Part XI, Chapter 49

Well-posedness for PDEs in mixed form

In Part XI, composed of Chapters 49 to 55, we study the well-posedness and the finite element approximation of PDEs formulated in mixed form. Mixed formulations are often obtained from elliptic PDEs by introducing one or more auxiliary variables. One reason for introducing these variables can be that they have some physical relevance. For instance, one can think of the flux in Darcy's equations (see Chapter 51). Another reason can be to relax a constraint imposed on a variational formulation. This is the case for the Stokes equations where the pressure results from the incompressibility constraint enforced on the velocity field (see Chapter 53). The PDEs considered in this part enjoy a coercivity property on the primal variable, but not on the auxiliary variable, so that the analysis relies on inf-sup conditions. The goal of the present chapter is to identify necessary and sufficient conditions for the well-posedness of a model problem that serves as a prototype for PDEs in mixed form. The finite element approximation of this model problem is investigated in Chapter 50.

49.1 Model problems

We introduce in this section some model problems illustrating the concept of PDEs in mixed form. Let D be a domain in $\mathbb{R}^d$, i.e., D is a nonempty, open, bounded, connected subset of $\mathbb{R}^d$ (see Definition 3.1).

49.1.1 Darcy

Consider the elliptic PDE $-\nabla\cdot(\mathbb{d}\nabla p) = f$ in D; see §31.1. Introducing the flux (or filtration velocity) $\boldsymbol{\sigma} := -\mathbb{d}\nabla p$ leads to the following mixed formulation known as Darcy's equations (see §24.1.2):

$$\mathbb{d}^{-1}\boldsymbol{\sigma} + \nabla p = \mathbf{0} \qquad \text{in } D, \tag{49.1a}$$
$$\nabla{\cdot}\boldsymbol{\sigma} = f \qquad \text{in } D. \tag{49.1b}$$

Here, (49.1a) is a phenomenological law relating the flux to the gradient of the primal unknown p (a nonzero right-hand side can be considered as well). The equation (49.1b) expresses mass conservation. For simplicity, we assume that (49.1) is equipped with the boundary condition $p_{|\partial D} = 0$.

Let us give a more abstract form to the above problem by setting

$$\boldsymbol{V} := \boldsymbol{L}^2(D), \qquad Q := H_0^1(D). \tag{49.2}$$

Consider the linear operators $A : \boldsymbol{V} \to \boldsymbol{V}' = \boldsymbol{L}^2(D)$ (owing to the Riesz–Fréchet representation theorem) and $B : \boldsymbol{V} \to Q' = H^{-1}(D)$ defined by setting $A(\boldsymbol{\tau}) := \mathbb{d}^{-1}\boldsymbol{\tau}$ and $B(\boldsymbol{\tau}) := -\nabla{\cdot}\boldsymbol{\tau}$. Under appropriate boundedness assumptions on $\mathbb{d}^{-1}$, the linear operators A and B are bounded. Using the identification $(H_0^1(D))'' = H_0^1(D)$, we have $B^* : Q \to \boldsymbol{V}'$ and $\langle B^*(q), \boldsymbol{\tau}\rangle_{\boldsymbol{V}',\boldsymbol{V}} = \langle q, B(\boldsymbol{\tau})\rangle_{H_0^1(D),H^{-1}(D)} = (\nabla q, \boldsymbol{\tau})_{\boldsymbol{L}^2(D)}$ for all $q \in H_0^1(D)$ and all $\boldsymbol{\tau} \in \boldsymbol{L}^2(D)$. Hence, $B^*(q) = \nabla q$ for all $q \in H_0^1(D)$.

An alternative point of view consists of setting

$$\boldsymbol{V} := \boldsymbol{H}(\text{div}; D), \qquad Q := L^2(D). \tag{49.3}$$

Then $A : \boldsymbol{V} \to \boldsymbol{V}'$ is defined by setting $A(\boldsymbol{\tau}) := \mathbb{d}^{-1}\boldsymbol{\tau}$ (where we use that $\boldsymbol{V} \hookrightarrow \boldsymbol{L}^2(D) \equiv \boldsymbol{L}^2(D)' \hookrightarrow \boldsymbol{V}'$) and $B : \boldsymbol{V} \to Q' = L^2(D)$ (owing to the Riesz–Fréchet representation theorem) is defined by setting $B(\boldsymbol{\tau}) := -\nabla{\cdot}\boldsymbol{\tau}$. The adjoint operator $B^* : Q \to \boldsymbol{V}'$ is s.t. $\langle B^*(q), \boldsymbol{\tau}\rangle_{\boldsymbol{V}',\boldsymbol{V}} = (q, B(\boldsymbol{\tau}))_{L^2(D)} = (q, -\nabla{\cdot}\boldsymbol{\tau})_{\boldsymbol{L}^2(D)}$ for all $q \in L^2(D)$ and all $\boldsymbol{\tau} \in \boldsymbol{H}(\text{div}; D)$. Let us have a closer look at B^*. Let $q \in L^2(D)$ and assume that there exists $\boldsymbol{g} \in \boldsymbol{L}^2(D)$ so that $\langle B^*(q), \boldsymbol{\tau}\rangle_{\boldsymbol{V}',\boldsymbol{V}} = (\boldsymbol{g}, \boldsymbol{\tau})_{\boldsymbol{L}^2(D)}$. This implies that $(q, -\nabla{\cdot}\boldsymbol{\tau})_{\boldsymbol{L}^2(D)} = (\boldsymbol{g}, \boldsymbol{\tau})_{\boldsymbol{L}^2(D)}$ for all $\boldsymbol{\tau} \in \boldsymbol{H}(\text{div}; D)$. Taking $\boldsymbol{\tau} \in \boldsymbol{C}_0^\infty(D)$ shows that q has a weak derivative and $\nabla q = \boldsymbol{g}$. Hence, $q \in H^1(D)$ and $(q, \nabla{\cdot}\boldsymbol{\tau})_{\boldsymbol{L}^2(D)} + (\nabla q, \boldsymbol{\tau})_{\boldsymbol{L}^2(D)} = 0$ for all $\boldsymbol{\tau} \in \boldsymbol{H}(\text{div}; D)$. Moreover, considering the trace of q on ∂D, $\gamma^{\text{g}}(q) \in H^{\frac12}(\partial D)$, and using the surjectivity of the normal trace operator $\gamma^{\text{d}} : \boldsymbol{H}(\text{div}; D) \to H^{-\frac12}(\partial D)$ (see Theorem 4.15), we infer that for all $\phi \in H^{-\frac12}(\partial D)$, there is $\boldsymbol{\tau}_\phi \in \boldsymbol{H}(\text{div}; D)$ s.t. $\gamma^{\text{d}}(\boldsymbol{\tau}_\phi) = \phi$. Then $\langle \phi, \gamma^{\text{g}}(q)\rangle_{H^{-\frac12},H^{\frac12}} = \langle \gamma^{\text{d}}(\boldsymbol{\tau}_\phi), \gamma^{\text{g}}(q)\rangle_{H^{-\frac12},H^{\frac12}} = (q, \nabla{\cdot}\boldsymbol{\tau}_\phi)_{L^2(D)} + (\nabla q, \boldsymbol{\tau}_\phi)_{\boldsymbol{L}^2(D)} = 0$ for all $\phi \in H^{-\frac12}(\partial D)$. Hence, $\gamma^{\text{g}}(q) = 0$, i.e., $q_{|\partial D} = 0$. This shows that $B^*(q) \in \boldsymbol{L}^2(D)$ encodes the boundary condition $q_{|\partial D} = 0$ in a weak sense.

In conclusion, regardless of the chosen setting, the Darcy problem (49.1) can be reformulated as follows:

$$\begin{cases} \text{Find } \boldsymbol{\sigma} \in \boldsymbol{V} \text{ and } p \in Q \text{ such that} \\ A(\boldsymbol{\sigma}) + B^*(p) = \mathbf{0}, \\ B(\boldsymbol{\sigma}) = -f. \end{cases} \tag{49.4}$$

The mixed finite element approximation of (49.4) is studied in Chapter 51.

49.1.2 Stokes

The Stokes equations model steady incompressible flows in which inertia forces are negligible. The problem is written in the following mixed form:

$$\nabla{\cdot}(-\mu\mathbb{e}(\boldsymbol{u})) + \nabla p = \boldsymbol{f} \qquad \text{in } D, \tag{49.5a}$$
$$\nabla{\cdot}\boldsymbol{u} = 0 \qquad \text{in } D, \tag{49.5b}$$

where $\mu > 0$ is the viscosity, $\boldsymbol{u} : D \to \mathbb{R}^d$ the velocity field with the (linearized) strain rate tensor $\mathbb{e}(\boldsymbol{u}) := \frac{1}{2}(\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^{\mathsf{T}}) : D \to \mathbb{R}^{d\times d}$, $p : D \to \mathbb{R}$ the pressure, and $\boldsymbol{f} : D \to \mathbb{R}^d$ the body force. The equation (49.5a) expresses the momentum balance in the flow, and (49.5b) the mass balance. For simplicity, we assume that (49.5) is equipped with the boundary condition $\boldsymbol{u}_{|\partial D} = \boldsymbol{0}$.

Let us set

$$\boldsymbol{V} := \boldsymbol{H}^1_0(D), \qquad Q := L^2(D), \tag{49.6}$$

and let us define $A : \boldsymbol{V} \to \boldsymbol{V}' = \boldsymbol{H}^{-1}(D)$, $B : \boldsymbol{V} \to Q' = L^2(D)$ (owing to the Riesz–Fréchet representation theorem) by setting $A(\boldsymbol{v}) := -\nabla{\cdot}(\mu\mathbb{e}(\boldsymbol{v}))$ and $B(\boldsymbol{v}) := -\nabla{\cdot}\boldsymbol{v}$. The adjoint operator $B^* : Q \to \boldsymbol{V}'$ is s.t. $\langle B^*(q), \boldsymbol{v}\rangle_{\boldsymbol{V}',\boldsymbol{V}} = (q, -\nabla{\cdot}\boldsymbol{v})_{L^2(D)}$ for all $\boldsymbol{v} \in \boldsymbol{V}$ and all $q \in Q$. This means that $B^*(q) = \nabla q$ for all $q \in L^2(D)$. In conclusion, the Stokes problem (49.5) can be reformulated as follows:

$$\begin{cases} \text{Find } \boldsymbol{v} \in V \text{ and } p \in Q \text{ such that} \\ A(\boldsymbol{v}) + B^*(p) = \boldsymbol{f}, \\ B(\boldsymbol{v}) = 0. \end{cases} \tag{49.7}$$

The finite element approximation of (49.7) is studied in Chapters 53 to 55.

49.1.3 Maxwell

Consider the model problem (43.4) for Maxwell's equations in the time-harmonic regime (see §43.1.1) in the limit $\omega \to 0$ and with the boundary condition $\boldsymbol{H}_{|\partial D}{\times}\boldsymbol{n} = \boldsymbol{0}$. Ampère's equation (43.4a) gives $-\boldsymbol{E} + \frac{1}{\sigma}\nabla{\times}\boldsymbol{H} = \frac{1}{\sigma}\boldsymbol{j}_{\mathrm{s}}$, and Faraday's equation (43.4b) gives $\nabla{\cdot}(\mu\boldsymbol{H}) = 0$ and $\nabla{\times}\boldsymbol{E} = \boldsymbol{0}$ (since $\omega \to 0$). Setting $\kappa := \sigma^{-1}$ the strong form of this problem consists of looking for a field $\boldsymbol{H} : D \to \mathbb{R}^3$ such that $\nabla{\times}(\kappa\nabla{\times}\boldsymbol{H}) = \boldsymbol{f}$, with $\boldsymbol{f} := \nabla{\times}(\kappa\boldsymbol{j}_{\mathrm{s}}) : D \to \mathbb{R}^3$, $\boldsymbol{H}_{|\partial D}{\times}\boldsymbol{n} = \boldsymbol{0}$, and under the constraint $\nabla{\cdot}(\mu\boldsymbol{H}) = \boldsymbol{0}$. The dual variable of this constraint is a scalar-valued function $\phi : D \to \mathbb{R}$ with the boundary condition $\phi_{|\partial D} = 0$, leading to the following mixed formulation (see, e.g., Kanayama et al. [263]):

$$\nabla{\times}(\kappa\nabla{\times}\boldsymbol{H}) + \nu\nabla\phi = \boldsymbol{f} \qquad \text{in } D, \tag{49.8a}$$
$$\nabla{\cdot}(\mu\boldsymbol{H}) = 0 \qquad \text{in } D. \tag{49.8b}$$

Let us set

$$\boldsymbol{V} := \boldsymbol{H}_0(\text{curl}; D), \qquad Q := H^1_0(D), \tag{49.9}$$

and let us define $A : \boldsymbol{V} \to \boldsymbol{V}'$, $B : \boldsymbol{V} \to Q' = H^{-1}(D)$ by setting $A(\boldsymbol{v}) := \nabla\times(\kappa\nabla\times\boldsymbol{v})$ and $B(\boldsymbol{v}) := -\nabla\cdot(\nu\boldsymbol{v})$. Using the identification $(H^1_0(D))'' = H^1_0(D)$, the adjoint operator $B^* : Q \to \boldsymbol{V}'$ is s.t. $\langle B^*(\psi), \boldsymbol{v}\rangle_{\boldsymbol{V}',\boldsymbol{V}} = \langle \psi, -\nabla\cdot(\nu\boldsymbol{v})\rangle_{H^1_0(D),H^{-1}(D)} = (\nu\nabla\psi, \boldsymbol{v})_{\boldsymbol{L}^2(D)}$ for all $\boldsymbol{v} \in V$ and all $\psi \in Q$. This means that $B^*(\psi) = \nu\nabla\psi$ for all $\psi \in H^1_0(D)$. In conclusion, the Maxwell problem (49.8) can be reformulated as follows:

$$\begin{cases} \text{Find } \boldsymbol{H} \in \boldsymbol{V} \text{ and } \phi \in Q \text{ such that} \\ A(\boldsymbol{H}) + B^*(\phi) = \boldsymbol{f}, \\ B(\boldsymbol{H}) = 0. \end{cases} \tag{49.10}$$

Some further aspects of this problem are considered in Exercise 49.6.

49.2 Well-posedness in Hilbert spaces

Consider two real Hilbert spaces V and Q, and two operators $A \in \mathcal{L}(V;V')$, $B \in \mathcal{L}(V;Q')$. We identify $Q'' = Q$. Our goal in this section is to investigate the well-posedness of the following mixed model problem:

$$\begin{cases} \text{Find } u \in V \text{ and } p \in Q \text{ such that} \\ A(u) + B^*(p) = f, \\ B(u) = g, \end{cases} \tag{49.11}$$

for all $(f,g) \in V'\times Q'$. We assume in the entire section that A is self-adjoint and coercive. This assumption simplifies many arguments. In particular, we establish the well-posedness of (49.11) by means of a coercivity argument on the Schur complement. The complete theory for well-posedness in Banach spaces is done in §49.4. Let α be the coercivity constant of A,

$$\inf_{v\in V} \frac{\langle A(v), v\rangle_{V',V}}{\|v\|_V^2} =: \alpha > 0. \tag{49.12}$$

We also assume that B is surjective, i.e., recalling Lemma C.40,

$$\inf_{q\in Q} \frac{\|B^*(q)\|_{V'}}{\|q\|_Q} =: \beta > 0. \tag{49.13}$$

We denote $\|a\| := \|A\|_{\mathcal{L}(V;V')}$ and $\|b\| := \|B\|_{\mathcal{L}(V';Q)}$.

49.2.1 Schur complement

Let $J_Q : Q \to Q'$ be the Riesz–Fréchet isometric isomorphism (see Theorem C.24), i.e., $\langle J_Q(q), r\rangle_{Q',Q} := (q, r)_Q$ for all $q, r \in Q$. We call Schur complement of A on Q the linear operator $S : Q \to Q$ defined by

$$S := J_Q^{-1} B A^{-1} B^*. \tag{49.14}$$

(S is sometimes defined with the opposite sign in the literature.)

Lemma 49.1 (Coercivity and boundedness of S). *Let S be defined in* (49.14)*. Then S is symmetric and bijective with*

$$\frac{\beta^2}{\|a\|}\|q\|_Q^2 \le (S(q), q)_Q \le \frac{\|b\|^2}{\alpha}\|q\|_Q^2, \qquad \forall q \in Q. \tag{49.15}$$

Proof. (1) Symmetry. Since A^{-1} is self-adjoint, we infer that for all $q, r \in Q$,

$$\begin{aligned}(S(q), r)_Q &= \langle BA^{-1}B^*(q), r\rangle_{Q',Q} = \langle A^{-1}B^*(q), B^*(r)\rangle_{V,V'} \\ &= \langle A^{-1}B^*(r), B^*(q)\rangle_{V,V'} = (S(r), q)_Q.\end{aligned}$$

(2) Bounds (49.15). The self-adjointness and coercivity of A imply that $\|A^{-1}\|_{\mathcal{L}(V';V)} = \alpha^{-1}$ (see Lemma C.51) and $\langle A^{-1}(\phi), \phi\rangle_{V,V'} \ge \frac{1}{\|a\|}\|\phi\|_{V'}^2$ for all $\phi \in V'$ (see Lemma C.63). Moreover, the definitions of $\|b\|$ and β mean that $\|B^*\|_{\mathcal{L}(Q;V')} = \|b\|$ and $\|B^*(q)\|_{V'} \ge \beta\|q\|_Q$ for all $q \in Q$. Since $(S(q), q)_Q = \langle A^{-1}(B^*(q)), B^*(q)\rangle_{V,V'}$ for all $q \in Q$, we conclude that (49.15) holds true. Finally, S is bijective since S is Q-coercive and bounded. □

Lemma 49.2 (Equivalence with (49.11)**).** *Let $(u, p) \in V \times Q$. Then the pair (u, p) solves* (49.11) *iff (u, p) solves*

$$S(p) = J_Q^{-1}(BA^{-1}(f) - g), \qquad A(u) = f - B^*(p). \tag{49.16}$$

Proof. Let $(u, p) \in V \times Q$ solve (49.11). Since A is bijective, we have $u = A^{-1}(f - B^*(p))$, so that $B(u) = BA^{-1}(f - B^*(p)) = g$. This in turn implies that $J_Q^{-1}BA^{-1}(f - B^*(p)) = J_Q^{-1}(g)$, finally giving $S(p) = J_Q^{-1}(BA^{-1}(f) - g)$ and $A(u) = f - B^*(p)$. This means that (u, p) solves (49.16). Conversely, assume that $(u, p) \in V \times Q$ solves (49.16). Then $BA^{-1}B^*(p) = BA^{-1}(f) - g$, that is, $BA^{-1}(f - B^*(p)) = g$. But $A^{-1}(f - B^*(p)) = u$. Hence, $B(u) = g$ and $A(u) = f - B^*(p)$, which means that (u, p) solves (49.11). □

Theorem 49.3 (Well-posedness). *The problem* (49.11) *is well-posed.*

Proof. Owing to Lemma 49.2, it suffices to show that (49.16) is well-posed, but this is a consequence of Lemma 49.1 and the Lax–Milgram lemma. □

49.2.2 Formulation with bilinear forms

We now reformulate the mixed problem (49.11) using bilinear forms. This formalism will be used in Chapters 50 to 55 where we consider various Galerkin approximations to (49.11). Let us set

$$a(v,w) := \langle A(v), w\rangle_{V',V}, \qquad b(w,q) := \langle q, B(w)\rangle_{Q,Q'}, \tag{49.17}$$

for all $v, w \in V$ and all $q \in Q$ (recall that we have identified Q'' and Q). The assumed boundedness of A and B implies that a and b are bounded on $V \times V$ and on $V \times Q$, respectively. The abstract problem (49.11) can then be reformulated in the following equivalent form:

$$\begin{cases} \text{Find } u \in V \text{ and } p \in Q \text{ such that} \\ a(u,w) + b(w,p) = f(w), & \forall w \in V, \\ b(u,q) = g(q), & \forall q \in Q, \end{cases} \tag{49.18}$$

with the shorthand notation $f(w) := \langle f, w\rangle_{V',V}$ and $g(q) := \langle g, q\rangle_{Q',Q}$. The definitions (49.12) and (49.13) of α and β are then equivalent to

$$\inf_{v\in V} \frac{|a(v,v)|}{\|v\|_V^2} =: \alpha > 0, \qquad \inf_{q\in Q} \sup_{v\in V} \frac{|b(v,q)|}{\|v\|_V \|q\|_Q} =: \beta > 0. \tag{49.19}$$

Let $X := V \times Q$ and consider the bilinear form $t : X \times X \to \mathbb{R}$ defined by

$$t((v,q),(w,r)) := a(v,w) + b(w,q) + b(v,r), \tag{49.20}$$

for all $(v,q),(w,r) \in X$. Then $(u,p) \in X$ solves (49.11) iff

$$t((u,p),(w,r)) = f(w) + g(r), \qquad \forall (w,r) \in X. \tag{49.21}$$

49.2.3 Sharper a priori estimates

We collect in this section some additional results regarding the operator S, and we give a priori estimates on the solution to the mixed problem (49.11). Recall from Definition 46.1 the notions of spectrum and eigenvalues.

Corollary 49.4 (Spectrum of S). *The spectrum of S is such that* $\sigma(S) \subset \left[\frac{\beta^2}{\|a\|}, \frac{\|b\|^2}{\alpha}\right]$, *and* $\|S\|_{\mathcal{L}(Q;Q)} \le \frac{\|b\|^2}{\alpha}$, $\|S^{-1}\|_{\mathcal{L}(Q;Q)} \le \frac{\|a\|}{\beta^2}$.

Proof. These statements are consequences of (49.15) and Theorem 46.17. Recall that Theorem 46.17 asserts in particular that $\sigma(S) \subset \mathbb{R}$, $\|S\|_{\mathcal{L}(Q;Q)} = \sup_{\lambda\in\sigma(S)} |\lambda|$ and $\|S^{-1}\|_{\mathcal{L}(Q;Q)} = \sup_{\lambda\in\sigma(S)} |\lambda|^{-1}$. See also Exercise 49.3. □

Remark 49.5 (Spectrum of S). Corollary 49.4 can be refined by equipping V with the energy norm $\|{\cdot}\|_a^2 := a(\cdot,\cdot)^{\frac12}$ (recall that a is symmetric and coercive) which is equivalent to $\|{\cdot}\|_V$. Setting $\beta_a := \inf_{q\in Q} \sup_{v\in V} \frac{|b(v,q)\|}{\|v\|_a \|q\|_Q}$

and $\|b\|_a := \sup_{q\in Q}\sup_{v\in V}\frac{|b(v,q)|}{\|v\|_a\|q\|_Q}$, we have $\|b\|_a^2 = \|S\|_{\mathcal{L}(Q;Q)}$, $\beta_a^{-2} = \|S^{-1}\|_{\mathcal{L}(Q;Q)}$, and $\{\beta_a^2, \|b\|_a^2\} \subset \sigma(S) \subset [\beta_a^2, \|b\|_a^2]$. □

We define the linear operator $T : X \to X$ such that

$$T(v,q) := (v + A^{-1}B^*(q), S^{-1}J_Q^{-1}B(v)), \tag{49.22}$$

for all $(v,q) \in V\times Q$. With a slight abuse of notation regarding the column vector convention, we can also write $T := \begin{pmatrix} I_V & A^{-1}B^* \\ S^{-1}J_Q^{-1}B & 0\end{pmatrix}$, where I_V is the identity in V. We have $T \in \mathcal{L}(X;X)$, and upon introducing the weighted inner product $(x,y)_{\widetilde{X}} := a(v,w) + (S(q),r)_Q$ for all $x := (v,q)$ and $y := (w,r)$ in X, we also have $(T(x),y)_{\widetilde{X}} = a(v,w) + b(v,r) + b(w,q)$, that is, $(T(x),y)_{\widetilde{X}} = t(x,y)$. This identity implies that T is symmetric with respect to the weighted inner product $(x,y)_{\widetilde{X}}$. The following result, due to Bacuta [42], provides a complete characterization of the spectrum of T. We refer the reader to §50.3.2 for the algebraic counterpart of this result.

Theorem 49.6 (Spectrum of T). *Let $\varrho := \frac{1+\sqrt{5}}{2}$ be the golden ratio. Assume that* $\ker(B)$ *is nontrivial. Then*

$$\sigma(T) = \sigma_{\mathrm{p}}(T) = \{-\varrho^{-1}, 1, \varrho\}. \tag{49.23}$$

Proof. Let $\lambda \in \sigma(T)$. Owing to Corollary 46.18, there is a sequence $(x_n)_{n\in\mathbb{N}}$ in X such that $\|x_n\|_X = 1$ for all $n \in \mathbb{N}$ and $T(x_n) - \lambda x_n \to 0$ in X as $n \to \infty$. Writing $x_n := (v_n, q_n)$, we infer that $(1-\lambda)v_n + A^{-1}B^*(q_n) \to 0$ in V and $S^{-1}J_Q^{-1}B(v_n) - \lambda q_n \to 0$ in Q as $n \to \infty$. Applying the bounded operator $J_Q^{-1}B$ to the first limit and the bounded operator S to the second one, we infer that

$$\begin{pmatrix} 1-\lambda & 1 \\ 1 & -\lambda \end{pmatrix}\begin{pmatrix} J_Q^{-1}B(v_n) \\ S(q_n)\end{pmatrix} \to 0.$$

Assume that $\lambda \notin \{-\varrho^{-1}, 1, \varrho\}$. The matrix on the left-hand side is invertible since $\lambda \notin \{-\varrho^{-1}, \varrho\}$. This implies that $S(q_n) \to 0$ in Q, so that $\|q_n\|_Q \to 0$ since S is a bounded bijective operator. Since $\lambda \neq 1$ and recalling that $(1-\lambda)v_n + A^{-1}B^*(q_n) \to 0$ in V, we conclude that $\|v_n\|_V \to 0$, providing the expected contradiction with $\|x_n\|_X = 1$. Hence, $\sigma(T) \subset \{-\varrho^{-1}, 1, \varrho\}$. Finally, we observe that $\lambda = 1$ is an eigenvalue associated with the eigenvectors $(v,0)^{\mathsf{T}}$ for all $v \in \ker(B)\setminus\{0\}$, and $\pm\varrho^{\pm 1}$ is an eigenvalue associated with the eigenvectors $(\pm\varrho^{\pm1}A^{-1}B^*(q), q)^{\mathsf{T}}$ for all $q \in Q\setminus\{0\}$. This proves (49.23). □

Theorem 49.6 allows us to derive sharp stability estimates for the solution of (49.18) in the weighted norm $\|(v,q)\|_{\widetilde{X}} := (a(v,v) + (S(q),q)_Q)^{\frac{1}{2}}$ induced by the weighted inner product for which T is symmetric. Equipping X with this weighted norm and since $\|T\|_{\mathcal{L}(X;X)} = \sup_{\lambda\in\sigma(T)}|\lambda|$ and $\|T^{-1}\|_{\mathcal{L}(X;X)} = \sup_{\lambda\in\sigma(T)}|\lambda|^{-1}$ (owing to Theorem 46.17), we infer from Theorem 49.6 that

$$\|T\|_{\mathcal{L}(X;X)} = \|T^{-1}\|_{\mathcal{L}(X;X)} = \varrho. \tag{49.24}$$

Recalling that $t((v,q),(w,r)) = (T(v,q),(w,r))_{\widetilde{X}}$, we infer that

$$\inf_{x\in X}\sup_{y\in Y}\frac{|t(x,y)|}{\|x\|_{\widetilde{X}}\|y\|_{\widetilde{X}}} = \|T^{-1}\|^{-1}_{\mathcal{L}(X;X)} = \varrho^{-1}, \tag{49.25a}$$

$$\sup_{x\in X}\sup_{y\in Y}\frac{|t(x,y)|}{\|x\|_{\widetilde{X}}\|y\|_{\widetilde{X}}} = \|T\|_{\mathcal{L}(X;X)} = \varrho. \tag{49.25b}$$

Corollary 49.7 (Stability). *Let* $(u,p) \in X$ *solve* (49.18). *The following holds true:*

$$\varrho^{-1}\left(\frac{1}{\|a\|}\|f\|^2_{V'} + \frac{\alpha}{\|b\|^2}\|g\|^2_{Q'}\right)^{\frac{1}{2}} \le \|(u,p)\|_{\widetilde{X}} \le \varrho\left(\frac{1}{\alpha}\|f\|^2_{V'} + \frac{\|a\|}{\beta^2}\|g\|^2_{Q'}\right)^{\frac{1}{2}}.$$

Proof. Let us set $x := (u,p)$. Then (49.18) amounts to $T(x) = y$ with $y := (A^{-1}(f), S^{-1}J_Q^{-1}(g))$. Hence, we have

$$\|T\|^{-1}_{\mathcal{L}(X;X)}\|y\|_{\widetilde{X}} \le \|x\|_{\widetilde{X}} \le \|T^{-1}\|_{\mathcal{L}(X;X)}\|y\|_{\widetilde{X}},$$

and we use (49.24) to infer that $\varrho^{-1}\|y\|_{\widetilde{X}} \le \|x\|_{\widetilde{X}} \le \varrho\|y\|_{\widetilde{X}}$. Finally, the bounds in the proof of Lemma 49.1 imply that $\|y\|^2_{\widetilde{X}} \ge \frac{1}{\|a\|}\|f\|^2_{V'} + \frac{\alpha}{\|b\|^2}\|g\|^2_{Q'}$ and $\|y\|^2_{\widetilde{X}} \le \frac{1}{\alpha}\|f\|^2_{V'} + \frac{\|a\|}{\beta^2}\|g\|^2_{Q'}$. □

Proposition 49.8 (Stability). *Let* $(u,p) \in X$ *solve* (49.18). *The following holds true:*

$$\frac{4}{((4\frac{\|b\|^2}{\alpha}+1)^{\frac{1}{2}}+1)^2}\left(\frac{1}{\|a\|}\|f\|^2_{V'} + \|g\|^2_{Q'}\right) \le a(u,u) + \|p\|^2_Q$$
$$\le \frac{4}{((4\frac{\beta^2}{\|a\|}+1)^{\frac{1}{2}}-1)^2}\left(\frac{1}{\alpha}\|f\|^2_{V'} + \|g\|^2_{Q'}\right). \tag{49.26}$$

Proof. The proof is similar to that of Corollary 49.7 but uses the operator $\widetilde{T} \in \mathcal{L}(X;X)$ s.t. $\widetilde{T}(v,q) := (v + A^{-1}B^*(q), J_Q^{-1}B(v))$ for all $(v,q) \in V\times Q$; see [42] and Exercise 49.4. □

49.3 Saddle point problems in Hilbert spaces

In this section, we assume again that V and Q are real Hilbert spaces and the bilinear form a is symmetric and coercive, i.e., A is self-adjoint and coercive. We show that the mixed problem (49.18) has a saddle point structure.

49.3.1 Finite-dimensional constrained minimization

We start by introducing some simple ideas in the finite-dimensional setting of linear algebra. Let N, M be two positive integers, let $\mathcal{A}$ be a symmetric positive definite matrix in $\mathbb{R}^{N\times N}$ and let $\mathsf{F} \in \mathbb{R}^N$. Consider the functional $\mathfrak{E} : \mathbb{R}^N \to \mathbb{R}$ such that $\mathfrak{E}(\mathsf{V}) := \frac{1}{2}(\mathcal{A}\mathsf{V}, \mathsf{V})_{\ell^2(\mathbb{R}^N)} - (\mathsf{F}, \mathsf{V})_{\ell^2(\mathbb{R}^N)}$. Then $\mathfrak{E}$ admits a unique global minimizer over $\mathbb{R}^N$, say U, which is characterized by the Euler condition $D\mathfrak{E}(\mathsf{U})(\mathsf{W}) = (\mathcal{A}\mathsf{U} - \mathsf{F}, \mathsf{W})_{\ell^2(\mathbb{R}^N)} = 0$ for all $\mathsf{W} \in \mathbb{R}^N$, i.e., $\mathcal{A}\mathsf{U} = \mathsf{F}$ (see Proposition 25.8 and Remark 26.5 for similar results expressed in terms of bilinear forms).

Let now $\mathcal{B} \in \mathbb{R}^{M\times N}$, let $\mathsf{G} \in \mathrm{im}(\mathcal{B}) \subset \mathbb{R}^M$, and consider the affine subspace $K := \{\mathsf{V} \in \mathbb{R}^N \mid \mathcal{B}\mathsf{V} = \mathsf{G}\}$. Then $\mathfrak{E}$ admits a unique global minimizer over K, say U, which is characterized by the Euler condition $D\mathfrak{E}(\mathsf{U})(\mathsf{W}) = (\mathcal{A}\mathsf{U} - F, \mathsf{W})_{\ell^2(\mathbb{R}^N)} = 0$ for all $\mathsf{W} \in \ker(\mathcal{B})$, i.e., $\mathcal{A}\mathsf{U} - \mathsf{F} \in \ker(\mathcal{B})^\perp$. Since $\ker(\mathcal{B})^\perp = \mathrm{im}(\mathcal{B}^\mathsf{T})$, we infer that there is $\mathsf{P} \in \mathbb{R}^M$ such that $\mathcal{A}\mathsf{U} + \mathcal{B}^\mathsf{T}\mathsf{P} = \mathsf{F}$. Recalling that $\mathcal{B}\mathsf{U} = \mathsf{G}$, the optimality condition is equivalent to solving the system

$$\begin{pmatrix} \mathcal{A} & \mathcal{B}^\mathsf{T} \\ \mathcal{B} & \mathbb{O} \end{pmatrix} \begin{pmatrix} \mathsf{U} \\ \mathsf{P} \end{pmatrix} = \begin{pmatrix} \mathsf{F} \\ \mathsf{G} \end{pmatrix}, \tag{49.27}$$

where $\mathbb{O}$ is the zero matrix in $\mathbb{R}^{M\times M}$. Moreover, P is unique if $\ker(\mathcal{B}^\mathsf{T}) = \{0\}$, i.e., if $\mathcal{B}$ has full row rank (note that this implies that $N \geq M$). This argument shows that the problem (49.27) (which is similar to (49.11)) is actually the optimality condition characterizing the minimizer of the functional $\mathsf{V} \mapsto \mathfrak{E}(\mathsf{V})$ under the constraint $\mathsf{V} \in K$.

Another way to look at the above problem consists of introducing the Lagrange multiplier associated with the constraint $\mathsf{V} \in K$, say Q, and considering the *Lagrangian* functional

$$\mathfrak{L}(\mathsf{V}, \mathsf{Q}) := \mathfrak{E}(\mathsf{V}) + (\mathsf{Q}, \mathcal{B}\mathsf{V} - \mathsf{G})_{\ell^2(\mathbb{R}^M)}. \tag{49.28}$$

Then the optimality conditions for a saddle point of $\mathfrak{L}$, say (U, P), are $D_\mathsf{V}\mathfrak{L}(\mathsf{U}, \mathsf{P})(\mathsf{W}) = (\mathcal{A}\mathsf{U} - \mathsf{F}, \mathsf{W})_{\ell^2(\mathbb{R}^N)} + (\mathcal{B}^\mathsf{T}\mathsf{P}, \mathsf{W})_{\ell^2(\mathbb{R}^M)} = 0$ for all $\mathsf{W} \in \mathbb{R}^N$, and $D_\mathsf{Q}\mathfrak{L}(\mathsf{U}, \mathsf{P})(\mathsf{R}) = (\mathsf{R}, \mathcal{B}\mathsf{U} - \mathsf{G})_{\ell^2(\mathbb{R}^M)} = 0$ for all $\mathsf{R} \in \mathbb{R}^M$, which again gives (49.27).

49.3.2 Lagrangian

Let us now reformulate in a more general framework the computations we have done in the previous section in finite dimension.

Definition 49.9 (Saddle point). *Let V and Q be two sets and consider a map $\mathcal{F} : V \times Q \to \mathbb{R}$. A pair $(u, p) \in V \times Q$ is said to be a* saddle point *of $\mathcal{F}$ if*

$$\forall q \in Q, \quad \mathcal{F}(u, q) \leq \mathcal{F}(u, p) \leq \mathcal{F}(v, p), \quad \forall v \in V. \tag{49.29}$$

Equivalently, we have $\sup_{q\in Q} \mathcal{F}(u, q) = \mathcal{F}(u, p) = \inf_{v\in V} \mathcal{F}(v, p)$.

Lemma 49.10 (Inf-sup). *The pair $(u,p) \in V\times Q$ is a saddle point of $\mathcal{F}$ iff*

$$\inf_{v\in V}\sup_{q\in Q}\mathcal{F}(v,q) = \mathcal{F}(u,p) = \sup_{q\in Q}\inf_{v\in V}\mathcal{F}(v,q). \tag{49.30}$$

Proof. According to Definition 49.9, $(u,p) \in V\times Q$ is a saddle point of $\mathcal{F}$ iff

$$\inf_{v\in V}\sup_{q\in Q}\mathcal{F}(v,q) \le \sup_{q\in Q}\mathcal{F}(u,q) = \mathcal{F}(u,p) = \inf_{v\in V}\mathcal{F}(v,p) \le \sup_{q\in Q}\inf_{v\in V}\mathcal{F}(v,q).$$

But independently of the existence of a saddle point, one can prove that

$$\sup_{q\in Q}\inf_{v\in V}\mathcal{F}(v,q) \le \inf_{v\in V}\sup_{q\in Q}\mathcal{F}(v,q). \tag{49.31}$$

Indeed, $\inf_{v\in V}\mathcal{F}(v,r) \le \mathcal{F}(w,r) \le \sup_{q\in Q}\mathcal{F}(w,q)$ for all $(w,r) \in V\times Q$. The assertion (49.31) follows by taking the supremum over $r \in Q$ on the left and the infimum over $w \in V$ on the right. Thus, the existence of a saddle point is equivalent to $\sup_{q\in Q}\inf_{v\in V}\mathcal{F}(v,q) = \inf_{v\in V}\sup_{q\in Q}\mathcal{F}(v,q)$. □

Proposition 49.11 (Lagrangian). *Let V and Q be two real Hilbert spaces. Let a be a bounded, symmetric, and coercive bilinear form on $V\times V$. Let b be a bounded bilinear form on $V\times Q$ satisfying* (49.19). *Let $f \in V'$ and $g \in Q'$. The following three statements are equivalent:* (i) *u minimizes the quadratic functional $\mathfrak{E}(v) := \frac{1}{2}a(v,v) - f(v)$ on the affine subspace $V_g := \{v \in V \mid b(v,q) = g(q), \forall q \in Q\}$.* (ii) *There is (a unique) $p \in Q$ such that the pair $(u,p) \in V\times Q$ is a saddle point of the Lagrangian $\mathcal{L}$ s.t.*

$$\mathcal{L}(v,q) := \frac{1}{2}a(v,v) + b(v,q) - f(v) - g(q). \tag{49.32}$$

(iii) *The pair (u,p) is the unique solution of* (49.18).

Proof. See Exercise 49.2. □

49.4 Babuška–Brezzi theorem

Let V and M be two real Banach spaces. Consider two bounded linear operators $A : V \to V'$ and $B : V \to M$, and the model problem

$$\begin{cases} \text{Find } u \in V \text{ and } p \in M' \text{ such that} \\ A(u) + B^*(p) = f, \\ B(u) = g, \end{cases} \tag{49.33}$$

where $B^* : M' \to V'$ is the adjoint of B, $f \in V'$, and $g \in M$. The goal of this section is to characterize the well-posedness of (49.33), reformulate it in terms of inf-sup conditions and bilinear forms associated with the operators A and

B, and relate this well-posedness result to the BNB theorem (Theorem 25.9). The theory exposed here is due to Babuška and Brezzi [34, 90]. (In the Hilbert setting considered in §49.2-§49.3, the spaces M and Q are related by $Q = M'$.)

49.4.1 Setting with Banach operators

Let $\ker(B)$ be the null space of B and let J_B be the canonical injection from $\ker(B)$ into V and $J_B^* : V' \to \ker(B)'$ be its adjoint. Let $A_\pi : \ker(B) \to \ker(B)'$ be such that $\langle A_\pi(v), w\rangle_{V',V} := \langle A(v), w\rangle_{V',V}$ for all $v, w \in \ker(B)$, i.e., $A_\pi := J_B^* A J_B$.

Theorem 49.12 (Well-posedness). *Problem* (49.33) *is well-posed if and only if A_π is an isomorphism and B is surjective.*

Proof. (1) Assume first that (49.33) is well-posed.
(1.a) Let $g \in M$ and let us denote by (u, p) the solution to (49.33) with data $(0, g)$. Since u satisfies $B(u) = g$, we infer that B is surjective.
(1.b) Let us show that A_π is surjective. Let $h \in \ker(B)'$. Owing to the Hahn–Banach theorem, there is an extension $\widetilde{h} \in V'$ s.t. $\langle \widetilde{h}, v\rangle_{V',V} = \langle h, v\rangle_{V',V}$ for all v in $\ker(B)$ and $\|\widetilde{h}\|_{V'} = \|h\|_{\ker(B)'}$. Let (u, p) be the solution to (49.33) with $f := \widetilde{h}$ and $g := 0$. Then $u \in \ker(B)$. Since $\langle B^*(p), v\rangle_{V',V} = \langle p, B(v)\rangle_{M',M} = 0$ for all $v \in \ker(B)$, we infer that $\langle A_\pi(u), v\rangle_{V',V} = \langle A(u), v\rangle_{V',V} = \langle \widetilde{h}, v\rangle_{V',V} = \langle h, v\rangle_{V',V}$ for all $v \in \ker(B)$. As a result, $A_\pi(u) = h$.
(1.c) Let us show that A_π is injective. Let $u \in \ker(B)$ be s.t. $A_\pi(u) = 0$. Then $\langle A(u), v\rangle_{V',V} = 0$ for all $v \in \ker(B)$, so that $A(u) \in \ker(B)^\perp$. B being surjective, $\mathrm{im}(B)$ is closed and owing to Banach's theorem (Theorem C.35), $\mathrm{im}(B^*) = \ker(B)^\perp$. As a result, $A(u) \in \mathrm{im}(B^*)$, i.e., there is $p \in M'$ such that $A(u) = -B^*(p)$. Hence, $A(u) + B^*(p) = 0$ and $B(u) = 0$, which shows that (u, p) solves (49.33) with $f := 0$ and $g := 0$. Uniqueness of the solution to (49.33) implies that $u = 0$.
(2) Conversely, assume that A_π is an isomorphism and B is surjective.
(2.a) For all $f \in V'$ and all $g \in M$, let us show that there is at least one solution to (49.33). The operator B being surjective, there is $u_g \in V$ s.t. $B(u_g) = g$. Denote by $h_{f,g}$ the bounded linear form on $\ker(B)$ s.t. $\langle h_{f,g}, v\rangle_{V',V} = \langle f, v\rangle_{V',V} - \langle A(u_g), v\rangle_{V',V}$ for all $v \in \ker(B)$. Since $A_\pi : \ker(B) \to \ker(B)'$ is an isomorphism, A_π is surjective, so that there is $\phi \in \ker(B)$ s.t. $A_\pi(\phi) = h_{f,g}$. Set $u := \phi + u_g$. The linear form $f - A(u)$ is in $\ker(B)^\perp$. Since B is surjective, $\ker(B)^\perp = \mathrm{im}(B^*)$, i.e., there is $p \in M'$ such that $B^*(p) = f - A(u)$. Moreover, $B(u) = B(\phi + u_g) = B(u_g) = g$. Hence, we have constructed a solution to (49.33).
(2.b) Let us show that the solution is unique. Let (u, p) be such that $B(u) = 0$ and $A(u) + B^*(p) = 0$, so that $u \in \ker(B)$ and $A_\pi(u) = 0$. Since πA is injective, $u = 0$. As a result, $B^*(p) = 0$. Since B is surjective, B^* is injective, which implies $p = 0$. □

49.4.2 Setting with bilinear forms and reflexive spaces

Let us now assume that V and M are reflexive Banach spaces and let us set $Q := M'$. Notice that this implies that $Q' = M'' = M$. Thus, we have $B \in \mathcal{L}(V;Q')$ and $B^* \in \mathcal{L}(Q;V')$. Consider the two bounded bilinear forms a and b defined, respectively, on $V \times V$ and on $V \times Q$ s.t. $a(v,w) := \langle A(v), w\rangle_{V',V}$ and $b(v,q) := \langle B(v), q\rangle_{Q',Q}$. Let us set

$$\|a\| := \sup_{v\in V}\sup_{w\in W} \frac{|a(v,w)|}{\|v\|_V\|w\|_V}, \qquad \|b\| := \sup_{v\in V}\sup_{q\in Q} \frac{|b(v,q)|}{\|v\|_V\|q\|_Q}. \tag{49.34}$$

Let $f \in V'$ and $g \in Q'$. With the shorthand notation $f(v) := \langle f, v\rangle_{V',V}$ and $g(q) := \langle g, q\rangle_{Q',Q}$, the abstract problem (49.33) is reformulated as follows:

$$\begin{cases} \text{Find } u \in V \text{ and } p \in Q \text{ such that} \\ a(u,w) + b(w,p) = f(w), \quad \forall w \in V, \\ b(u,q) = g(q), \quad \forall q \in Q. \end{cases} \tag{49.35}$$

Theorem 49.13 (Babuška–Brezzi). (49.35) *is well-posed if and only if*

$$\begin{cases} \displaystyle\inf_{v\in\ker(B)}\sup_{w\in\ker(B)} \frac{|a(v,w)|}{\|v\|_V\|w\|_V} =: \alpha > 0, \\ \forall w \in \ker(B), \quad [\,\forall v \in \ker(B),\, a(v,w) = 0\,] \implies [\,w = 0\,], \end{cases} \tag{49.36}$$

and the following inequality, usually called Babuška–Brezzi condition*, holds:*

$$\inf_{q\in Q}\sup_{v\in V} \frac{|b(v,q)|}{\|v\|_V\|q\|_Q} =: \beta > 0. \tag{49.37}$$

Furthermore, we have the following a priori estimates:

$$\|u\|_V \le c_1\,\|f\|_{V'} + c_2\,\|g\|_{Q'}, \tag{49.38a}$$

$$\|p\|_Q \le c_3\,\|f\|_{V'} + c_4\,\|g\|_{Q'}, \tag{49.38b}$$

with $c_1 := \frac{1}{\alpha}$, $c_2 := \frac{1}{\beta}(1 + \frac{\|a\|}{\alpha})$, $c_3 := \frac{1}{\beta}(1 + \frac{\|a\|}{\alpha})$, *and* $c_4 := \frac{\|a\|}{\beta^2}(1 + \frac{\|a\|}{\alpha})$.

Proof. (1) Since $\ker(B) \subset V$ is reflexive, we infer form Theorem C.49 that (49.36) is equivalent to A_π being an isomorphism. Furthermore, (49.37) is equivalent to B being surjective owing to (C.17) in Lemma C.40 since Q is reflexive. We invoke Theorem 49.12 to conclude that (49.35) is well-posed iff (49.36)-(49.37) hold true.

(2) Let us now prove the a priori estimates (49.38). From the condition (49.37) and Lemma C.42 (since Q is reflexive), we deduce that there exists $u_g \in V$ such that $B(u_g) = g$ and $\beta\|u_g\|_V \le \|g\|_{Q'}$. Setting $\phi := u - u_g \in \ker(B)$ yields $a(\phi, v) = f(v) - a(u_g, v)$ for all $v \in \ker(B)$. Since

$$|a(\phi,v)| \le (\|f\|_{V'} + \|a\|\,\|u_g\|_V)\|v\|_V \le \left(\|f\|_{V'} + \frac{\|a\|}{\beta}\|g\|_{Q'}\right)\|v\|_V,$$

taking the supremum over v in $\ker(B)$ yields $\alpha\|\phi\|_V \le \|f\|_{V'} + \frac{\|a\|}{\beta}\|g\|_{Q'}$ owing to (49.36). The estimate on u results from this inequality and the triangle inequality. To prove the estimate on p, we deduce from (49.37) and Lemma C.40 that $\beta\|p\|_Q \le \|B^*(p)\|_{V'}$, yielding $\beta\|p\|_Q \le \|a\|\|u\|_V + \|f\|_{V'}$. The estimate on $\|p\|_Q$ then results from that on $\|u\|_V$. □

Remark 49.14 (Coercivity). The conditions in (49.36) are automatically fulfilled if the bilinear form a is coercive on $\ker(B)$ or coercive on X. □

Let us recall that we have adopted the convention that suprema and infima in expressions like (49.34)-(49.36)-(49.37) are taken over nonzero arguments. To relate the conditions (49.36) and (49.37) with the conditions (BNB1) and (BNB2) from the BNB theorem (Theorem 25.9), let us introduce the space $X := V\times Q$ equipped with the norm $\|(v,q)\|_X := \|v\|_V + \|q\|_Q$ and let us recall from (49.20) the bounded bilinear form t on $X\times X$ defined by

$$t((v,q),(w,r)) := a(v,w) + b(w,q) + b(v,r). \tag{49.39}$$

Theorem 49.15 (Link with BNB). *The bilinear form t satisfies the conditions* (BNB1) *and* (BNB2) *if and only if* (49.36) *and* (49.37) *are satisfied.*

Proof. (1) Let us prove that (49.36) and (49.37) imply (BNB1) and (BNB2).
(1a) Proof of (BNB1). Let $(v,q) \in X$ and set $\mathbb{S} := \sup_{(w,r)\in X} \frac{|t((v,q),(w,r))|}{\|(w,r)\|_X}$. Lemma C.42 implies that there exists $\widehat{v} \in V$ such that $B(\widehat{v}) = B(v)$ and $\beta\|\widehat{v}\|_V \le \|B(v)\|_{Q'}$. We infer that

$$\beta\|\widehat{v}\|_V \le \|B(v)\|_{Q'} = \sup_{r\in Q}\frac{|b(v,r)|}{\|r\|_Q} = \sup_{(0,r)\in X}\frac{|t((v,q),(0,r))|}{\|(0,r)\|_X} \le \mathbb{S}.$$

Observing that $v - \widehat{v} \in \ker(B)$ we also infer that

$$\begin{aligned}
\alpha\|v-\widehat{v}\|_V &\le \sup_{w\in\ker(B)}\frac{|a(v-\widehat{v},w)|}{\|w\|_V} \\
&= \sup_{w\in\ker(B)}\frac{|a(v-\widehat{v},w)+b(w,q)+b(v,0)|}{\|w\|_V} \\
&\le \sup_{(w,0)\in X}\frac{|t((v,q),(w,0))|}{\|(w,0)\|_X} + \sup_{w\in\ker(B)}\frac{|a(\widehat{v},w)|}{\|w\|_V} \\
&\le \mathbb{S} + \|a\|\,\|\widehat{v}\|_V \le \left(1+\frac{\|a\|}{\beta}\right)\mathbb{S}.
\end{aligned}$$

Using the triangle inequality yields $\|v\|_V \le \left(\frac{1}{\beta} + \frac{1}{\alpha}\left(1+\frac{\|a\|}{\beta}\right)\right)\mathbb{S}$. Then we proceed as follows to bound $\|q\|_Q$:

$$\beta\|q\|_Q \leq \sup_{w\in V}\frac{|b(w,q)|}{\|w\|_V} \leq \sup_{w\in V}\frac{|a(v,w)+b(w,q)+b(v,0)|}{\|(w,0)\|_X} + \sup_{w\in V}\frac{|a(v,w)|}{\|w\|_V}.$$

This estimate implies that $\beta\|q\|_Q \leq \mathbb{S} + \|a\|\|v\|_V$, and we conclude that

$$\|q\|_Q \leq \frac{1}{\beta}\left(1+\|a\|\left(\frac{1}{\beta}+\frac{1}{\alpha}\left(1+\frac{\|a\|}{\beta}\right)\right)\right)\mathbb{S}.$$

This proves (BNB1).
(1b) Let $(w,r)\in X$ be s.t. $t((v,q),(w,r))=0$ for all $(v,q)\in X$, i.e.,

$$a(v,w)+b(v,r)=0, \quad \forall v\in V, \tag{49.40a}$$
$$b(w,r)=0, \quad \forall r\in Q. \tag{49.40b}$$

Then (49.40b) implies that $w\in\ker(B)$, and taking $v\in\ker(B)$ in (49.40a), we infer that $a(v,w)=0$, for all $v\in\ker(B)$. The second statement in (49.36) implies that $w=0$. Finally, (49.40a) yields $b(v,r)=0$ for all $v\in V$, and (49.37) implies that $r=0$. This proves (BNB2).
(2) Let us now prove that the conditions (BNB1) and (BNB2) on the bilinear form t imply the conditions (49.36) and (49.37) on the bilinear forms a and b. Let γ denote the inf-sup constant of the bilinear form t on $X\times X$.
(2a) Let us start with (49.37). For all $q\in Q$, we have

$$\begin{aligned}\gamma\|q\|_Q = \gamma\|(0,q)\|_X &\leq \sup_{(w,r)\in X}\frac{|t((0,q),(w,r))|}{\|(w,r)\|_X} = \sup_{(w,r)\in X}\frac{|b(w,q)|}{\|(w,r)\|_X}\\ &= \sup_{w\in V}\sup_{r\in Q}\frac{|b(w,q)|}{\|(w,r)\|_X} = \sup_{w\in V}\frac{|b(w,q)|}{\|w\|_V},\end{aligned}$$

since the supremum over $r\in Q$ is reached for $r=0$. This proves (49.37) with $\beta\geq\gamma>0$.
(2b) Let us prove the first statement in (49.36). For all $w\in V$, we define $(w'_w,r'_w)\in X$ to be the solution to the adjoint problem $t((v,q),(w'_w,r'_w)) = a(v,w)$ for all $(v,q)\in X$. Owing to Lemma C.53, this problem is well-posed. Moreover, we have $w'_w\in\ker(B)$ and $\gamma\|w'_w\|_V\leq\|a\|\|w\|_V$. Let $v\in\ker(B)$. We have $a(v,w)=t((v,q),(w'_w,r'_w))=a(v,w'_w)$ for all $q\in Q$. We infer that

$$\begin{aligned}\gamma\,\|v\|_V = \gamma\,\|(v,0)\|_X &\leq \sup_{(w,r)\in X}\frac{|t((v,0),(w,r))|}{\|(w,r)\|_X} = \sup_{(w,r)\in X}\frac{|a(v,w)|}{\|(w,r)\|_X}\\ &= \sup_{w\in V}\frac{|a(v,w)|}{\|w\|_V} = \sup_{w\in V}\frac{|a(v,w'_w)|}{\|w\|_V} \leq \frac{\|a\|}{\gamma}\sup_{w\in V}\frac{|a(v,w'_w)|}{\|w'_w\|_V}.\end{aligned}$$

Since $w'_w\in\ker(B)$, this finally gives $\frac{\gamma^2}{\|a\|}\|v\|_V\leq\sup_{w\in\ker(B)}\frac{|a(v,w)|}{\|w\|_V}$, which is the first statement in (49.36) with $\alpha\geq\frac{\gamma^2}{\|a\|}>0$.
(2c) Let us now prove the second statement in (49.36). We first recall that

we have already established that (49.37) holds true. From Lemma C.40, we then infer that $\mathrm{im}(B^*)$ is closed in V'. Let $w \in \ker(B)$ be s.t. $a(v,w) = 0$ for all $v \in \ker(B)$. This implies that $A^*(w) \in \ker(B)^\perp = \overline{\mathrm{im}(B^*)} = \mathrm{im}(B^*)$. Then there is $r_w \in Q$ s.t. $B^*(r_w) = A^*(w)$. For all $(v,q) \in X$, we then have

$$t((v,q),(w,-r_w)) = a(v,w) + b(w,q) - b(v,r_w) = a(v,w) - b(v,r_w)$$
$$= \langle A^*(w) - B^*(r_w), v\rangle_{V',V} = 0.$$

The condition (BNB2) on t implies that $(w,-r_w) = 0$, so that $w = 0$. □

Remark 49.16 (Sharper estimate). Sharper estimates on the inf-sup stability constant of t have been derived in Corollary 49.7 and Proposition 49.8 under the assumption that the bilinear form a is symmetric and coercive. □

Remark 49.17 (Direct sums). Notice that the map $V \ni w \mapsto w'_w \in \ker(B)$ introduced in step (2b) of the proof of Theorem 49.15 implies that any $w \in V$ can be uniquely decomposed into $w = w'_w + (A^*)^{-1}B^*r'_w$. This means that we have the direct decomposition $V = \ker(B) \oplus \mathrm{im}((A^*)^{-1}B^*)$. Note also that the same argument implies that $V = \ker(B) \oplus \mathrm{im}(A^{-1}B^*)$. □

Exercises

Exercise 49.1 (Algebraic setting). (i) Derive the counterpart of Theorem 49.12 in the setting of §49.3.1. (*Hint*: assume that the matrix $\mathcal{B}$ has full row rank and consider a basis of $\ker(\mathcal{B})$.) (ii) What happens if the matrix $\mathcal{A}$ is symmetric positive definite?

Exercise 49.2 (Constrained minimization). The goal is to prove Proposition 49.11. (i) Prove that if u minimizes $\mathfrak{E}$ over V_g, there is (a unique) $p \in Q$ such that (u,p) solves (49.35). (*Hint*: proceed as in §49.3.1.) (ii) Prove that (u,p) solves (49.35) if and only if (u,p) is a saddle point of $\mathcal{L}$. (*Hint*: consider $\mathfrak{E}_p : V \to \mathbb{R}$ s.t. $\mathfrak{E}_p(v) := \mathcal{L}(v,p)$ with fixed $p \in Q$.) (iii) Prove that if (u,p) is a saddle point of $\mathcal{L}$, then u minimizes $\mathfrak{E}$ over V_g. (iv) Application: minimize $\mathfrak{E}(v) := 2v_1^2 + 2v_2^2 - 6v_1 + v_2$ over $\mathbb{R}^2$ under the constraint $2v_1 + 3v_2 = -1$.

Exercise 49.3 (Symmetric operator). Let X be a Hilbert space and let $T \in \mathcal{L}(X;X)$ be a bijective symmetric operator. (i) Prove that T^{-1} is symmetric. (ii) Prove that $[\,\lambda \in \sigma(T)\,] \iff [\,\lambda^{-1} \in \sigma(T^{-1})\,]$. (*Hint*: use Corollary 46.18.) (iii) Prove that $\sigma(T) \subset \mathbb{R}$. (*Hint*: consider the sesquilinear form $t_\lambda(x,y) := ((T - \lambda I_X)(x), y)_X$ and use the Lax–Milgram lemma.)

Exercise 49.4 (Sharp stability). The goal is to prove Proposition 49.8. (i) Assume that $\ker(B)$ is nontrivial. Verify that $1 \in \sigma_{\mathrm{p}}(\widetilde{T})$. (ii) Let $\lambda \neq 1$ be in $\sigma(\widetilde{T})$. Prove that $\lambda(\lambda - 1) \in \sigma(S)$. (*Hint*: consider the sequence $x_n := (v_n, q_n)$ in X from Corollary 46.18, then observe that $(S(q_n), q_n)_Q = (1 -$

$\lambda)^2\langle A(v_n),v_n\rangle_{V',V}+\delta_n$, with $\delta_n:=\langle B^*(q_n)+(1-\lambda)A(v_n),A^{-1}B^*(q_n)-(1-\lambda)v_n\rangle_{V',V}$, and prove that $S(q_n)-\lambda(\lambda-1)q_n\to 0$ and $\liminf_{n\to\infty}\|q_n\|_Q>0$.) (iii) Prove that $\sigma(\widetilde{T})\subset[\lambda_\sharp^-,\lambda_\flat^-]\cup\{1\}\cup[\lambda_\flat^+,\lambda_\sharp^+]$ with $\lambda_\flat^\pm=\frac{1}{2}(1\pm(4\frac{\beta^2}{\|a\|}+1)^{\frac{1}{2}})$, and $\lambda_\sharp^\pm=\frac{1}{2}(1\pm(4\frac{\|b\|^2}{\alpha}+1)^{\frac{1}{2}})$. (*Hint*: use Lemma 49.1.) (iv) Conclude. (*Hint*: $\widetilde{T}$ is symmetric with respect to the weighted inner product $(x,y)_{\widetilde{X}}:=a(v,w)+(q,r)_Q$.)

Exercise 49.5 (Abstract Helmholtz decomposition). Consider the setting of §49.2 and equip V with the bilinear form a as inner product. (i) Prove that $\mathrm{im}(A^{-1}B^*)$ is closed and that $V=\ker(B)\oplus\mathrm{im}(A^{-1}B^*)$, the sum being a-orthogonal. (*Hint*: use Lemma C.39.) (ii) Let $f\in\ker(B)^\perp$. Prove that solving $b(v,p)=f(v)$ for all $v\in V$ is equivalent to solving $(S(p),q)_Q=(J_Q^{-1}BA^{-1}(f),q)_Q$ for all $q\in Q$.

Exercise 49.6 (Maxwell's equations). Consider the following problem: For $\boldsymbol{f}\in\boldsymbol{L}^2(D)$, find $\boldsymbol{A}$ and ϕ such that

$$\begin{cases}\nabla\times(\kappa\nabla\times\boldsymbol{A})+\nu\nabla\phi=\boldsymbol{f},\\ \nabla\cdot(\nu\boldsymbol{A})=0,\\ \boldsymbol{A}_{|\partial D_{\mathrm{d}}}\times\boldsymbol{n}=\boldsymbol{0},\ \phi_{|\partial D_{\mathrm{d}}}=0,\ (\kappa\nabla\times\boldsymbol{A})_{|\partial D_{\mathrm{n}}}\times\boldsymbol{n}=\boldsymbol{0},\ \boldsymbol{A}_{|\partial D_{\mathrm{n}}}\cdot\boldsymbol{n}=0,\end{cases}$$

where κ, ν are real and positive constants (for simplicity), and $|\partial D_{\mathrm{d}}|>0$ (see §49.1.3; here we write $\boldsymbol{A}$ in lieu of $\boldsymbol{H}$ and we consider mixed Dirichlet–Neumann conditions). (i) Give a mixed weak formulation of this problem. (*Hint*: use the spaces $\boldsymbol{V}_{\mathrm{d}}:=\{\boldsymbol{v}\in\boldsymbol{H}(\mathrm{curl};D)\mid\gamma^{\mathrm{c}}(\boldsymbol{v})_{|\partial D_{\mathrm{d}}}=\boldsymbol{0}\}$, where the meaning of the boundary condition is specified in §43.2.1, and $Q_{\mathrm{d}}:=\{q\in H^1(D)\mid\gamma^{\mathrm{g}}(q)_{|\partial D_{\mathrm{d}}}=0\}$.) (ii) Let $B:\boldsymbol{V}_{\mathrm{d}}\to Q_{\mathrm{d}}'$ be s.t. $\langle B(\boldsymbol{v}),q\rangle_{Q_{\mathrm{d}}',Q_{\mathrm{d}}}:=(\nu\boldsymbol{v},\nabla q)_{\boldsymbol{L}^2(D)}$. Let $\boldsymbol{v}\in\ker(B)$. Show that $\nabla\cdot\boldsymbol{v}=0$ and, if $\boldsymbol{v}\in\boldsymbol{H}^1(D)$, $\gamma^{\mathrm{g}}(\boldsymbol{v})_{|\partial D_{\mathrm{n}}}\cdot\boldsymbol{n}=0$. (*Hint*: recall that ν is constant.) (iii) Accept as a fact that D, ∂D_{d}, ∂D_{n} have topological and smoothness properties such that there exists $c>0$ s.t. $\ell_D\|\nabla\times\boldsymbol{v}\|_{\boldsymbol{L}^2(D)}\geq c\|\boldsymbol{v}\|_{\boldsymbol{L}^2(D)}$, for all $\boldsymbol{v}\in\ker(B)$, with $\ell_D:=\mathrm{diam}(D)$. Show that the above weak problem is well-posed. (*Hint*: use Theorem 49.13.) (iv) Let $(\mathcal{T}_h)_{h\in\mathcal{H}}$ be a shape-regular sequence of affine meshes. Let $k\geq 0$, let $\boldsymbol{V}_h:=\boldsymbol{P}_k^{\mathrm{c}}(\mathcal{T}_h)\cap\boldsymbol{V}_{\mathrm{d}}$, and let $Q_h:=P_{k+1}^{\mathrm{g}}(\mathcal{T}_h)\cap Q_{\mathrm{d}}$. Show that $\nabla Q_h\subset\boldsymbol{V}_h$. (v) Show that the discrete mixed problem is well-posed in $\boldsymbol{V}_h\times Q_h$ assuming that $\partial D_{\mathrm{d}}=\partial D$. (*Hint*: invoke Theorem 44.6.)