Chapter 2

The Babuška-Brezzi Theory

In this chapter we present the main results forming part of the Babuška-Brezzi theory, which allows to analyze a large family of mixed variational formulations and their respective Galerkin approximations. Our main references here include [16], [41], [50], and [52]. We begin by introducing the specific kind of operator equations which we are interested in.

2.1 The operator equation

Let $(H, \langle\cdot,\cdot\rangle_H)$ and $(Q, \langle\cdot,\cdot\rangle_Q)$ be real Hilbert spaces with induced norms $\|\cdot\|_H$ and $\|\cdot\|_Q$, respectively, and let $a : H \times H \to \mathbb{R}$ and $b : H \times Q \to \mathbb{R}$ be bounded bilinear forms. Then, given $F \in H'$ and $G \in Q'$, we are interested in the following problem: Find $(\sigma, u) \in H \times Q$ such that

$$\begin{aligned} a(\sigma,\tau) + b(\tau,u) &= F(\tau) \qquad \forall\, \tau \in H\,, \\ b(\sigma,v) &= G(v) \qquad \forall\, v \in Q\,. \end{aligned} \tag{2.1}$$

Next, let $\mathbf{A} : H \to H$ and $\mathbf{B} : H \to Q$ be the linear and bounded operators induced by a and b, respectively. Equivalently, according to the analysis in Section 1.1 (cf. (1.2), (1.3), and (1.4)), there holds

$$\mathbf{A} := \mathcal{R}_H \circ \mathcal{A} \quad \text{and} \quad \mathbf{B} := \mathcal{R}_Q \circ \mathcal{B}\,,$$

where $\mathcal{R}_H : H' \to H$ and $\mathcal{R}_Q : Q' \to Q$ are the respective Riesz mappings, and the operators $\mathcal{A} : H \to H'$ and $\mathcal{B} : H \to Q'$ are defined by:

$$\mathcal{A}(\sigma)(\tau) := a(\sigma,\tau) \qquad \forall\, \sigma \in H\,, \quad \forall\, \tau \in H\,,$$

and

$$\mathcal{B}(\tau)(v) := b(\tau,v) \qquad \forall\, \tau \in H\,, \quad \forall\, v \in Q\,.$$

It follows that

$$a(\sigma,\tau) = \langle \mathbf{A}(\sigma), \tau\rangle_H \qquad \forall\, (\sigma,\tau) \in H \times H \tag{2.2}$$

and

$$b(\tau,v) = \langle \mathbf{B}(\tau), v\rangle_Q = \langle \mathbf{B}^*(v), \tau\rangle_H \qquad \forall\, (\tau,v) \in H \times Q\,, \tag{2.3}$$

where $\mathbf{B}^* : Q \to H$ is the adjoint operator of $\mathbf{B}$.

In this way, (2.1) is rewritten, equivalently, as: Find $(\sigma,u) \in H \times Q$ such that

$$\begin{aligned} \langle \mathbf{A}(\sigma), \tau\rangle_H + \langle \mathbf{B}^*(u), \tau\rangle_H &= \langle \mathcal{R}_H(F), \tau\rangle_H \qquad \forall\, \tau \in H\,, \\ \langle \mathbf{B}(\sigma), v\rangle_Q &= \langle \mathcal{R}_Q(G), v\rangle_Q \qquad \forall\, v \in Q\,, \end{aligned}$$

or: Find $(\sigma,u) \in H \times Q$ such that

$$\begin{aligned} \mathbf{A}(\sigma) + \mathbf{B}^*(u) &= \mathcal{R}_H(F)\,, \\ \mathbf{B}(\sigma) &= \mathcal{R}_Q(G)\,, \end{aligned} \tag{2.4}$$

which, denoting the null operator by 0, reduces to the following matrix operator equation: Find $(\sigma,u) \in H \times Q$ such that

$$\begin{pmatrix} \mathbf{A} & \mathbf{B}^* \\ \mathbf{B} & 0 \end{pmatrix} \begin{pmatrix} \sigma \\ u \end{pmatrix} = \begin{pmatrix} \mathcal{R}_H(F) \\ \mathcal{R}_Q(G) \end{pmatrix}\,. \tag{2.5}$$

We now aim to provide necessary and sufficient conditions for (2.1) (equivalently (2.4) or (2.5)) to be well-posed.

2.2 The inf-sup condition

We recall first that this condition was already introduced in Section 1.1 (cf. (1.16), (1.17), (1.18), and (1.19)). Indeed, we say that the bounded bilinear form $b : H \times Q \to \mathbb{R}$

satisfies the continuous inf-sup condition if there exists a constant $\beta > 0$ such that :

$$\sup_{\substack{\tau \in H \\ \tau \neq 0}} \frac{b(\tau, v)}{\|\tau\|_H} \;\geq\; \beta \, \|v\|_Q \qquad \forall\, v \,\in\, Q\,. \tag{2.6}$$

Note, similarly as it was established for the pairs of conditions (1.16) - (1.18) and (1.17) - (1.19), that (2.6) is equivalent to

$$\inf_{\substack{v \in Q \\ v \neq 0}} \; \sup_{\substack{\tau \in H \\ \tau \neq 0}} \frac{b(\tau, v)}{\|\tau\|_H \, \|v\|_Q} \;\geq\; \beta\,,$$

which explains again the name INF-SUP. This hypothesis is also known as the LADYZHENSKAYA-BABUŠKA-BREZZI condition, or simply BABUŠKA-BREZZI condition. In addition, utilizing the adjoint operator $\mathbf{B}^*$, we have

$$\sup_{\substack{\tau \in H \\ \tau \neq 0}} \frac{b(\tau, v)}{\|\tau\|_H} \;=\; \sup_{\substack{\tau \in H \\ \tau \neq 0}} \frac{\langle \mathbf{B}^*(v), \tau \rangle_H}{\|\tau\|_H} \;=\; \|\mathbf{B}^*(v)\|_H\,,$$

and therefore the condition (2.6) is written also as:

$$\|\mathbf{B}^*(v)\|_H \;\geq\; \beta \, \|v\|_Q \qquad \forall\, v \,\in\, Q\,. \tag{2.7}$$

Moreover, the following lemma establishes equivalent conditions for (2.6) (or (2.7)).

LEMA 2.1 *The following statements are equivalent:*

i) *There exists $\beta > 0$ such that*

$$\sup_{\substack{\tau \in H \\ \tau \neq 0}} \frac{b(\tau, v)}{\|\tau\|_H} \;\geq\; \beta \, \|v\|_Q \qquad \forall\, v \,\in\, Q\,.$$

ii) $\mathbf{B}^*$ *is an isomorphism (linear bijection) from Q into $N(\mathbf{B})^\perp$ and*

$$\|\mathbf{B}^*(v)\|_H \;\geq\; \beta \, \|v\|_Q \qquad \forall\, v \,\in\, Q\,.$$

iii) $\mathbf{B}$ *is an isomorphism (linear bijection) from $N(\mathbf{B})^\perp$ into Q and*

$$\|\mathbf{B}(\tau)\|_Q \;\geq\; \beta \, \|\tau\|_H \qquad \forall\, \tau \,\in\, N(\mathbf{B})^\perp\,. \tag{2.8}$$

iv) $\mathbf{B} : H \to Q$ *is surjective.*

Demostración.
i) $\Rightarrow$ ii): Suppose that there exists $\beta > 0$ such that (2.6) (equivalently, (2.7)) is satisfied. It follows from (2.7) that $N(\mathbf{B}^*) = \{0\}$ and $R(\mathbf{B}^*)$ is closed, which implies that $\mathbf{B}^*$ is injective and $R(\mathbf{B}^*) = N((\mathbf{B}^*)^*)^\perp = N(\mathbf{B})^\perp$. Thus, $\mathbf{B}^*$ is a linear bijection from Q into $N(\mathbf{B})^\perp$.
ii) $\Rightarrow$ iii): Suppose that $\mathbf{B}^*$ is a linear bijection from Q into $N(\mathbf{B})^\perp$ and that (2.7) is satisfied. It follows again from (2.7) that $N(\mathbf{B}^*) = \{0\}$ and $R(\mathbf{B}^*)$ is closed. The latter implies, in virtue of a known result from functional analysis, that $R(\mathbf{B})$ is also closed, and therefore $R(\mathbf{B}) = N(\mathbf{B}^*)^\perp = \{0\}^\perp = Q$. In this way, $\mathbf{B}$ is a linear bijection from $N(\mathbf{B})^\perp$ into Q. In addition, it is clear from (2.7) that $\|(\mathbf{B}^*)^{-1}\| \le \dfrac{1}{\beta}$, and hence

$$\|\mathbf{B}^{-1}\| = \|(\mathbf{B}^{-1})^*\| = \|(\mathbf{B}^*)^{-1}\| \le \frac{1}{\beta},$$

which yields (2.8).
iii) $\Rightarrow$ iv): It follows directly from the fact that $\mathbf{B} : N(\mathbf{B})^\perp \to Q$ is bijective (in particular, surjective), recalling also that $H = N(\mathbf{B}) \oplus N(\mathbf{B})^\perp$.
iv) $\Rightarrow$ i): Suppose now that $\mathbf{B} : H \to Q$ is surjective. Since $R(\mathbf{B}) = Q$ is obviously closed, we have that $R(\mathbf{B}^*)$ is closed as well. Furthermore, applying orthogonality to the identity $Q = R(\mathbf{B}) = N(\mathbf{B}^*)^\perp$ we obtain that $N(\mathbf{B}^*) = \{0\}$, which says that $\mathbf{B}^*$ is injective. Hence, the characterization result for operators with closed range implies the inequality (2.7), which is exactly i).

$\square$

2.3 The main result

The characterization of the inf-sup condition given by Lemma 2.1 is essential for the proof of the following theorem, which establishes sufficient conditions for (2.1) to be well-posed. In the statement of this theorem and in what follows throughout the rest of the section, we assume the same notations and definitions from the previous sections.

Teorema 2.1 *Let $V := N(\mathbf{B})$ and let $\Pi : H \to V$ be the orthogonal projection operator. Assume that:*

i) $\Pi\,\mathbf{A} : V \to V$ *is a bijection.*

ii) *The bilinear form* b *satisfies the inf-sup condition* (2.6) (*equivalently,* (2.7)).

Then, for each pair $(F, G) \in H' \times Q'$ *there exists a unique* $(\sigma, u) \in H \times Q$ *solution of* (2.1) (*equivalently* (2.4) *or* (2.5)). *Moreover, there exists a constant* $C > 0$*, which depends on* $\|\mathbf{A}\|$, $\|(\Pi\,\mathbf{A})^{-1}\|$ *and* β*, such that*

$$\|(\sigma, u)\|_{H\times Q} \;\leq\; C \left\{ \|F\|_{H'} \,+\, \|G\|_{Q'} \right\}. \tag{2.9}$$

Demostración. Since $\mathbf{B}$ is a bijection from $V^{\perp}$ into Q (which is consequence of (2.6) and Lemma 2.1, part iii)), we deduce that there exists a unique $\sigma_g \in V^{\perp}$ such that

$$\mathbf{B}(\sigma_g) \;=\; \mathcal{R}_Q(G)\,, \tag{2.10}$$

and according to (2.8), there holds

$$\|\sigma_g\|_H \;\leq\; \frac{1}{\beta}\,\|\mathbf{B}(\sigma_g)\|_Q \;=\; \frac{1}{\beta}\,\|G\|_{Q'}\,. \tag{2.11}$$

In turn, since $\Pi\,\mathbf{A} : V \to V$ is a bijection and $\Pi(\mathcal{R}_H(F) - \mathbf{A}(\sigma_g))$ belongs to V, there exists a unique $\sigma_0 \in V$ such that $\Pi\,\mathbf{A}(\sigma_0) = \Pi(\mathcal{R}_H(F) - \mathbf{A}(\sigma_g))$. In addition, the bounded inverse theorem guarantees the existence of $\tilde{C} := \|(\Pi\,\mathbf{A})^{-1}\|$ such that

$$\|\sigma_0\|_H \;\leq\; \tilde{C}\,\|\Pi(\mathcal{R}_H(F) - \mathbf{A}(\sigma_g))\|_H \;\leq\; \tilde{C}\,\|\mathcal{R}_H(F) - \mathbf{A}(\sigma_g)\|_H\,,$$

from which, using the bound for $\|\sigma_g\|_H$, we obtain that

$$\|\sigma_0\|_H \;\leq\; \tilde{C} \left\{ \|F\|_{H'} \,+\, \frac{1}{\beta}\,\|\mathbf{A}\|\,\|G\|_{Q'} \right\}. \tag{2.12}$$

Now, thanks to the orthogonality condition of the projector Π, it is easy to see that the identity $\Pi\,\mathbf{A}(\sigma_0) = \Pi(\mathcal{R}_H(F) - \mathbf{A}(\sigma_g))$ is equivalent to saying that the vector $\mathbf{A}(\sigma_0 + \sigma_g) - \mathcal{R}_H(F)$ belongs to $V^{\perp}$. Thus, it follows from Lemma 2.1, part ii), that there exists a unique $u \in Q$ such that

$$\mathbf{B}^*(u) \;=\; \mathcal{R}_H(F) \,-\, \mathbf{A}(\sigma_0 + \sigma_g) \tag{2.13}$$

and

$$\|u\|_Q \;\leq\; \frac{1}{\beta}\,\|\mathbf{B}^*(u)\|_H \;=\; \frac{1}{\beta}\,\|\mathcal{R}_H(F) \,-\, \mathbf{A}(\sigma_0 + \sigma_g)\|_H\,,$$

from where

$$\|u\|_Q \;\leq\; \frac{1}{\beta}\left\{ \|F\|_{H'} \,+\, \|\mathbf{A}\|\left(\|\sigma_0\|_H + \|\sigma_g\|_H\right) \right\}. \tag{2.14}$$

In this way, defining $\sigma := \sigma_0 + \sigma_g \in H$, and noting that $\mathbf{B}(\sigma_0) = 0$, we deduce from (2.10), (2.13), and the estimates (2.11), (2.12), and (2.14), that (σ, u) solves (2.4) and satisfies (2.9).

For the uniqueness, let $(\sigma, u) \in H \times Q$ be solution of the homogeneous problem:

$$\begin{aligned} \mathbf{A}(\sigma) + \mathbf{B}^*(u) &= 0\,, \\ \mathbf{B}(\sigma) &= 0\,. \end{aligned}$$

It is clear from the second equation that $\sigma \in V$, and then, applying the projector Π to the first one, and recalling that $\mathbf{B}^*(u) \in V^\perp$, we get $\Pi\,\mathbf{A}(\sigma) = 0$. Thus, since $\Pi\,\mathbf{A} : V \to V$ is a bijection, it follows that $\sigma = 0$, and then from the first equation we obtain that $\mathbf{B}^*(u) = 0$. Finally, since $\mathbf{B}^* : Q \to V^\perp$ is also a bijection, we conclude that $u = 0$.

$\square$

We show next that the conditions i) and ii) from Theorem 2.1 are also **necessary**. Indeed, we have the following result.

Teorema 2.2 *Let* $V := N(\mathbf{B})$ *and let* $\Pi : H \to V$ *be the orthogonal projection operator. Assume that for each pair* $(F, G) \in H' \times Q'$ *there exists a unique solution* $(\sigma, u) \in H \times Q$ *of* (2.1) *(equivalently* (2.4) *or* (2.5)*), which satisfies*

$$\|(\sigma, u)\|_{H\times Q} \;\le\; C \left\{ \|F\|_{H'} + \|G\|_{Q'} \right\},$$

with a constant $C > 0$ *independent of* F *and* G*. Then:*

i) $\Pi\,\mathbf{A} : V \to V$ *is a bijection.*

ii) *The bilinear form* b *satisfies the inf-sup condition* (2.6)*.*

Demostración. First we prove ii). For this purpose, in virtue of Lemma 2.1, it suffices to show that $\mathbf{B}$ is surjective. In fact, given $g \in Q$, we know from the hypotheses that there exists a unique pair $(\sigma_g, u_g) \in H \times Q$ such that

$$\begin{aligned} \mathbf{A}(\sigma_g) + \mathbf{B}^*(u_g) &= 0\,, \\ \mathbf{B}(\sigma_g) &= g\,, \end{aligned}$$

and it is clear that the second equation of this system confirms the surjectivity of $\mathbf{B}$. Hence, knowing that b satisfies the continuous inf-sup condition, we can use the equivalences given by Lemma 2.1 to show that $\Pi\,\mathbf{A} : V \to V$ is a bijection.

Indeed, given $f \in V$, we know also from the hypotheses that there exists a unique $(\sigma_f, u_f) \in H \times Q$ such that

$$\begin{aligned} \mathbf{A}(\sigma_f) \,+\, \mathbf{B}^*(u_f) &= f\,, \\ \mathbf{B}(\sigma_f) &= 0\,, \end{aligned}$$

which shows, according to the second equation, that $\sigma_f \in V$. Then, applying the orthogonal projector Π to the first equation, and using, from part ii) of Lemma 2.1, that $\mathbf{B}^*(u_f) \in V^\perp$, we obtain $\Pi\,\mathbf{A}(\sigma_f) = \Pi(f) = f$, thus proving that $\Pi\,\mathbf{A} : V \to V$ is surjective. In turn, let $\sigma_0 \in V$ be such that $\Pi\,\mathbf{A}(\sigma_0) = 0$. It follows that $\mathbf{A}(\sigma_0) \in V^\perp$, and since, according to part ii) of Lemma 2.1, $\mathbf{B}^* : Q \to V^\perp$ is a bijection, we deduce that there exists a unique $u_0 \in Q$ such that $\mathbf{B}^*(u_0) = -\,\mathbf{A}(\sigma_0)$. In this way we have

$$\begin{aligned} \mathbf{A}(\sigma_0) \,+\, \mathbf{B}^*(u_0) &= 0\,, \\ \mathbf{B}(\sigma_0) &= 0\,, \end{aligned}$$

and thanks again to the hypotheses, we get $(\sigma_0, u_0) = (0, 0)$, which gives the injectivity of $\Pi\,\mathbf{A} : V \to V$.

$\square$

On the other hand, it is easy to see, according to Lemma 1.2, the inequalities (1.16) and (1.17), and the orthogonality characterizing the projector $\Pi : H \to V$, that the hypothesis i) in Theorems 2.1 and 2.2 is equivalent to each one of the following pairs of conditions:

i-1) there exists $\alpha > 0$ such that

$$\sup_{\substack{\tau \in V \\ \tau \neq 0}} \frac{a(\sigma, \tau)}{\|\tau\|_H} \;\geq\; \alpha\,\|\sigma\|_H \qquad \forall\,\sigma \in V\,,$$

i-2) for each $\tau \in V$, $\tau \neq 0$, there holds $\displaystyle\sup_{\sigma \in V} a(\sigma, \tau) > 0$.

and

i-1)' there exists $\alpha > 0$ such that

$$\sup_{\substack{\tau \in V \\ \tau \neq 0}} \frac{a(\tau,\sigma)}{\|\tau\|_H} \;\geq\; \alpha\, \|\sigma\|_H \qquad \forall\, \sigma \,\in\, V\,,$$

i-2)' for each $\tau \,\in\, V$, $\tau \neq 0$, there holds $\displaystyle\sup_{\sigma \in V} a(\tau,\sigma) \,>\, 0$.

More precisely, hypothesis i-1) (resp. i-1)') is an inf-sup condition for the bilinear form a, which is the same as requiring that the operator $\Pi\,\mathbf{A}$ (resp. $(\Pi\,\mathbf{A})^*$) be injective and with closed range, which, in addition, is equivalent to the surjectivity of the operator $(\Pi\,\mathbf{A})^*$ (resp. $\Pi\,\mathbf{A}$). In turn, i-2) (resp. i-2)') is equivalent to the injectivity of $(\Pi\,\mathbf{A})^*$ (resp. $\Pi\,\mathbf{A}$).

Certainly, when a is a symmetric bilinear form on $V \times V$, the operator $\Pi\,\mathbf{A}$ becomes self-adjoint, and in this case i-2) and i-2)' are redundant and therefore unnecessary.

Furthermore, it is important to remark that a sufficient condition (but not necessary) for i), which appears very often in applications, is the V-ellipticity of the bilinear form a, which means (cf. Definition 1.3) that there exists $\alpha > 0$ such that

$$a(\tau,\tau) \;\geq\; \alpha\, \|\tau\|_H^2 \qquad \forall\, \tau \,\in\, V\,. \tag{2.15}$$

In fact, the result that usually appears in the literature, even more frequently than Theorem 2.1, is the following.

Teorema 2.3 *Let* $V \,:=\, N(\mathbf{B})$ *and assume that:*

i) *The bilinear form* a *is* V*-elliptic* (*cf.* (2.15)).

ii) *The bilinear form* b *satisfies the inf-sup condition* (2.6) (*equivalently,* (2.7)).

Then, for each pair $(F,G) \,\in\, H' \times Q'$ *there exists a unique* $(\sigma,u) \in H \times Q$ *solution of* (2.1) (*equivalently* (2.4) *or* (2.5)). *Moreover, there exists a constant* $C > 0$*, which depends on* $\|\mathbf{A}\|$, α *and* β*, such that*

$$\|(\sigma,u)\|_{H\times Q} \;\leq\; C \left\{ \|F\|_{H'} \,+\, \|G\|_{Q'} \right\}.$$

Demostración. It suffices to see, for instance in virtue of the Lax-Milgram Lemma, that the V-ellipticity of a implies the hypothesis i) of Theorem 2.1.

□

Further extensions of the Babuška-Brezzi theory to other classes of linear and nonlinear abstract variational problems have been developed in several places (see, e.g. [13], [26], [31], [45]). In addition, interesting characterizations of the inf-sup condition for bilinear forms defined on product spaces can be found in [40] and also in [45].

2.4 Application Examples

In this section we illustrate the applicability of the Babuška-Brezzi theory with the classical examples given by the Poisson and elasticity problems.

2.4.1 The Poisson problem

Let Ω be a bounded domain of $\mathbb{R}^n, n \geq 2$, with Lipschitz-continuous boundary Γ. Then, given $f \in L^2(\Omega)$ and $g \in H^{1/2}(\Gamma)$, we consider the same problem introduced in Section 1.2.2, that is:

$$-\Delta u = f \quad \text{in} \quad \Omega, \quad u = g \quad \text{on} \quad \Gamma. \tag{2.16}$$

Then, as in that section, we introduce the additional unknown $\boldsymbol{\sigma} := \nabla u$ in Ω, so that problem (2.16) is rewritten as the first order system:

$$\boldsymbol{\sigma} = \nabla u \quad \text{in} \quad \Omega, \quad \operatorname{div} \boldsymbol{\sigma} = -f \quad \text{in} \quad \Omega, \quad u = g \quad \text{on} \quad \Gamma.$$

Then, multiplying the equation $\boldsymbol{\sigma} = \nabla u$ in Ω by $\boldsymbol{\tau} \in H(\operatorname{div};\Omega)$, and applying the Green identity (1.50) (cf. Lemma 1.4), we obtain:

$$\int_\Omega \boldsymbol{\sigma} \cdot \boldsymbol{\tau} = \int_\Omega \nabla u \cdot \boldsymbol{\tau} = -\int_\Omega u \operatorname{div} \boldsymbol{\tau} + \langle \gamma_{\boldsymbol{\nu}}(\boldsymbol{\tau}), \gamma_0(u) \rangle,$$

from which, using that the Dirichlet boundary condition says that $\gamma_0(u) = g$ on Γ, we deduce that

$$\int_\Omega \boldsymbol{\sigma} \cdot \boldsymbol{\tau} + \int_\Omega u \operatorname{div} \boldsymbol{\tau} = \langle \gamma_{\boldsymbol{\nu}}(\boldsymbol{\tau}), g \rangle \quad \forall\, \boldsymbol{\tau} \in H(\operatorname{div};\Omega). \tag{2.17}$$

Recall here that $\gamma_0 : H^1(\Omega) \to H^{1/2}(\Gamma)$ and $\gamma_{\boldsymbol{\nu}} : H(\operatorname{div};\Omega) \to H^{-1/2}(\Gamma)$ are the trace operators studied in Sections 1.3.1 and 1.3.4, and that $\langle \cdot, \cdot \rangle$ denotes the duality between $H^{-1/2}(\Gamma)$ and $H^{1/2}(\Gamma)$ with respect to the inner product of $L^2(\Gamma)$. Note also that the Green identity (1.50) and (2.17) justify the "*integration by parts*" employed in the deduction of (1.31).