nc^t}p)$ *are nonempty. Then,*

$$|||u - u_h|||_{V_h} \lesssim \inf_{v_h \in V_h} |||u - v_h|||_{V_*}.$$

Proof. As in section 2.5, for $u_I \in Z_h(B^{nc}u)$, $|||\mathcal{G}^{nc}||| = 0$, and for $p_I \in Z_h^*(B^{nc^t}p)$,

$$|||u_h - u_I|||_{V_h} \lesssim |||\mathcal{F}^{nc}||| \lesssim \sup_{v_h \in V_h} \frac{a_h(u - u_I, v_h)}{|||v_h|||_{V_h}} \lesssim |||u - u_I|||_{V_*}, \tag{3.14}$$

where the last inequality is by boundedness of $a_h(\cdot,\cdot)$ as defined in definition 3.2. By the triangle inequality,

$$|||u - u_h|||_{V_h} \leq |||u_h - u_I|||_{V_h} + |||u - u_I|||_{V_h} \lesssim |||u - u_I|||_{V_*}.$$

Thus,

$$|||u - u_h|||_{V_h} \lesssim \inf_{u_I \in Z_h(B^{nc}u)} |||u - u_I|||_{V_*}. \tag{3.15}$$

The result follows by proposition 3.1. □

The Q-robustness result of theorem 3.3 hinges on the nonemptiness of $Z_h(B^{nc}u)$ and $Z_h^*(B^{nc^t}p)$. However, the nonemptiness of $Z_h(B^{nc}u)$ is equivalent to the surjectivity of B_h^{nc}, i.e., the discrete inf-sup conditon of $b_h(\cdot,\cdot)$, which is always required for well-posedness. Hence, to obtain Q-robustness requires to show nonemptiness of $Z_h^*(B^{nc^t}p)$.

Recall that kernel inclusion $Ker(B_h) \subseteq Ker(B)$ is a sufficient condition for $Z_h^*(B^t p)$ to be nonempty in the conforming case (see remark 2.13). For the nonconforming case we have a similar condition as given in the next proposition.

Proposition 3.2. *The following statements are equivalent:*

(i) $Ker(B_h^{nc}) \subseteq Ker(B^{nc})$.

(ii) $\Pi_{V_h'} Im(B^{nc^t}) \subseteq Im(B_h^{nc^t})$.

(iii) There exists a linear operator $\Phi_h : Q_{*h} \to Q_h$ *such that*

$$b(v_h, q - \Phi_h q) = 0 \quad \forall q \in Q_{*h}, \ \forall v_h \in V_h. \tag{3.16}$$

Proof. For a proof of $(i) \Leftrightarrow (ii)$ see [4, Prop. 4.1.6]. We prove $(ii) \Leftrightarrow (iii)$.

$(iii) \Rightarrow (ii)$ $\forall q \in Q_{*h}$, eq. (3.16) can be rewritten as

$$\Pi_{V_h'} B^{nc^t} q = \Pi_{V_h'} B^{nc^t} E_{Q_h} \Phi_h q = B_h^{nc^t} \Phi_h q \in Im(B_h^{nc^t}).$$

$(ii) \Rightarrow (iii)$ $\forall q \in Q_{*h}$,

$$\Pi_{V_h'} B^{nc^t} q \in Im(B_h^{nc^t}),$$

Let $L_{B_h^{nc^t}}$ be the right inverse of $B_h^{nc^t}$. Define now

$$\Phi_h q := L_{B_h^{nc^t}}(\Pi_{V_h'} B^{nc^t} q).$$

It follows that

$$B_h^{nc^t} \Phi_h q = \Pi_{V_h'} B^{nc^t} q,$$

which is the same as eq. (3.16). □

Chapter 4

The Stokes Problem

In this chapter we apply the theory of chapter 2 and chapter 3 to the Stokes problem. We first briefly introduce this problem.

The incompressible Navier–Stokes equations are the fundamental equations of fluid dynamics. The steady Navier–Stokes equations are given by:

$$-\nu\nabla^2\boldsymbol{u} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla p = \boldsymbol{f}, \tag{4.1a}$$

$$\nabla\cdot\boldsymbol{u} = 0, \tag{4.1b}$$

$$\boldsymbol{u}|_{\partial\Omega} = 0, \tag{4.1c}$$

$$\langle p\rangle_\Omega = 0. \tag{4.1d}$$

Here eq. (4.1a) is the equation of momentum conservation while eq. (4.1b) is the equation for mass conservation. Furthermore, we consider only homogeneous Dirichlet boundary conditions eq. (4.1c). Since the pressure is unique up to a constant, we impose that the mean of the pressure over the whole domain is zero by eq. (4.1d).

In the low Reynolds number limit the Navier–Stokes equations simplify to the Stokes equations:

$$-\nu\nabla^2\boldsymbol{u} + \nabla p = \boldsymbol{f}, \tag{4.2a}$$

$$-\nabla\cdot\boldsymbol{u} = g, \tag{4.2b}$$

$$\boldsymbol{u}|_{\partial\Omega} = 0, \tag{4.2c}$$

$$\langle p\rangle_\Omega = 0. \tag{4.2d}$$

We remark that g in eq. (4.2b) originates from transforming inhomogeneous Dirichlet boundary conditions to no-slip boundary conditions [11]. And $\int_\Omega g\,\mathrm{d}\boldsymbol{x} = 0$ because the divergence theorem implies $\int_\Omega \nabla \cdot \boldsymbol{u}\, \mathrm{d}\boldsymbol{x} = \int_{\partial\Omega} \boldsymbol{u} \cdot \boldsymbol{n}\, \mathrm{d}\boldsymbol{s} = 0$. We also remark that there are variations on boundary conditions besides eqs. (4.2c) to (4.2d) (see [8, Section 4.1.4] and [7, Chapter 3]) and the rest of the analysis in this chapter can be altered accordingly for a treatment to general boundary conditions.

4.1 Variational Formulation

Let $\Omega \subset \mathbb{R}^d$ be a polygonal ($d = 2$) or polyhedral ($d = 3$) domain and let $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ and $g \in L^2(\Omega)$. Let $L^2_0(\Omega)$ be the space of square integrable functions with vanishing mean on Ω and let $\boldsymbol{H}^1_0(\Omega) := \left[H^1_0(\Omega)\right]^d$. The variational formulation of the Stokes problem eq. (4.2) is given by: find $(\boldsymbol{u}, p) \in \boldsymbol{V} \times Q := \boldsymbol{H}^1_0(\Omega) \times L^2_0(\Omega)$ such that

$$a(\boldsymbol{u}, \boldsymbol{v}) + b(\boldsymbol{v}, p) = (\boldsymbol{f}, \boldsymbol{v}), \quad \forall \boldsymbol{v} \in \boldsymbol{V}, \tag{4.3a}$$

$$b(\boldsymbol{u}, q) = (g, q), \quad \forall q \in Q, \tag{4.3b}$$

where

$$a(\boldsymbol{u}, \boldsymbol{v}) := \nu(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}) := \nu \int_\Omega \nabla \boldsymbol{u} : \nabla \boldsymbol{v}\, \mathrm{d}\boldsymbol{x},$$

$$b(\boldsymbol{v}, q) := -(\nabla \cdot \boldsymbol{v}, q) = -\int_\Omega q \nabla \cdot \boldsymbol{v}\, \mathrm{d}\boldsymbol{x},$$

where $(\cdot, \cdot)$ denotes the L^2 inner product over Ω. These bilinear forms are clearly bounded:

$$a(\boldsymbol{u}, \boldsymbol{v}) = \nu(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}) \le \nu \|\nabla \boldsymbol{u}\|_{\boldsymbol{L}^2(\Omega)} \|\nabla \boldsymbol{v}\|_{\boldsymbol{L}^2(\Omega)} \le \nu \|\boldsymbol{u}\|_{\boldsymbol{H}^1(\Omega)} \|\boldsymbol{v}\|_{\boldsymbol{H}^1(\Omega)},$$

$$b(\boldsymbol{v}, q) = -(\nabla \cdot \boldsymbol{v}, q) \le \|\nabla \cdot \boldsymbol{v}\|_{\boldsymbol{L}^2(\Omega)} \|q\|_{L^2(\Omega)} \le \|\nabla \boldsymbol{v}\|_{\boldsymbol{L}^2(\Omega)} \|q\|_{L^2(\Omega)} \le \|\boldsymbol{v}\|_{\boldsymbol{H}^1(\Omega)} \|q\|_{L^2(\Omega)}.$$

The operator B associated with the bilinear form b is given by

$$B = \nabla \cdot \ : \boldsymbol{H}^1_0(\Omega) \to L^2_0(\Omega), \tag{4.4}$$

and its kernel is given by

$$Ker(B) = \left\{ \boldsymbol{v} \in \boldsymbol{H}^1_0(\Omega) : B\boldsymbol{v} = \nabla \cdot \boldsymbol{v} = 0 \right\}. \tag{4.5}$$

Recall theorem 2.5. For problem eq. (4.2) to be well-posed, we require

- $a(\cdot, \cdot)$ satisfies the double inf-sup condition on $Ker(B)$ (see lemma 2.9).
- $b(\cdot, \cdot)$ satisfies the inf-sup condition (see lemma 2.8).

For the first condition, it suffices that $a(\cdot,\cdot)$ is coercive on $\boldsymbol{H}_0^1(\Omega)$. To show this we require Poincaré's inequality, as defined next.

Lemma 4.1 (Poincaré inequality)**.** *[9, Theorem 1 and remark, Section 5.7] Let $1 \leq p \leq \infty$ and let Ω be a bounded open set. Then there exists constant $C_{p,\Omega} > 0$ such that*

$$\forall v \in W_0^{1,p}(\Omega), \quad C_{p,\Omega}\|v\|_{L^p(\Omega)} \leq \|\nabla v\|_{L^p(\Omega)} . \tag{4.6}$$

To show now that $a(\cdot,\cdot)$ is coercive on $\boldsymbol{H}_0^1(\Omega)$, note that

$$a(\boldsymbol{u},\boldsymbol{u}) = \nu(\nabla\boldsymbol{u},\nabla\boldsymbol{u}) = \nu\|\nabla\boldsymbol{u}\|_{\boldsymbol{L}^2(\Omega)} \geq \alpha\|\boldsymbol{u}\|_{\boldsymbol{H}^1(\Omega)} , \quad \forall\boldsymbol{u} \in \boldsymbol{H}_0^1(\Omega), \tag{4.7}$$

where we used Poincaré's inequality for the last step. Furthermore, $\alpha = \nu C$ where C is a constant that depends on the Poincaré constant.

To show that $b(\cdot,\cdot)$ satisfies the inf-sup condition (equivalently, the surjectivity of the divergence operator eq. (4.4)), we require the following definition.

Definition 4.1 (Star-shaped w.r.t. a ball)**.** *An open bounded set Ω is said to be star-shaped w.r.t. a ball B if for any $x \in \Omega$ and $z \in B \subset \Omega$, the segment joining x and z is contained in Ω.*

Using this definition, it was shown [10, Lemma 3.1] that the divergence operator is surjective:

Lemma 4.2 (Bogovskiĭ)**.** *Let Ω is star-shaped w.r.t. a ball $B \subset \Omega$. Then the divergence operator $B = \nabla\cdot\ : \boldsymbol{H}_0^1(\Omega) \to L_0^2(\Omega)$ is surjective.*

We remark that if $p = 2$ in lemma 4.1 then lemma 4.2 holds also for a domain Ω of which the boundary is Lipschitz (instead of Ω being star-shaped), see [8, Remark B.70]. This is the case for the Stokes problem.

Denoting the inf-sup constant by $\beta > 0$, the inf-sup condition for the divergence constraint eq. (4.3b) is given by:

$$\inf_{q\in L_0^2(\Omega)} \sup_{\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega)} \frac{(\nabla\cdot\boldsymbol{v},q)}{\|\boldsymbol{v}\|_{\boldsymbol{H}^1(\Omega)}\|q\|_{L^2(\Omega)}} \geq \beta. \tag{4.8}$$

We may now state well-posedness of the Stokes problem.

Theorem 4.1 (Well-posedness of the Stokes problem)**.** *The Stokes problem (4.2) is well-posed. Furthermore, the solution satisfies the following bounds: $\forall\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ and $g \in L_0^2(\Omega)$,*

$$\|\boldsymbol{u}\|_{1,\Omega} \leq \frac{1}{\alpha}\|\boldsymbol{f}\|_{-1,\Omega} + \frac{2\|a\|^{1/2}}{\alpha^{1/2}\beta}\|g\|_{0,\Omega} , \tag{4.9a}$$

$$\|p\|_{0,\Omega} \leq \frac{2\|a\|^{1/2}}{\alpha^{1/2}\beta}\|\boldsymbol{f}\|_{-1,\Omega} + \frac{2\|a\|^{1/2}}{\alpha^{1/2}\beta}\|g\|_{0,\Omega} , \tag{4.9b}$$

where $\|a\| = \nu$ *and* $\alpha = C\nu$ *(see eq. (4.7)).*

4.2 Standard Mixed FEM Methods

Let $\boldsymbol{V}_h \times Q_h \subset \boldsymbol{V} \times Q$ denote a pair of stable conforming piecewise polynomial spaces w.r.t. a mesh $\mathcal{T}_h$ of Ω. A Galerkin mixed finite element method for the Stokes equations seeks $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h \times Q_h$ such that,

$$a(\boldsymbol{u}_h, \boldsymbol{v}_h) + b(\boldsymbol{v}_h, p_h) = (\boldsymbol{f}, \boldsymbol{v}_h), \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{4.10a}$$

$$b(\boldsymbol{u}_h, q_h) = (g, q_h), \quad \forall q_h \in Q_h. \tag{4.10b}$$

Using the theory introduced in section 2.4 we immediately find the following error estimate for the velocity:

$$\begin{aligned} \|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{H}^1(\Omega)} &\leq \left(\frac{2\|a\|}{\alpha} + \frac{2\|a\|^{1/2}}{\alpha^{1/2}\beta}\right)\|\boldsymbol{u} - \boldsymbol{u}_I\|_{\boldsymbol{H}^1(\Omega)} + \frac{\|b\|}{\alpha}\|p - p_I\|_{L^2(\Omega)}\,, \\ &= \underbrace{\frac{2}{\tilde{C}_\Omega}\left(\frac{1}{\tilde{C}_\Omega} + \frac{1}{\beta}\right)}_{C_1}\|\boldsymbol{u} - \boldsymbol{u}_I\|_{\boldsymbol{H}^1(\Omega)} + \underbrace{\frac{1}{\tilde{C}_\Omega^2}}_{C_2}\nu^{-1}\|p - p_I\|_{L^2(\Omega)}\,, \end{aligned}$$

where we used that $\|b\| = 1$. We therefore find that

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_{\boldsymbol{H}^1(\Omega)} \leq C_1 \inf_{\boldsymbol{v}_h \in \boldsymbol{V}_h} \|\boldsymbol{u} - \boldsymbol{v}_h\|_{\boldsymbol{H}^1(\Omega)} + C_2\nu^{-1} \inf_{q_h \in Q_h} \|p - q_h\|_{L^2(\Omega)}\,, \tag{4.11}$$

where the constants C_1, C_2 depend only on the domain Ω. This estimate immediately shows the dependence of the velocity error estimate on the best approximation error of the pressure scaled by the inverse of the viscosity, see example 1.1. However, as we saw in section 2.5, if we use a stable conforming finite element pair with the kernel inclusion property proposition 2.5(ii), this result may be improved. We, however, do not pursue this here. Instead, in the next section we consider a non-conforming finite element method.

4.3 $H(\mathbf{div})$-Conforming DG Methods

In this section we consider an $H(\text{div})$-conforming discontinuous Galerkin (DG) finite element method for the Stokes equations eq. (4.2). We consider the interior penalty DG version of the method originally introduced in [6].

The new theory for nonconforming methods developed in chapter 3 will provide guidance in proving a pressure-robust error estimate for the velocity.

4.3.1 The Raviart–Thomas Space

To define the $H(\text{div})$-conforming DG method, we require $H(\text{div})$-conforming finite element spaces. We consider the Raviart–Thomas (RT) space [4, Chapter 2]. However, we start by defining $\boldsymbol{H}(\text{div}, \Omega)$.

Definition 4.2. *The Hilbert space of vector fields that possess a weak divergence (defined in the same way as weak derivatives) is defined by*

$$\boldsymbol{H}(\text{div}, \Omega) := \{\mathrm{w} \in \mathrm{L}^2(\Omega) : \nabla \cdot \mathrm{w} \in \mathrm{L}^2(\Omega)\}. \tag{4.12}$$

Let $\mathcal{T}_h$ denote a shape-regular triangulation of Ω. Let $\mathcal{F}_h$ denote the set of all the edges ($d = 2$) or faces ($d = 3$) of the mesh $\mathcal{T}_h$. The set of boundary edges/faces is denoted by $\mathcal{F}_h^B \subset \mathcal{F}_h$ while $\mathcal{F}_h^I \subset \mathcal{F}_h$ is the set of interior edges/faces. Consider an arbitrary element $K \in \mathcal{T}_h$. We will next define the Raviart–Thomas space on K.

Definition 4.3 (Polynomials)**.** *[19, Section 2] We introduce preliminary notations in order to define and investigate the RT polynomial space on K.*

- $\mathcal{P}_k(K)$ *is the space of polynomials of degree at most k defined on K.*
- $\boldsymbol{\mathcal{P}}_k(K) := \mathcal{P}_k(K)^d$.
- $\tilde{\mathcal{P}}_k(K)$ *are homogeneous polynomials of degree k.*
- $\mathcal{P}_k(e)$ *with $e \in \partial K$ is the space of $(d-1)$-variate polynomials on tangential coordinates.*
- $\mathcal{R}_k(\partial K) = \Pi_{e \in \partial K} \mathcal{P}_k(e)$ *are piecewise polynomial functions on ∂K.*

The Raviart–Thomas space is now defined as:

Definition 4.4 (Raviart–Thomas Space)**.** *The RT space on K is defined as*

$$\mathcal{RT}_k(K) := \boldsymbol{\mathcal{P}}_k(K) \oplus \mathbf{x}\tilde{\mathcal{P}}_k(K). \tag{4.13}$$

The RT space satisfies the following useful property which will play an important role in deriving pressure-robust error estimates in section 4.3.4:

Proposition 4.1. $\mathrm{div}(\mathcal{RT}_\mathrm{k}(\mathrm{K})) = \mathcal{P}_\mathrm{k}(\mathrm{K})$.

For other useful properties of the RT space we refer to [19, Prop. 2.3].

For the RT space to be a proper finite element, the RT space requires the definition of local degrees of freedom and, accordingly, the local interpolation/projection operator.

Definition 4.5 (The Raviart-Thomas Projection)**.** *[19] Let* $\boldsymbol{v} : K \to \mathbb{R}^d$ *be sufficiently smooth. The RT projection* $\boldsymbol{\Pi}_k^{RT}\boldsymbol{v} \in \mathcal{RT}_k(K)$ *is defined by the equations*

$$(\boldsymbol{\Pi}_k^{RT}\boldsymbol{v}, \boldsymbol{w})_K = (\boldsymbol{v}, \boldsymbol{w})_K \quad \forall \boldsymbol{w} \in \boldsymbol{\mathcal{P}}_{k-1}(K), \tag{4.14a}$$

$$\langle \boldsymbol{\Pi}_k^{RT}\boldsymbol{v} \cdot \boldsymbol{n}, \mu \rangle_{\partial K} = \langle \boldsymbol{v} \cdot \boldsymbol{n}, \mu \rangle_{\partial K} \quad \forall \mu \in \mathcal{R}_k(\partial K). \tag{4.14b}$$

Proposition 4.2. *Equation* (4.14) *are uniquely solvable, i.e., the RT space is a unisolvent finite element [19, Prop. 2.4].*

The RT projection $\boldsymbol{\Pi}_k^{RT}$ satisfies the following approximation results.

Proposition 4.3 (Approximation estimates of RT projection)**.** *Let* K *be an affine element. There exists a constant* C *depending only on* k *and the shape of* K *such that, for* $1 \le m \le k+1$, $s = 0$ *or* 1 *and for any* $\boldsymbol{v}$ *in* $\boldsymbol{H}^m(K)$*, we have*

$$\|\boldsymbol{v} - \boldsymbol{\Pi}_k^{RT}\boldsymbol{v}\|_{s,K} \le C h_K^{m-s} |\boldsymbol{v}|_{m,K} \tag{4.15}$$

Proof. This a special case of [4, Prop. 2.5.1]. □

The following commutativity property will play an important role in deriving the inf-sup condition in lemma 4.8. This commutativity property enables the RT projection to be a Fortin operator as we will see in lemma 4.7.

Proposition 4.4 (Commutativity Property)**.** *[19, Eq. (17)] Consider* $\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}, \Omega)$*, the following commutativity property holds:*

$$\nabla \cdot \boldsymbol{\Pi}_k^{RT}\boldsymbol{v} = \Pi_k \nabla \cdot \boldsymbol{v}, \tag{4.16}$$

where Π_k *is the* $L^2(K)$*-projection onto* $\mathcal{P}_k(K)$*, i.e., for all* $u \in L^2(K)$*,*

$$(\Pi_k u, v)_K = (u, v)_K \quad \forall v \in \mathcal{P}_k(K).$$

4.3.2 Weak Formulation

We will now discuss the nonconforming DG weak formulation for the Stokes problem eq. (4.2).

As finite dimensional function spaces we consider

$$\begin{aligned} \boldsymbol{V}_h &:= \mathcal{RT}_k := \{\boldsymbol{v}_h \in \boldsymbol{H}_0(\mathrm{div}, \Omega) : \boldsymbol{v}_h|_K \in \mathcal{RT}_k(K) \ \forall K \in \mathcal{T}_h\}, \\ Q_h &:= \{q_h \in L_0^2(\Omega) : q_h|_K \in \mathcal{P}_k \ \forall K \in \mathcal{T}_h\}, \end{aligned} \tag{4.17}$$

where $\boldsymbol{H}_0(\mathrm{div},\Omega) := \{\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div},\Omega) : \boldsymbol{v}\cdot\boldsymbol{n} = 0 \text{ on } \partial\Omega\}$. We note that $\boldsymbol{V}_h \not\subset \boldsymbol{H}_0^1(\Omega)$, i.e., we are considering a nonconforming finite element for the Stokes problem. However, since $\boldsymbol{V}_h \subset \boldsymbol{H}(\mathrm{div},\Omega)$, this method is called $H(\mathrm{div})$-conforming.

The DG weak formulation using $H(\mathrm{div})$-conforming finite element methods is now given by: Find $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h \times Q_h$ such that

$$a_h(\boldsymbol{u}_h,\boldsymbol{v}_h) + b(\boldsymbol{v}_h,p_h) = (\boldsymbol{f},\boldsymbol{v}_h) \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \tag{4.18a}$$

$$b(\boldsymbol{u}_h,q_h) = (g,q_h) \quad \forall q_h \in Q_h. \tag{4.18b}$$

The bilinear form $a_h(\boldsymbol{u}_h,\boldsymbol{v}_h)$ is the 'standard' interior-penalty DG method for the vector-Laplacian modified for $H(\mathrm{div})$-conforming finite element methods (see appendix A for a heuristic derivation):

$$\begin{aligned} a_h(\boldsymbol{u}_h,\boldsymbol{v}_h) &:= \int_\Omega \nabla_h \boldsymbol{u}_h : \nabla_h \boldsymbol{v}_h \,\mathrm{d}\boldsymbol{x} \\ &- \sum_{F\in\mathcal{F}_h} \left(\int_F \{\!\{\epsilon(\boldsymbol{u}_h)\}\!\} [\![\boldsymbol{v}_h]\!] \,\mathrm{d}\boldsymbol{s} + \int_F \{\!\{\epsilon(\boldsymbol{v}_h)\}\!\} [\![\boldsymbol{u}_h]\!] \,\mathrm{d}\boldsymbol{s} - \frac{\sigma}{h_F}\int_F [\![\boldsymbol{u}_h]\!][\![\boldsymbol{v}_h]\!]\,\mathrm{d}\boldsymbol{s} \right). \end{aligned} \tag{4.19}$$

Here we used the standard notation for the average and jump operators. These are defined on interior facets $F \in \mathcal{F}_h^I$, $F = \partial K_+ \cap \partial K_-$ by

$$\{\!\{\epsilon(\boldsymbol{u})\}\!\}|_F := \frac{1}{2}\Big(\frac{\partial(\boldsymbol{u}_+\cdot\boldsymbol{\tau}_{K_+})}{\partial \boldsymbol{n}_{K_+}} + \frac{\partial(\boldsymbol{u}_-\cdot\boldsymbol{\tau}_{K_-})}{\partial \boldsymbol{n}_{K_-}} \Big), \tag{4.20a}$$

$$[\![\boldsymbol{u}]\!]|_F := \boldsymbol{u}_+\cdot\boldsymbol{\tau}_{K_+} + \boldsymbol{u}_-\cdot\boldsymbol{\tau}_{K_-}, \tag{4.20b}$$

where $\boldsymbol{u}_\pm = \boldsymbol{u}|_{K_\pm}$ and where $\boldsymbol{\tau}_K$ denotes the tangential unit vector of ∂K, obtained by rotating the normal $\boldsymbol{n}_K$ by 90 degrees counterclockwise. For a boundary edge $F \in \mathcal{F}_h^B$ these operators are defined as

$$\{\!\{\epsilon(\boldsymbol{u})\}\!\}|_F := \frac{\partial(\boldsymbol{u}\cdot\boldsymbol{\tau}_K)}{\partial \boldsymbol{n}_K}, \tag{4.20c}$$

$$[\![\boldsymbol{u}]\!]|_F := \boldsymbol{u}\cdot\boldsymbol{\tau}_K. \tag{4.20d}$$

To prove consistency, boundedness and stability of the DG method eq. (4.18), we require the definition of extended function spaces with norms. Let $\boldsymbol{V}_{*h} := \boldsymbol{V}_h + \boldsymbol{V}_* := \boldsymbol{V}_h + (\boldsymbol{H}^s(\Omega)\cap\boldsymbol{H}_0^1(\Omega))$, with $s > \frac{3}{2}$, and let $Q_{*h} := Q_h + Q_* = Q_h + (H^1(\Omega)\cap L_0^2(\Omega))$. We let the two bilinear forms $a_h(\cdot,\cdot)$ and $b(\cdot,\cdot)$ be extended, respectively, on $\boldsymbol{V}_{*h}\times\boldsymbol{V}_{*h}$ and $\boldsymbol{V}_{*h}\times Q_{*h}$. Note that since $Q_{*h} \subseteq L^2(\Omega)$, the domain of the second argument of the bilinear form $b(\cdot,\cdot)$ needs not to be extended. We introduce two norms $|\!|\!|\cdot|\!|\!|$ and $|\!|\!|\cdot|\!|\!|_*$ on $\boldsymbol{V}_{*h}$ [21, eqs. (4.5) and (4.6)] since the $\boldsymbol{H}^1$-norm cannot be used for $\boldsymbol{H}(\mathrm{div}, \Omega)$. These two norms are defined as:

$$|||\boldsymbol{v}|||^2 = |\boldsymbol{v}|_{1,h}^2 + \sum_{F\in\mathcal{F}_h} h_F^{-1} \big\| [\![\boldsymbol{v}]\!] \big\|_F^2 , \tag{4.21a}$$

$$|||\boldsymbol{v}|||_*^2 = |||\boldsymbol{v}|||^2 + \sum_{F\in\mathcal{F}_h} h_F \big\| \{\!\{\epsilon(\boldsymbol{v})\}\!\} \big\|_F^2 , \tag{4.21b}$$

where $|\boldsymbol{v}|_{1,h}^2 = \sum_{K\in\mathcal{T}_h} |\boldsymbol{v}|_{1,K}^2$ and $\|\cdot\|_F$ is the L^2-norm on F. On Q_{*h}, since $Q_{*h} \subseteq L^2(\Omega)$, we use the L^2-norm and it will be denoted as $\|\cdot\|$.

To obtain an error estimate, recall from section 3.3 that we require that for all $\boldsymbol{v}_h \in \boldsymbol{V}_h$, $|||\boldsymbol{v}_h||| \le |||\boldsymbol{v}_h|||_* \le C|||\boldsymbol{v}_h|||$ with C independent of h. We verify that this is true.

Let K be an element with F as an edge. By the continuous trace inequality [15, Lemma 1.49] and Young's inequality we can show the following variant of the trace inequality [21, eq. (4.7)]. (See appendix B for a proof.)

Lemma 4.3 (Variant of Trace Inequality)**.** *There exists a constant C, independent of the mesh size h, such that for any $v \in H^1(K)$*

$$\|v\|_F^2 \le C(h_K^{-1}\|v\|_K^2 + h_K|v|_{1,K}^2). \tag{4.22}$$

We now show the equivalence of the norms $|||\cdot|||$ and $|||\cdot|||_*$ on $\boldsymbol{V}_h$. First of all, the inequality $|||\boldsymbol{v}_h||| \le |||\boldsymbol{v}_h|||_*$ is trivial for all $\boldsymbol{v}_h \in \boldsymbol{V}_h$ so that we only show that $|||\boldsymbol{v}_h|||_* \le C|||\boldsymbol{v}_h|||$. We first note that the second term on the right hand side of eq. (4.21b) can be bounded using eq. (4.22) as:

$$\sum_{F\in\mathcal{F}_h} h_F \big\| \{\!\{\epsilon(\boldsymbol{v}_h)\}\!\} \big\|_F^2 \lesssim \sum_{K\in\mathcal{T}_h} (|\boldsymbol{v}_h|_{1,K}^2 + h_K^2 |\nabla \boldsymbol{v}_h|_{1,K}^2) \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h. \tag{4.23}$$

We now apply an inverse inequality [22, Eq. (3.3)] to the second term on the right hand side to obtain

$$\sum_{F\in\mathcal{F}_h} h_F \big\| \{\!\{\epsilon(\boldsymbol{v}_h)\}\!\} \big\|_F^2 \lesssim \sum_{K\in\mathcal{T}_h} |\boldsymbol{v}_h|_{1,K}^2. \tag{4.24}$$

It follows that indeed

$$|||\boldsymbol{v}_h|||_* \lesssim |||\boldsymbol{v}_h|||, \tag{4.25}$$

so that $|||\cdot|||$ and $|||\cdot|||_*$ on $\boldsymbol{V}_h$.

We are now ready to prove consistency, boundedness of $a_h(\cdot,\cdot)$, coercivity of $a_h(\cdot,\cdot)$, and the inf-sup condition of $b(\cdot,\cdot)$.

Lemma 4.4 (Consistency)**.** *Let $(u, p) \in \boldsymbol{V}_* \times L_0^2(\Omega)$ be the exact solution to the Stokes problem*

(4.3). Then,

$$\begin{cases} a_h(\boldsymbol{u}, \boldsymbol{v}_h) + b(\boldsymbol{v}_h, p) &= (\boldsymbol{f}, \boldsymbol{v}_h), \quad \forall \boldsymbol{v}_h \in \boldsymbol{V}_h, \\ b(\boldsymbol{u}, q_h) &= (g, q_h), \quad \forall q_h \in Q_h, \end{cases} \tag{4.26}$$

Proof. This is shown in appendix A. □

Lemma 4.5 (Boundedness of $a_h(\cdot,\cdot)$)**.** *$a_h(\cdot,\cdot)$ is bounded on $\boldsymbol{V}_{*h}$, i.e.,*

$$|a_h(\boldsymbol{u}, \boldsymbol{v})| \lesssim |||\boldsymbol{u}|||_* |||\boldsymbol{v}|||_* \quad \forall \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{V}_{*h}. \tag{4.27}$$

Proof. This is an immediate consequence of the Cauchy–Schwarz inequality. For all $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{V}_{*h}$ it holds that

$$\begin{aligned} |a_h(\boldsymbol{u}, \boldsymbol{v})| \lesssim \Bigg\{ & |\boldsymbol{u}|_{1,h} |\boldsymbol{v}|_{1,h} + \Bigg(\sum_{F \in \mathcal{F}_h} h_F \big\| \{\!\{ \epsilon(\boldsymbol{u}) \}\!\} \big\|_F^2 \Bigg)^{\frac{1}{2}} \Bigg(\sum_{F \in \mathcal{F}_h} h_F^{-1} \big\| [\![\boldsymbol{v}]\!] \big\|_F^2 \Bigg)^{\frac{1}{2}} \\ & + \Bigg(\sum_{F \in \mathcal{F}_h} h_F \big\| \{\!\{ \epsilon(\boldsymbol{v}) \}\!\} \big\|_F^2 \Bigg)^{\frac{1}{2}} \Bigg(\sum_{F \in \mathcal{F}_h} h_F^{-1} \big\| [\![\boldsymbol{u}]\!] \big\|_F^2 \Bigg)^{\frac{1}{2}} \\ & + \sigma \Bigg(\sum_{F \in \mathcal{F}_h} h_F^{-1} \big\| [\![\boldsymbol{u}]\!] \big\|_F^2 \Bigg)^{\frac{1}{2}} \Bigg(\sum_{F \in \mathcal{F}_h} h_F^{-1} \big\| [\![\boldsymbol{v}]\!] \big\|_F^2 \Bigg)^{\frac{1}{2}} \Bigg\} \\ \lesssim\ & |||\boldsymbol{u}|||_* |||\boldsymbol{v}|||_*. \end{aligned}$$

□

Remark 4.1. *We note that when $\boldsymbol{v}$ is restricted to $\boldsymbol{V}_h$, by eq. (4.25), lemma 4.5 reduces to*

$$|a_h(\boldsymbol{u}, \boldsymbol{v}_h)| \lesssim |||\boldsymbol{u}|||_* |||\boldsymbol{v}_h|||. \tag{4.28}$$

Lemma 4.6 (Coercivity of $a_h(\cdot,\cdot)$)**.** *There exists a constant C_{ell*}, independent of h, such that $\forall \boldsymbol{v}_h \in \boldsymbol{V}_h$ we have*

$$a_h(\boldsymbol{v}_h, \boldsymbol{v}_h) \geq C_{ell*} |||\boldsymbol{v}_h|||_*^2, \tag{4.29}$$

provided that the penalty parameter σ is sufficiently large.

Proof. It follows from the Cauchy–Schwarz inequality, eq. (4.24), and Young's inequality, that there exists a constant C such that

$$\begin{aligned}\left|\sum_{F\in\mathcal{F}_h}\int_F \{\!\{\epsilon(\boldsymbol{u}_h)\}\!\}[\![\boldsymbol{v}_h]\!]\,\mathrm{d}\boldsymbol{s}\right| &\le \Big(\sum_{F\in\mathcal{F}_h} h_F \big\|\{\!\{\epsilon(\boldsymbol{u}_h)\}\!\}\big\|_F^2\Big)^{\frac{1}{2}}\Big(\sum_{F\in\mathcal{F}_h} h_F^{-1}\big\|[\![\boldsymbol{v}_h]\!]\big\|_F^2\Big)^{\frac{1}{2}}\\
&\le C|\boldsymbol{u}_h|_{1,h}\Big(\sum_{F\in\mathcal{F}_h} h_F^{-1}\big\|[\![\boldsymbol{v}_h]\!]\big\|_F^2\Big)^{\frac{1}{2}}\\
&\le \frac{1}{2}|\boldsymbol{u}_h|_{1,h}^2 + \frac{C^2}{2}\sum_{F\in\mathcal{F}_h} h_F^{-1}\big\|[\![\boldsymbol{v}_h]\!]\big\|_F^2 .\end{aligned}$$

Using the above inequality and eq. (4.25), we obtain

$$\begin{aligned}a_h(\boldsymbol{v}_h,\boldsymbol{v}_h) &= |\boldsymbol{v}_h|_{1,h}^2 + \sigma\sum_{F\in\mathcal{F}_h} h_F^{-1}\big\|[\![\boldsymbol{v}_h]\!]\big\|_F^2 - 2\sum_{F\in\mathcal{F}_h}\int_F \{\!\{\epsilon(\boldsymbol{v}_h)\}\!\}[\![\boldsymbol{v}_h]\!]\,\mathrm{d}\boldsymbol{s}\\
&\ge |\boldsymbol{v}_h|_{1,h}^2 + \sigma\sum_{F\in\mathcal{F}_h} h_F^{-1}\big\|[\![\boldsymbol{v}_h]\!]\big\|_F^2 - \frac{1}{2}|\boldsymbol{v}_h|_{1,h}^2 - \frac{C^2}{2}\sum_{F\in\mathcal{F}_h} h_F^{-1}\big\|[\![\boldsymbol{v}_h]\!]\big\|_F^2\\
&= \frac{1}{2}|\boldsymbol{v}_h|_{1,h}^2 + \Big(\sigma - \frac{C^2}{2}\Big)\sum_{F\in\mathcal{F}_h} h_F^{-1}\big\|[\![\boldsymbol{v}_h]\!]\big\|_F^2\\
&\ge C_{ell}|||\boldsymbol{v}_h|||^2\\
&\ge C_{ell*}|||\boldsymbol{v}_h|||_*^2,\end{aligned}$$

where the second inequality holds for sufficiently large σ, i.e., if $\sigma > \frac{C^2}{2}$. □

To prove the discrete inf-sup condition of $b(\cdot,\cdot)$, we first recall Fortin's trick [4, Prop. 5.4.2]:

Definition 4.6 (Fortin Operator)**.** *Fortin's operator* $\boldsymbol{\Pi}_h : \boldsymbol{V} \to \boldsymbol{V}_h$ *is defined as*

$$b(\boldsymbol{v} - \boldsymbol{\Pi}_h\boldsymbol{v}, q_h) = 0 \quad \forall q_h \in Q_h. \tag{4.30}$$

Proposition 4.5 (Fortin's trick)**.** *Assume that the continuous inf-sup condition eq. (4.8) is satisfied. If the Fortin operator is bounded with a constant C_Π, independent of h, i.e.,*

$$|||\boldsymbol{\Pi}_h\boldsymbol{v}||| \le C_\Pi\|\boldsymbol{v}\|_{\boldsymbol{V}} \quad \forall \boldsymbol{v} \in \boldsymbol{V}, \tag{4.31}$$

then

$$\sup_{\boldsymbol{v}_h\in\boldsymbol{V}_h}\frac{b(\boldsymbol{v}_h,q_h)}{|||\boldsymbol{v}_h|||} \ge \frac{\beta}{C_\Pi}\|q_h\| \quad \forall q_h \in Q_h. \tag{4.32}$$

Proof. The proof is as follows:

$$\sup_{\boldsymbol{v}_h \in \boldsymbol{V}_h} \frac{b(\boldsymbol{v}_h, q_h)}{|||\boldsymbol{v}_h|||} \geq \sup_{\boldsymbol{v} \in \boldsymbol{V}} \frac{b(\boldsymbol{\Pi}_h \boldsymbol{v}, q_h)}{|||\boldsymbol{\Pi}_h \boldsymbol{v}|||} = \sup_{\boldsymbol{v} \in \boldsymbol{V}} \frac{b(\boldsymbol{v}, q_h)}{|||\boldsymbol{\Pi}_h \boldsymbol{v}|||} \geq \frac{1}{C_\Pi} \sup_{\boldsymbol{v} \in \boldsymbol{V}} \frac{b(\boldsymbol{v}, q_h)}{\|\boldsymbol{v}\|_{\boldsymbol{V}}} \geq \frac{\beta}{C_\Pi} \|q_h\| \,. \tag{4.33}$$

□

We apply Fortin's trick now to our $H(\mathrm{div})$-conforming DG finite element method. We note first that the Fortin operator in our case is the Raviart–Thomas projection:

Lemma 4.7 (The Raviart–Thomas projection $\boldsymbol{\Pi}_k^{RT}$ is a Fortin Operator)**.** *For all $\boldsymbol{v} \in \boldsymbol{V}_{*h}$, $\boldsymbol{\Pi}_k^{RT} : \boldsymbol{V}_{*h} \to \boldsymbol{V}_h$ is a Fortin operator. Furthermore, the Raviart–Thomas projection $\boldsymbol{\Pi}_k^{RT}$ is bounded.*

Proof. An immediate consequence of the commutativity property of $\boldsymbol{\Pi}_k^{RT}$ proposition 4.4 is that

$$b(\boldsymbol{v} - \boldsymbol{\Pi}_k^{RT} \boldsymbol{v}, q_h) = 0 \quad \forall q_h \in Q_h. \tag{4.34}$$

We are left to show boundedness, i.e., that

$$|||\boldsymbol{\Pi}_k^{RT} \boldsymbol{v}||| \lesssim \|\boldsymbol{v}\|_{\boldsymbol{V}} \quad \forall \boldsymbol{v} \in \boldsymbol{V}. \tag{4.35}$$

By the approximation estimates of the RT projection operator eq. (4.15) and the trace inequality eq. (4.22), we see that

$$\begin{aligned}
|||\boldsymbol{v} - \boldsymbol{\Pi}_k^{RT} \boldsymbol{v}|||^2 &= |\boldsymbol{v} - \boldsymbol{\Pi}_k^{RT} \boldsymbol{v}|_{1,h}^2 + \sum_{F \in \mathcal{F}_h} h_F^{-1} \left\| [\![\boldsymbol{v} - \boldsymbol{\Pi}_k^{RT} \boldsymbol{v}]\!] \right\|_F^2 \\
&\leq C |\boldsymbol{v}|_1^2 + C \sum_{K \in \mathcal{T}_h} \sum_{F \in \partial K} h_F^{-1} \left\| \boldsymbol{v} - \boldsymbol{\Pi}_k^{RT} \boldsymbol{v} \right\|_F^2 \\
&\leq C \|\boldsymbol{v}\|_1^2 + C \sum_{K \in \mathcal{T}_h} (h_K^{-2} \left\| \boldsymbol{v} - \boldsymbol{\Pi}_k^{RT} \boldsymbol{v} \right\|_K^2 + |\boldsymbol{v} - \boldsymbol{\Pi}_k^{RT} \boldsymbol{v}|_{1,K}^2) \\
&\leq C \|\boldsymbol{v}\|_1^2 + C \sum_{K \in \mathcal{T}_h} (h_K^{-2} h_K^2 |\boldsymbol{v}|_{1,K}^2 + |\boldsymbol{v}|_{1,K}^2) \\
&\leq C \|\boldsymbol{v}\|_1^2 \quad \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega).
\end{aligned} \tag{4.36}$$

Furthermore,

$$|||\boldsymbol{v}||| = |\boldsymbol{v}|_1 \leq \|\boldsymbol{v}\|_1 \quad \forall \boldsymbol{v} \in \boldsymbol{V} = \boldsymbol{H}_0^1(\Omega).$$

Equation (4.35) then follows from the triangle inequality. □

We may now prove the discrete inf-sup condition.

Lemma 4.8 (Inf-sup condition of $b(\cdot,\cdot)$)**.** *There exists constant β_* independent of h such that*

$$\sup_{\boldsymbol{v}_h \in \boldsymbol{V}_h} \frac{b(\boldsymbol{v}_h, q_h)}{|||v_h|||_*} \geq \beta_* \|q_h\| \quad \forall q_h \in Q_h. \tag{4.37}$$

Proof. By the equivalence result eq. (4.25) and eq. (4.35),

$$|||\boldsymbol{\Pi}_k^{RT} \boldsymbol{v}|||_* \leq C \|\boldsymbol{v}\|_1 . \tag{4.38}$$

The result follows by proposition 4.5. □

4.3.3 Well-posedness

We now prove well-posedness of the $H(\text{div})$-conforming DG method eq. (4.18).

Recall from theorem 3.1 that we need to fulfill two conditions for our $H(\text{div})$-conforming finite element method to be well-posed. The first condition in theorem 3.1 is satisfied because a_h is coercive on $\boldsymbol{V}_h$ lemma 4.6. The second condition in theorem 3.1 is also satisfied, see lemma 4.8. Note that in both cases we use the equivalence of the norms $|||\cdot|||$ and $|||\cdot|||_*$ on $\boldsymbol{V}_h$. We therefore have the following result.

Theorem 4.2. *The discrete formulation eq. (4.18) has a unique solution $(\boldsymbol{u}_h, p_h) \in \boldsymbol{V}_h \times Q_h$. Furthermore, this solution is stable, i.e., the solution is bounded by the data:*

$$\begin{aligned} |||\boldsymbol{u}_h||| \leq & \quad C_1 |||f|||_{\boldsymbol{V}_h'} + C_2 \|g\|_{Q_h'} , \\ \|p_h\| \leq & \quad C_3 |||f|||_{\boldsymbol{V}_h'} + C_4 \|g\|_{Q_h'} . \end{aligned} \tag{4.39}$$

4.3.4 Pressure-Robust Estimate

We end section 4.3 by proving a pressure-robust velocity error estimate, i.e., we prove

$$|||\boldsymbol{u} - \boldsymbol{u}_h||| \lesssim \inf_{\boldsymbol{v}_h \in \boldsymbol{V}_h} |||\boldsymbol{u} - \boldsymbol{v}_h|||_*. \tag{4.40}$$

We know from theorem 3.3 that eq. (4.40) holds provided that the following two affine manifolds are nonempty:

$$Z_h := \left\{ \boldsymbol{v}_h \in \boldsymbol{V}_h,\ b(\boldsymbol{u} - \boldsymbol{v}_h, q_h) = \int_\Omega \nabla \cdot (\boldsymbol{u} - \boldsymbol{v}_h) q_h \, \mathrm{d}\boldsymbol{x} = 0,\ \forall q_h \in Q_h \right\},$$
$$Z_h^* := \left\{ q_h \in Q_h,\ b(\boldsymbol{v}_h, p - q_h) = \int_\Omega \nabla \cdot \boldsymbol{v}_h (p - q_h) \, \mathrm{d}\boldsymbol{x} = 0,\ \forall \boldsymbol{v}_h \in \boldsymbol{V}_h \right\}.$$

The nonemptiness of Z_h can be easily shown due to the Raviart–Thomas projection operator $\mathbf{\Pi}_k^{RT}$. To see this, note that $\mathbf{\Pi}_k^{RT}\boldsymbol{u} \in Z_h$.

We now show the nonemptiness of Z_h^*. For this we require the following proposition.

Proposition 4.6. $\nabla \cdot \boldsymbol{V}_h \subseteq Q_h$.

Proof. For $\forall \boldsymbol{v}_h \in \boldsymbol{V}_h$, by proposition 4.1, we have $\nabla \cdot \boldsymbol{v}_h|_K \in \mathcal{P}_k$ and so $\nabla \cdot \boldsymbol{v}_h \in L^2(\Omega)$. Moreover,

$$\int_\Omega \nabla \cdot \boldsymbol{v}_h \, \mathrm{d}\boldsymbol{x} = \int_{\partial\Omega} \boldsymbol{v}_h \cdot \boldsymbol{n} \, \mathrm{d}\boldsymbol{s} = 0,$$

since $\boldsymbol{v}_h \in \boldsymbol{H}_0(\mathrm{div}, \Omega)$. We obtain $\nabla \cdot \boldsymbol{v}_h \in L_0^2(\Omega)$. Since $\nabla \cdot \boldsymbol{v}_h|_K \in \mathcal{P}_k$ and $\nabla \cdot \boldsymbol{v}_h \in L_0^2(\Omega)$ the result follows. □

We remark that $\boldsymbol{V}_h$ is a Hilbert space. Furthermore, since $\boldsymbol{V}_h$ is finite dimensional and the divergence operator is a bounded linear operator, $\nabla \cdot \boldsymbol{V}_h$ is a finite dimensional subspace of $L^2(\Omega)$, thus closed. Using that a closed subspace of a Hilbert space is itself a Hilbert space [4, Section 4.1.2] with respect to the same inner product, we note that $\nabla \cdot \boldsymbol{V}_h$ is a Hilbert space. We may now apply the Riesz Representation Theorem, theorem 2.4, to show that we have a unique $w_h \in \nabla \cdot \boldsymbol{V}_h \subseteq Q_h$ such that

$$b(\boldsymbol{v}_h, p - w_h) = \int_\Omega \nabla \cdot \boldsymbol{v}_h (p - w_h) \, \mathrm{d}\boldsymbol{x} = 0, \ \forall \boldsymbol{v}_h \in \boldsymbol{V}_h.$$

Therefore, Z_h^* is nonempty. The pressure robust estimate eq. (4.40) now follows.

Chapter 5

Conclusions

Pressure-robust finite element methods have become very popular in recent years. This is because the velocity error is independent of the pressure error scaled by the viscosity. In chapter 1 we compared a pressure-robust discretization for the Stokes problem against a discretization that is not pressure-robust. We saw a significant improvement in the velocity.

The analysis of pressure-robust methods has also gained much popularity recently. However, the result, i.e., a velocity error estimate that is independent of the pressure error, is not new. Such results may be found for example in the book of Brezzi and Fortin (Mixed and Hybrid Finite Element Methods, 1991). There, an analysis for general saddle-point problems is presented and conditions are presented on conforming mixed finite element spaces under which one may achieve pressure-robustness. This research paper extended these ideas to non-conforming finite element methods for general saddle-point problems. Since in the general case the scalar unknown is not necessarily a pressure, we refer to our analysis as 'Q-robustness'.

We give sufficient conditions for 'Q-robustness' for general nonconforming mixed finite element methods for saddle point problems. These conditions are the nonemptiness of the following affine manifolds:

$$\begin{aligned} Z_h &:= \left\{ v_h \in V_h,\ b_h(u - v_h, q_h) = 0,\ \forall q_h \in Q_h \right\}, \\ Z_h^* &:= \left\{ q_h \in Q_h,\ b_h(v_h, p - q_h) = 0,\ \forall v_h \in V_h \right\}. \end{aligned}$$

In general it may be difficult to show that Z_h and Z_h^* are not empty. For this reason we presented equivalent, but easier to show, conditions under which we obtain Q-robustness. In particular, we showed:

- The discrete inf-sup condition is a necessary condition for the discrete saddle-point problem to be well-posed. We showed that it also guarantees that the first manifold is nonempty.

- A kernel inclusion is sufficient for the nonemptiness of the affine manifold Z_h^*.

We applied our extended theory to an $H(\text{div})$-conforming discontinuous Galerkin method for the Stokes equations. We showed well-posedness of the method and provided a simple proof of Q-robustness. This is the main application of the extended theory in this research paper.

Future Work

In the future we plan to:

- We want to apply our extended analysis framework to an $H(\text{div})$-conforming hybridized discontinuous Galerkin method for the Stokes problem. In this case, since the method introduces Lagrangian multipliers for both velocity and pressure approximations, we need remark that $a_h(\cdot,\cdot) \neq a(\cdot,\cdot)$ and $b_h(\cdot,\cdot) \neq b(\cdot,\cdot)$ (in this research paper we only considered the case $a_h(\cdot,\cdot) \neq a(\cdot,\cdot)$ and $b_h(\cdot,\cdot) = b(\cdot,\cdot)$). Although pressure-robustness of the discretization has already been shown in the literature, we expect that our new analysis will provide a shorter proof. In particular, we will analyze the nonemptiness of the affine manifold Z_h^*,
$$Z_h^* := \left\{ q_h \in Q_h,\ b_h(v_h, p - q_h) = 0,\ \forall v_h \in V_h \right\}.$$

- It can be shown that both $H(\text{div})$-conforming discontinuous Galerkin method and $H(\text{div})$-conforming hybridized discontinuous Galerkin method are nice methods in the sense that the velocity approximation is $H(\text{div})$-conforming and pointwise divergence free. It can be further shown that these two nice properties will guarantee the nonemptiness of the affine manifold Z_h^*.
  A very nice result of the analysis we did is that our analysis provides sufficient conditions under which Q-robustness may be achieved. These may be weaker than requiring the velocity approximation to be $H(\text{div})$-conforming and pointwise divergence free. This suggests that it might be possible to design new Q-robust finite element methods (or modify existing finite element methods to be Q-robust) that do not necessarily result in a velocity approximation that is $H(\text{div})$-conforming and/or pointwise divergence free.

- Instead of imposing nonemptiness of Z_h and Z_h^* on the finite element method, it may be possible to weaken the conditions for Q-robustness even further. We believe this might be achieved by considering the Helmholtz decomposition and closed range theorem. This may again result in new Q-robust finite element methods that do not necessarily result in a velocity approximation that is $H(\text{div})$-conforming and/or pointwise divergence free.

References

[1] D. J. Acheson. *Elementary Fluid Dynamics*. Oxford Applied Mathematics and Computing Science Series, 1990.

[2] D. N. Arnold. An interior penalty finite element method with discontinuous elements. *SIAM Journal on Numerical Analysis*, 19:742–760, 1982.

[3] D. N. Arnold and N. Douglas. An interior penalty finite element method with discontinuous elements. *SIAM Journal on Numerical Analysis*, 19, 08 1982.

[4] D. Boffi, F. Brezzi, and M. Fortin. *Mixed Finite Element Methods and Applications*. Springer, 2013.

[5] S. Brenner and R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer, 2008.

[6] B. Cockburn, G. Kanschat, and D. Schötzau. A note on discontinuous Galerkin divergence-free solutions of the Navier–Stokes equations. *J. Sci. Comput.*, 31(1/2):61–73, 2007.

[7] H. Elman, D. Silvester, and A. Wathen. *FInite Elements and Fast Iterative Solvers with Applications in Incompressible Fluid Dynamics second edition*. 2014.

[8] A. Ern and J. Guermond. *Theory and practice of finite elements*. Springer, 2004.

[9] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, 2010.

[10] G. P. Galdi. *An Introduction to the Mathematical Theory of the Navier-Stokes Equations Volume I*, volume 38. Springer Tracts in Natural Philosophy, 1994.

[11] V. John, A. Linke, C. Merdon, M. Neilan, and L. G. Rebholz. On the divergence constraint in mixed finite element methods for incompressible flows. *SIAM Journal on Numerical Analysis*, 59(3):492–544, 2017.

[12] R. J. Labeur and G. Wells. Energy stable and momentum conserving hybrid finite element method for the incompressible Navier-Stokes equations. *SIAM Journal on Scientific Computing*, 34(2):A889–A913, 2012.

[13] A Linke and Christian Merdon. Pressure-robustness and discrete helmholtz projectors in mixed finite element methods for the incompressible navierstokes equations. *Computer Methods in Applied Mechanics and Engineering*, 311, 08 2016.

[14] A Linke, Christian Merdon, M Neilan, and Felix Neumann. Quasi-optimality of a pressure-robust nonconforming finite element method for the stokes-problem. *Mathematics of Computation*, 87:1, 12 2017.

[15] D. A. Di Pietro and A. Ern. *Mathematical Aspects of Discontinuous Galerkin Methods*. Springer, 2012.

[16] S. Rhebergen and G. Wells. Analysis of a hybridized/interface stabilized finite element method for the stokes equations. *SIAM Journal on Numerical Analysis*, 55(4):1982–2003, August 2017.

[17] S. Rhebergen and G. Wells. An embedded-hybridized discontinuous galerkin finite element method for the stokes equations. 2018.

[18] B. Riviere. *Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations*. SIAM, 2008.

[19] F. Sayas. From raviart-thomas to hdg. 2013.

[20] J. Schöberl. C++11 implementation of finite elements in NGSolve. Technical Report ASC Report 30/2014, Institute for Analysis and Scientific Computing, Vienna University of Technology, 2014.

[21] J. Wang and X. Ye. New finite element methods in computational fluid dynamics by h(div) elements. *SIAM Journal on Numerical Analysis*, 45(3):1269–1286, 2007.

[22] G. N. Wells. Analysis of an interface stabilized finite element method: the advection-diffusion reaction equation. *SIAM J. Numer. Anal.*, 49:87–109, 2011.

APPENDICES

Appendix A

The DG Method for Stokes

We derive the bilinear form a_h for an $H(\mathrm{div})$-conforming DG method (see eq. (4.19)). It is based on the Symmetric Interior Penalty (SIP) method introduced by Arnold [3]. In contrast with the derivation in [15, section 4.2], here we introduce only the tangential jumps across the element interfaces since $H(\mathrm{div})$-conformity implies normal continuity.

For simplicity, consider the non-dimensionalized Stokes equations:

$$-\nabla^2 \boldsymbol{u} + \nabla p = \boldsymbol{f}, \tag{A.1a}$$

$$\nabla \cdot \boldsymbol{u} = g, \tag{A.1b}$$

$$\boldsymbol{u}|_{\partial\Omega} = 0, \tag{A.1c}$$

$$\langle p \rangle_\Omega = 0. \tag{A.1d}$$

Multiply the eq. (A.1a) by an arbitrary $\boldsymbol{v}_h \in \boldsymbol{V}_h$ (recall the definition in eq. (4.17)) and rewrite the integral as the sum of integrals on each element:

$$\sum_{K\in\mathcal{T}_h} \left[-\int_K \nabla^2 \boldsymbol{u} \cdot \boldsymbol{v}_h \,\mathrm{d}\boldsymbol{x} + \int_K \nabla p \cdot \boldsymbol{v}_h \,\mathrm{d}\boldsymbol{x} \right] = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}_h \,\mathrm{d}\boldsymbol{x},$$

Apply the divergence theorem to both integrals inside the square bracket. The left hand side becomes:

$$\sum_{K\in\mathcal{T}_h} \left[\int_K \nabla \boldsymbol{u} : \nabla \boldsymbol{v} \,\mathrm{d}\boldsymbol{x} - \int_{\partial K} \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{n}_K} \cdot \boldsymbol{v}_h \,\mathrm{d}\boldsymbol{s} - \int_K p \nabla \cdot \boldsymbol{v}_h \,\mathrm{d}\boldsymbol{x} + \int_{\partial K} p \boldsymbol{v}_h \cdot \boldsymbol{n}_K \,\mathrm{d}\boldsymbol{s} \right], \tag{A.2}$$

where

$$\sum_{K\in\mathcal{T}_h}\int_{\partial K} p\boldsymbol{v}_h\cdot\boldsymbol{n}_K\,\mathrm{d}\boldsymbol{s} = \sum_{F\in\mathcal{F}_h^B}\int_F p\boldsymbol{v}_h\cdot\boldsymbol{n}_K\,\mathrm{d}\boldsymbol{s} = 0,$$

where the first equality is due to the normal continuity of $\boldsymbol{v}_h$ across an element boundary and the assumption that $p\in H^1(\Omega)$; the second equality is because of the boundary condition of the weak problem. (Again, recall the definition in eq. (4.17).) For the sum of the second integral in eq. (A.2), we first consider the decomposition of $\boldsymbol{v}_h$ in normal and tangential directions respectively,

$$\begin{aligned}\boldsymbol{v}_h &= (\boldsymbol{v}_h\cdot\boldsymbol{n}_K)\boldsymbol{n}_K + (\boldsymbol{v}_h\cdot\boldsymbol{\tau}_K)\boldsymbol{\tau}_K,\\ \frac{\partial\boldsymbol{u}}{\partial\boldsymbol{n}_K}\cdot\boldsymbol{v}_h &= \frac{\partial\boldsymbol{u}\cdot\boldsymbol{n}_K}{\partial\boldsymbol{n}_K}(\boldsymbol{v}_h\cdot\boldsymbol{n}_K) + \frac{\partial\boldsymbol{u}\cdot\boldsymbol{\tau}_K}{\partial\boldsymbol{n}_K}(\boldsymbol{v}_h\cdot\boldsymbol{\tau}_K).\end{aligned}$$

Thus,

$$\begin{aligned}\sum_{K\in\mathcal{T}_h}\int_{\partial K}\frac{\partial\boldsymbol{u}}{\partial\boldsymbol{n}_K}\cdot\boldsymbol{v}_h\,\mathrm{d}\boldsymbol{s} &= \sum_{K\in\mathcal{T}_h}\int_{\partial K}\left[\frac{\partial\boldsymbol{u}\cdot\boldsymbol{n}_K}{\partial\boldsymbol{n}_K}(\boldsymbol{v}_h\cdot\boldsymbol{n}_K) + \frac{\partial\boldsymbol{u}\cdot\boldsymbol{\tau}_K}{\partial\boldsymbol{n}_K}(\boldsymbol{v}_h\cdot\boldsymbol{\tau}_K)\right]\mathrm{d}\boldsymbol{s}\\ &= \sum_{K\in\mathcal{T}_h}\int_{\partial K}\frac{\partial\boldsymbol{u}\cdot\boldsymbol{\tau}_K}{\partial\boldsymbol{n}_K}(\boldsymbol{v}_h\cdot\boldsymbol{\tau}_K)\,\mathrm{d}\boldsymbol{s},\end{aligned}\tag{A.3}$$

where the second equality is again due to normal continuity of $\boldsymbol{v}_h$ across an elment boundary and the assumption that $\boldsymbol{u}\in\boldsymbol{H}^2(\Omega)$ (which implies the continuity of $\boldsymbol{u}$). We next simplify eq. (A.3).

Consider a single edge/facet $F\in\mathcal{F}_h^I$ as illustrated in fig. A.1. Then

$$\begin{aligned}&\int_F\frac{\partial\boldsymbol{u}|_{K_1}\cdot\boldsymbol{\tau}_1}{\partial\boldsymbol{n}_1}(\boldsymbol{v}_h|_{K_1}\cdot\boldsymbol{\tau}_1)\,\mathrm{d}\boldsymbol{s} + \int_F\frac{\partial\boldsymbol{u}|_{K_2}\cdot\boldsymbol{\tau}_2}{\partial\boldsymbol{n}_2}(\boldsymbol{v}_h|_{K_2}\cdot\boldsymbol{\tau}_2)\,\mathrm{d}\boldsymbol{s}\\ &\qquad= -\int_F\frac{\partial\boldsymbol{u}|_{K_1}\cdot\boldsymbol{\tau}_1}{\partial\boldsymbol{n}_1}(\boldsymbol{v}_h|_{K_1}\cdot\boldsymbol{\tau}_F)\,\mathrm{d}\boldsymbol{s} + \int_F\frac{\partial\boldsymbol{u}|_{K_2}\cdot\boldsymbol{\tau}_2}{\partial\boldsymbol{n}_2}(\boldsymbol{v}_h|_{K_2}\cdot\boldsymbol{\tau}_F)\,\mathrm{d}\boldsymbol{s},\\ &\qquad= \int_F\frac{1}{2}\left(\frac{\partial\boldsymbol{u}|_{K_1}\cdot\boldsymbol{\tau}_1}{\partial\boldsymbol{n}_1} + \frac{\partial\boldsymbol{u}|_{K_2}\cdot\boldsymbol{\tau}_2}{\partial\boldsymbol{n}_2}\right)\left(\boldsymbol{v}_h|_{K_2}\cdot\boldsymbol{\tau}_F - \boldsymbol{v}_h|_{K_1}\cdot\boldsymbol{\tau}_F\right)\mathrm{d}\boldsymbol{s}\\ &\qquad\quad+ \int_F\frac{1}{2}\underbrace{\left(\frac{\partial\boldsymbol{u}|_{K_1}\cdot\boldsymbol{\tau}_1}{\partial\boldsymbol{n}_1} - \frac{\partial\boldsymbol{u}|_{K_2}\cdot\boldsymbol{\tau}_2}{\partial\boldsymbol{n}_2}\right)}_{=0\text{ due to continuity}}\left(\boldsymbol{v}_h|_{K_2}\cdot\boldsymbol{\tau}_F + \boldsymbol{v}_h|_{K_1}\cdot\boldsymbol{\tau}_F\right)\mathrm{d}\boldsymbol{s}\\ &\qquad= \int_F\frac{1}{2}\left(\frac{\partial\boldsymbol{u}|_{K_1}\cdot\boldsymbol{\tau}_1}{\partial\boldsymbol{n}_1} + \frac{\partial\boldsymbol{u}|_{K_2}\cdot\boldsymbol{\tau}_2}{\partial\boldsymbol{n}_2}\right)\left(\boldsymbol{v}_h|_{K_2}\cdot\boldsymbol{\tau}_2 + \boldsymbol{v}_h|_{K_1}\cdot\boldsymbol{\tau}_1\right)\mathrm{d}\boldsymbol{s},\end{aligned}$$

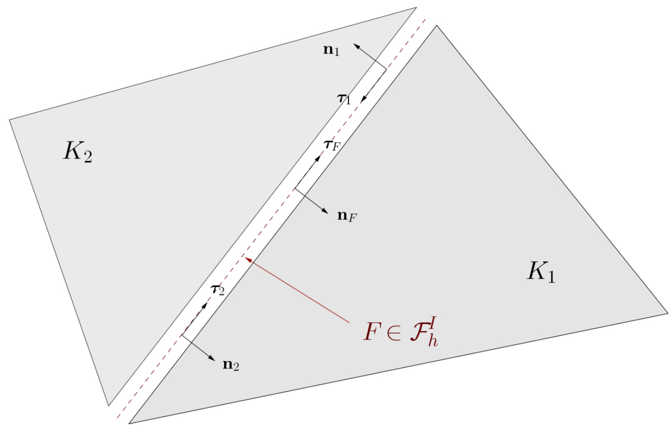


Figure A.1: Notation for an edge/facet. Note that the normal direction of an arbitrary $F \in \mathcal{F}_h^I$ is determined once there is a mesh and does not change afterwards.

where the second equality uses the identity:

$$a_2 b_2 - a_1 b_1 = \frac{1}{2}(b_2 + b_1)(a_2 - a_1) + \frac{1}{2}(a_2 + a_1)(b_2 - b_1).$$

To simplify notation we define average and jump operators on interior edges/facets as:

$$\{\!\{\epsilon(\boldsymbol{u})\}\!\}|_F := \frac{1}{2}\left(\frac{\partial \boldsymbol{u}|_{K_1} \cdot \boldsymbol{\tau}_1}{\partial \boldsymbol{n}_1} + \frac{\partial \boldsymbol{u}|_{K_2} \cdot \boldsymbol{\tau}_2}{\partial \boldsymbol{n}_2}\right) \tag{A.4a}$$

$$[\![\boldsymbol{u}]\!]|_F := \boldsymbol{u}|_{K_1} \cdot \boldsymbol{\tau}_1 + \boldsymbol{u}|_{K_2} \cdot \boldsymbol{\tau}_2. \tag{A.4b}$$

On boundary edges/facets, the average and jump operators are defined as:

$$\{\!\{\epsilon(\boldsymbol{u})\}\!\}|_F := \frac{\partial \boldsymbol{u}|_{K_1} \cdot \boldsymbol{\tau}_1}{\partial \boldsymbol{n}_1} \tag{A.5a}$$

$$[\![\boldsymbol{u}]\!]|_F := \boldsymbol{u}|_{K_1} \cdot \boldsymbol{\tau}_1. \tag{A.5b}$$

We can therefore write eq. (A.3) as:

$$\sum_{K \in \mathcal{T}_h} \int_{\partial K} \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{n}_K} \cdot \boldsymbol{v}_h \, \mathrm{d}\boldsymbol{s} = \sum_{F \in \mathcal{F}_h} \int_F \{\!\{\epsilon(\boldsymbol{u})\}\!\} [\![\boldsymbol{v}_h]\!] \, \mathrm{d}\boldsymbol{s}.$$

We observe from eq. (A.2) that we can define the bilinear form $a_h(\cdot,\cdot)$ as:

$$a_h(\boldsymbol{u}_h,\boldsymbol{v}_h) := \int_\Omega \nabla_h \boldsymbol{u}_h : \nabla_h \boldsymbol{v}_h \,\mathrm{d}\boldsymbol{x} - \sum_{F\in\mathcal{F}_h} \int_F \{\!\{\epsilon(\boldsymbol{u}_h)\}\!\} [\![\boldsymbol{v}_h]\!] \,\mathrm{d}\boldsymbol{s}. \tag{A.6}$$

However, $a_h(\cdot,\cdot)$ defined this way is not symmetric. To obtain symmetry we add a consistent term:

$$a_h(\boldsymbol{u}_h,\boldsymbol{v}_h) := \int_\Omega \nabla_h \boldsymbol{u}_h : \nabla_h \boldsymbol{v}_h \,\mathrm{d}\boldsymbol{x} - \sum_{F\in\mathcal{F}_h} \int_F \{\!\{\epsilon(\boldsymbol{u}_h)\}\!\} [\![\boldsymbol{v}_h]\!] \,\mathrm{d}\boldsymbol{s} - \sum_{F\in\mathcal{F}_h} \int_F \{\!\{\epsilon(\boldsymbol{v}_h)\}\!\} [\![\boldsymbol{u}_h]\!] \,\mathrm{d}\boldsymbol{s}. \tag{A.7}$$

To see that the extra term is consistent, we observe that for $\boldsymbol{u} \in \boldsymbol{H}^2(\Omega) \cap \boldsymbol{H}^1_0(\Omega)$:

$$\begin{aligned} [\![\boldsymbol{u}]\!]|_{F\in\mathcal{F}_h^I} &= \boldsymbol{u}|_{K_1}\cdot\boldsymbol{\tau}_1 + \boldsymbol{u}|_{K_2}\cdot\boldsymbol{\tau}_2 = 0,\\ [\![\boldsymbol{u}]\!]|_{F\in\mathcal{F}_h^B} &= \boldsymbol{u}|_{K_1}\cdot\boldsymbol{\tau}_1 = 0. \end{aligned}$$

Unfortunately, $a_h(\cdot,\cdot)$ given in eq. (A.7) is not coercive. To see this, note that

$$a_h(\boldsymbol{v}_h,\boldsymbol{v}_h) := \int_\Omega \nabla_h \boldsymbol{v}_h : \nabla_h \boldsymbol{v}_h \,\mathrm{d}\boldsymbol{x} - 2\sum_{F\in\mathcal{F}_h} \int_F \{\!\{\epsilon(\boldsymbol{v}_h)\}\!\} [\![\boldsymbol{v}_h]\!] \,\mathrm{d}\boldsymbol{s}.$$

We note that the sign of $a_h(\boldsymbol{v}_h,\boldsymbol{v}_h)$ is undetermined. To fix this, we add a stabilization term:

$$\begin{aligned} a_h(\boldsymbol{u}_h,\boldsymbol{v}_h) := \int_\Omega \nabla_h \boldsymbol{u}_h : \nabla_h \boldsymbol{v}_h \,\mathrm{d}\boldsymbol{x} - \sum_{F\in\mathcal{F}_h} \int_F \{\!\{\epsilon(\boldsymbol{u}_h)\}\!\} [\![\boldsymbol{v}_h]\!] \,\mathrm{d}\boldsymbol{s} - \sum_{F\in\mathcal{F}_h} \int_F \{\!\{\epsilon(\boldsymbol{v}_h)\}\!\} [\![\boldsymbol{u}_h]\!] \,\mathrm{d}\boldsymbol{s} \\ + \frac{\sigma}{h_F} \sum_{F\in\mathcal{F}_h} \int_F [\![\boldsymbol{u}_h]\!][\![\boldsymbol{v}_h]\!] \,\mathrm{d}\boldsymbol{s}, \end{aligned}$$

where σ is the stablization parameter that needs to be chosen sufficiently large to ensure stability. Note again that the extra term is consistent.

Appendix B

Variant of the Trace Inequality

Lemma B.1 (Young's Inequality)**.** *Let* $a, b, \epsilon \in \mathbb{R}$*. Then,*

$$ab \leq \frac{a^2}{2\epsilon} + \frac{\epsilon b^2}{2}.$$

Proof. This is direct consequence of

$$0 \leq (a - \epsilon b)^2 = a^2 - 2\epsilon ab + (\epsilon b)^2.$$

□

Now we use Young's Inequality to prove lemma 4.3.

Proof. By the continuous trace inequality [15, Lemma 1.49], we have

$$\begin{aligned}
\|v\|_F^2 &\leq C \left(|v|_{1,K} + h_K^{-1} \|v\|_K \right) \|v\|_K \,, \\
&\leq C \left(|v|_{1,K} \|v\|_K + h_K^{-1} \|v\|_K^2 \right), \\
&\leq C \left(\frac{h_K |v|_{1,K}^2}{2} + \frac{\|v\|_K^2}{2h_K} + h_K^{-1} \|v\|_K^2 \right), \\
&\leq C \left(h_K^{-1} \|v\|_K^2 + h_K |v|_{1,K}^2 \right).
\end{aligned}$$

□