Chapter 4
Weak formulation of elliptic problems

"Alan Turing is reported as saying that PDE’s are made by God, the boundary conditions by the Devil! The situation has changed, Devil has changed places... We can say that the main challenges are in the interfaces, with Devil not far away from them..."
Jacques-Louis Lions (1928-2001)

In this chapter, we consider linear elliptic problems that are commonly found in mechanical and physical partial differential models. The aim is to introduce the notion of weak formulation that give access to existence and uniqueness results for the solutions and that is well suited for the numerical approximation of such problems.

Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Dirichlet problem</td>
<td>53</td>
</tr>
<tr>
<td>4.2</td>
<td>Neumann problem</td>
<td>54</td>
</tr>
<tr>
<td>4.3</td>
<td>Abstract variational problems</td>
<td>55</td>
</tr>
<tr>
<td>4.4</td>
<td>Variational approximation of elliptic problems</td>
<td>66</td>
</tr>
<tr>
<td>4.5</td>
<td>Application to boundary-value problems in $\mathbb{R}$</td>
<td>71</td>
</tr>
<tr>
<td>4.6</td>
<td>Maximum principles</td>
<td>74</td>
</tr>
<tr>
<td>4.7</td>
<td>Exercises</td>
<td>75</td>
</tr>
</tbody>
</table>

4.1 Dirichlet problem

Let consider a bounded domain $\Omega \subset \mathbb{R}^d$ with a piecewise $C^1$ continuous boundary $\partial \Omega$ and let denote by $\nu$ the unit outer normal vector to $\Omega$. We first recall a Green’s identity.

**Lemma 4.1** For every function $u \in H^2(\Omega)$ and every function $v \in H^1(\Omega)$, we have the Green’s identity:

$$-\int_{\Omega} (\Delta u) v \, dx = \sum_{i=1}^{d} \int_{\Omega} \partial_{x_i} u \partial_{x_i} v \, dx - \int_{\partial \Omega} \frac{\partial u}{\partial \nu} v \, d\sigma.$$  \hspace{1cm} (4.1)
We consider the following model problem:

*Given a real-valued function \( f \in L^2(\Omega) \), find a function \( u \) defined in \( \Omega \) solving*

\[
-\Delta u = f \quad \text{in } \Omega \\
u = 0 \quad \text{on } \partial \Omega.
\] (4.2) (4.3)

Suppose the solution \( u \) is sufficiently smooth, for example \( u \in H^2(\Omega) \). Then, we can multiply each side of the equation (4.2) by a test function \( v \in H_0^1(\Omega) \) and integrate on \( \Omega \) to obtain:

\[-\int_{\Omega} (\Delta u)v \, dx = \int_{\Omega} fv \, dx.\]

Thanks to Green’s identity (4.1) and by noticing that \( v|_{\partial \Omega} = 0 \), we obtain:

\[\forall v \in H_0^1(\Omega), \quad \sum_{i=1}^{d} \int_{\Omega} \frac{\partial u}{\partial x_i} v \, dx = \int_{\Omega} fv \, dx. \] (4.4)

The specific boundary conditions and the Theorem 2.21 yield to conclude that the solution \( u \) of the problem (4.4) is such that:

\[u \in H_0^1(\Omega)\]

and the initial problem is now replaced by the following problem:

*Given a function \( f \in L^2(\Omega) \), find a function \( u \in H_0^1(\Omega) \) such that identity (4.4) holds.*

This is the weak formulation of the Dirichlet problem (4.2).

We have seen that any solution \( u \) of the Dirichlet problem (4.2), sufficiently smooth, is solution of the problem (4.4). Conversely, a solution \( u \in H_0^1(\Omega) \) is a solution of the problem (4.4) if and only if:

\[\forall \varphi \in D(\Omega), \quad \sum_{i=1}^{d} \int_{\Omega} \frac{\partial u}{\partial x_i} \varphi \, dx = \int_{\Omega} f \varphi \, dx,\]

since \( H_0^1(\Omega) \) is the closure of \( D(\Omega) \) in \( H^1(\Omega) \) (cf. Definition 2.14). Consequently, if \( u \) satisfies the equation (4.4), then the initial equation \(-\Delta u = f\) is satisfied on \( \Omega \) in the distributional sense and since \( f \in L^2(\Omega) \), this equation is verified in \( L^2(\Omega) \), thus almost everywhere in \( \Omega \). Finally, \( u \in H_0^1(\Omega) \) leads to satisfy the boundary condition according to the trace theorem on the boundary \( \partial \Omega \).

### 4.2 Neumann problem

Let consider an open bounded domain \( \Omega \subset \mathbb{R}^d \) with a piecewise \( C^1 \) continuous boundary \( \partial \Omega \). We now focus on the following problem:

*Given a function \( f \in L^2(\Omega) \), find a function \( u \) defined on \( \Omega \) solving*

\[
-\Delta u + u = f \quad \text{in } \Omega \\
\frac{\partial u}{\partial \nu} = 0 \quad \text{on } \partial \Omega.
\] (4.5) (4.6)

Suppose the solution \( u \in H^2(\Omega) \). Then, following the same procedure as for the Dirichlet problem, we obtain:

\[\forall v \in H^1(\Omega), \quad \sum_{i=1}^{d} \int_{\Omega} \frac{\partial u}{\partial x_i} v \, dx + \int_{\Omega} uv \, dx = \int_{\Omega} fv \, dx. \] (4.7)
The original boundary-value problem is then replaced by the following:

\( \forall \varphi \in \mathcal{D}(\Omega), \quad \sum_{i=1}^{d} \int_{\Omega} \partial_{x_i} u \partial_{x_i} \varphi \, dx + \int_{\Omega} w \varphi \, dx = \int_{\Omega} f \varphi \, dx , \)

and since \( u \) and \( f \) are \( L^2(\Omega) \) functions, the initial equation \(-\Delta u + u = f\) is satisfied in the distributional sense, in \( L^2(\Omega) \). When \( u \in H^2(\Omega) \), we can show, using the Green’s formula, that equation 4.7 is equivalent to the initial equation and that:

\( \forall v \in H^1(\Omega), \quad \int_{\partial \Omega} \frac{\partial u}{\partial \nu} v \, d\sigma = 0 , \)

and, since the trace space on \( \Omega \) is dense in \( L^2(\Omega) \), we have:

\( \forall w \in L^2(\partial \Omega), \quad \int_{\partial \Omega} \frac{\partial u}{\partial \nu} w \, d\sigma = 0 , \)

which leads to the Neumann boundary condition in the \( L^2(\Omega) \) sense. Any smooth solution of equation (4.7) is thus solving the initial Neumann problem.

### 4.3 Abstract variational problems

The previous examples fit in a general abstract framework that is well suited for the resolution of a large class of boundary value elliptic problems. To this end, we consider:

(i) a Hilbert space \( V \) on \( \mathbb{R} \), endowed with a norm \( \| \cdot \| \),

(ii) a bilinear form \( a \) defined and continuous on \( V \times V \), i.e. such that there exists a constant \( M > 0 \) such that:

\( \forall u \in V, \forall v \in V, \quad a(u, v) \leq M \| u \| \| v \| , \)

(iii) a linear form \( l \) continuous on \( V \), i.e. an element \( l \) of the topological dual space \( V' = \mathcal{L}(V, \mathbb{R}) \) of \( V \) endowed with the dual norm

\( \| l \|_{V'} = \sup_{v \in V, v \neq 0} \frac{l(v)}{\| v \|} . \)

Then, we consider the general variational problem:

\[
(P) \quad \begin{cases} 
\text{Given } l \in V', \text{ find } u \in V \text{ solving the problem} \\
\forall v \in V, \quad a(u, v) = l(v) 
\end{cases}
\] (4.8)

As such, this problem is certainly not well-posed. We will give a sufficient condition on the bilinear form \( a \) to have a unique solution. We first introduce a definition.

**Definition 4.1** The problem \( (P) \) is said to be well-posed if there is a unique solution to this problem and if the following linear stability property is satisfied:

\( \exists C > 0, \forall l \in V', \quad \| u \| \leq C \| l \|_{V'} . \)

The objective is now to determine if and when a variational problem is well-posed. Two fundamental results will help us to solve this question.
4.3.1 The Lax-Milgram theorem

**Definition 4.2** The bilinear form \( a(\cdot, \cdot) \) is said to be \( V \)-elliptic if there exists a constant \( \alpha > 0 \) such that
\[
\forall v \in V, \quad a(v, v) \geq \alpha \|v\|^2.
\]

**Theorem 4.1 (Lax-Milgram lemma)** Let \( V \) be a Hilbert space, \( a \in \mathcal{L}(V \times V, \mathbb{R}) \) be a bilinear form, and \( l \in V' \) be a linear form continuous on \( V \). Then, under the condition that:

(c0) the bilinear form \( a \) is continuous and \( V \)-elliptic with constant \( \alpha \),

the problem \( \mathcal{P} \) is well-posed (i.e. has a unique solution) and we have the estimate:
\[
\forall l \in V', \quad \|u\|_V \leq \frac{1}{\alpha} \|l\|_{V'},
\]

(i.e. the mapping \( u \mapsto l \) is an isomorphism of \( V \) on \( V' \)).

**Proof.** Since the linear form \( l \) is continuous on \( V \), the Riesz representation theorem asserts the existence of a unique element \( w_l \in V \) such that
\[
\forall v \in V, \quad l(v) = (w_l, v).
\]
This relation defines a linear one-to-one isometric mapping \( l \mapsto w_l \) from \( V' \) on \( V \). Similarly, \( u \) being fixed in \( V \), the linear form \( v \mapsto a(u, v) \) is continuous on \( V \). It is thus an isomorphism. Invoking the representation theorem again, there exists a unique element \( w \in V \) such that:
\[
\forall v \in V, \quad a(u, v) = (w, v),
\]
and we pose \( w = Au \). We claim that \( A : V \rightarrow V \) is a bounded continuous linear operator. It follows that the problem \( (\mathcal{P}) \) is equivalent to finding \( u \in V \) solving
\[
Au = w_l.
\]
The existence and uniqueness are obtained by establishing that the linear operator \( A \) is a bijective mapping from \( V \) to \( V \). \( A \) being continuous, the inverse \( A^{-1} \) is continuous from \( V \) to \( V \).

(i) \( A \) is an injection, i.e. \( \text{Ker}(A) = \{0\} \). Actually, since \( a \) is \( V \)-elliptic, we have
\[
\forall v \in V, \quad \alpha \|v\|^2 \leq a(v, v) = (Av, v) \leq \|Av\| \|v\|,
\]
and thus,
\[
\forall v \in V, \quad \|Av\| \geq \alpha \|v\|.
\]
This shows that \( A \) is an injection, since:
\[
v \in \text{Ker}(A) \Leftrightarrow Av = 0 \Rightarrow (Av, v) = 0
\]
thus \( v = 0 \) and, consequently, \( \text{Ker}(A) = \{0\} \).

(ii) \( A \) is a surjection, i.e. \( \overline{\text{Im} A} = R(A) = V \). We need to show that \( R(A) \) is closed and dense in \( V \).

Let consider \( w \in \overline{R(A)} \) (the closure of \( R(A) \) in \( V \)) and let \( (Av_m)_{m \in \mathbb{N}} \) be a sequence of \( R(A) \) that converges to \( w \) in \( V \). We have:
\[
\|Av_m - Av_n\| \geq \alpha \|v_m - v_n\|
\]
such that \( (v_m) \) is a Cauchy sequence in the Hilbert space \( V \): it converges to an element \( v \in V \) and \( (Av_m) \) converges to \( Av \), since \( A \) is continuous. Hence, we have:
\[
w = Av \in R(A),
\]
and we can conclude that \( R(A) \) is closed in \( V \).

Now, let consider \( v_0 \in (R(A))^\perp \), we have
\[
\alpha \|v_0\|^2 \leq a(v_0, v_0) = (Av_0, v_0) = 0,
\]
anf thus \( v_0 = 0 \). This shows that \( (R(A))^\perp \) is reduced to \( \{0\} \) and thus \( A \) is a surjection.
4.3. Abstract variational problems

The linear operator $A$ is a one-to-one mapping from $V$ to $V$. From the $V$-ellipticity property, we deduce
\[ \forall v \in V, \quad \| A^{-1}v \| \leq \frac{1}{\alpha} \| v \|, \]
and hence that $A^{-1}$ is continuous. By noticing that:
\[ \| w \| = \sup_{v \in V, v \neq 0} \frac{\langle w, v \rangle}{\| v \|} = \sup_{v \in V, v \neq 0} \frac{l(v)}{\| v \|} = \| l \|, \]
we obtain that the solution $u$ to the problem $(P)$ is such that:
\[ \| u \| = \| A^{-1}w \| \leq \frac{1}{\alpha} \| w \| = \frac{1}{\alpha} \| l \|, \]
and the result follows. $\Box$

When the bilinear form $a(\cdot, \cdot)$ is symmetric and positive, the problem $(P)$ can be interpreted as a minimization problem.

**Theorem 4.2** Consider a Hilbert space $V$, $a \in \mathcal{L}(V \times V, \mathbb{R})$ and $l \in V'$. Suppose the bilinear form $a(\cdot, \cdot)$ is symmetric and positive:
\[ \forall u, v \in V, \quad a(u, v) = a(v, u) \quad \text{and} \quad \forall u \in V, \quad a(u, u) \geq 0. \]

Then, $u$ is solution of the problem $(P)$ if and only if $u$ minimizes on $V$ the functional
\[ \forall v \in V, \quad J(v) = \frac{1}{2} a(v, v) - l(v), \]
i.e. $J(u) = \inf_{v \in V} J(v)$.

**Proof.** Suppose $u$ is solution of the problem $(P)$. A direct calculation gives:
\[ \forall v \in V, \quad J(v) = J(u) + a(u, v - u) - l(v - u) + \frac{1}{2} a(v - u, v - u) \]
\[ = J(u) + \frac{1}{2} a(v - u, v - u) \]
\[ \geq J(u), \]
which attests that $u$ minimizes the function $J$ on $V$. Conversely, let consider $v \in V$, $\lambda \in [0, 1]$ and we pose $w = u + \lambda v$. A direct calculation gives:
\[ J(w) - J(u) = \lambda (a(u, v) - l(v)) + \frac{\lambda^2}{2} a(v, v) \geq 0. \]

Dividing by $\lambda$ and letting $\lambda \to 0$ yields
\[ \forall v \in V, \quad a(u, v) \geq l(v). \]
Substituting $v$ by $-v$ yields to the equality and thus to conclude that $u$ is solution of the problem $(P)$. $\Box$

**Remark 4.1** (i) When the bilinear form $a(\cdot, \cdot)$ is symmetric, the problem $(P)$ corresponds to the minimization of a quadratic functional on a Hilbert space $V$, which is the abstract formulation of numerous problems in calculus of variations. This explains why $(P)$ is called a variational problem.
(ii) When the bilinear form $a(\cdot, \cdot)$ is symmetric and $V$-elliptic, the Lax-Milgram theorem indicates that the optimization problem $\inf_{v \in V} J(v)$ has a unique solution. The $V$-ellipticity of $a$ can be seen as a property of strong convexity of the functional $J$.

(iii) In many applications, the functional $J$ corresponds to an energy (cf. the deformation of an elastic membrane).

**Example 4.1** We consider the Dirichlet problem (4.2). Its variational (weak) formulation is that of problem (P), if we pose

$$V = H^1_0(\Omega) \quad \text{endowed with} \quad \|v\|_1, \Omega = \|v\|_{1, \Omega},$$

$$a(u, v) = \sum_{i=1}^d \int \partial_{x_i} u \partial_{x_i} v \, dx,$$

$$l(v) = \int f v \, dx.$$

It is easy to see that the bilinear form $a(\cdot, \cdot)$ (resp. the linear form $l(\cdot)$) is continuous on $H^1_0(\Omega) \times H^1_0(\Omega)$ (resp. on $H^1_0(\Omega)$) since we have

$$a(u, v) \leq |u|_{1, \Omega} |v|_{1, \Omega} \leq \|u\|_{1, \Omega} \|v\|_{1, \Omega},$$

$$l(v) \leq \|f\|_{0, \Omega} \|v\|_{0, \Omega} \leq \|f\|_{0, \Omega} \|v\|_{1, \Omega},$$

where $| \cdot |_{1, \Omega}$ denotes the semi-norm on $H^1(\Omega)$:

$$|v|_{1, \Omega} = \left( \sum_{i=1}^d \|\partial_{x_i} v\|^2_{0, \Omega} \right)^{1/2}.$$

Furthermore, since $\Omega$ is an open bounded subset of $\mathbb{R}^d$, the bilinear form $a(\cdot, \cdot)$ is $H^1_0(\Omega)$-elliptic, as we have:

$$\forall v \in H^1_0(\Omega), \quad a(v, v) \geq \frac{1}{1 + C_2^2} \|v\|^2_{1, \Omega},$$

thanks to Poincaré’s inequality. According to Lax-Milgram theorem, there exists a unique function $u \in H^1_0(\Omega)$ such that

$$\forall v \in H^1_0(\Omega), \quad \sum_{i=1}^d \int \Omega \partial_{x_i} u \partial_{x_i} v \, dx = \int \Omega f v \, dx,$$

that is precisely the form (4.4).

**Remark 4.2** We could have considered the Dirichlet boundary-value problem in the following formulation: given $f \in H^{-1}(\Omega)$, find $u$ such that

$$-\Delta u = f \quad \text{in} \quad \Omega \quad \text{and} \quad u = 0 \quad \text{on} \quad \partial \Omega.$$

As the right-hand side is not smooth, we will not look for a classical solution ($u \in C^2(\Omega)$) for this problem. Since $\Delta u$ is in $H^{-1}(\Omega)$, it seems reasonable to look for $u \in H^1(\Omega)$ and thus in $H^1_0(\Omega)$ because of the boundary condition. Let consider $v \in H^1_0(\Omega)$. The density of $D(\Omega)$ in $H^1_0(\Omega)$ leads to write:

$$\forall v \in H^1_0(\Omega), \forall w \in H^1(\Omega), \quad \langle -\Delta w, v \rangle = \langle \nabla w, \nabla v \rangle_{L^2(\Omega)}.$$
4.3. Abstract variational problems

Hence, if \( u \) is a solution of the problem in \( H^1(\Omega) \), then \( u \) is also a solution of the weak formulation: find \( u \in H^1_0(\Omega) \) such that
\[
\forall v \in H^1_0(\Omega), \quad \langle \nabla u, \nabla v \rangle_{L^2(\Omega)} = \langle f, v \rangle.
\]
Conversely, suppose \( u \) is a solution of the problem above and consider \( v = \varphi \in \mathcal{D}(\Omega) \):
\[
\forall \varphi \in \mathcal{D}(\Omega), \quad \langle f, v \rangle = \langle \nabla u, \nabla \varphi \rangle_{L^2(\Omega)} = \langle \Delta u, \varphi \rangle,
\]
hence, in the distributional sense:
\[
-\Delta u = f,
\]
and since \( f \in H^{-1}(\Omega) \), this equality is posed in \( H^{-1}(\Omega) \). To solve this problem, we have to show that the weak formulation has a unique solution. This can be achieved by invoking Lax-Miilgram theorem with \( V = H^1_0(\Omega), V' = H^{-1}(\Omega) \), \( a(u, v) = \langle \nabla u, \nabla v \rangle_{L^2(\Omega)} \) and \( l(v) = \langle f, v \rangle \). \( H^1_0(\Omega) \) is a Hilbert space endowed with the norm \( \| \cdot \|_{1, \Omega} \) (equivalent to the norm \( \| \cdot \|_{1, \Omega} \thanks{Poincaré’s inequality}) and \( \langle \nabla u, \nabla v \rangle_{L^2(\Omega)} \) is the inner product associated with this norm. Hence, the hypothesis of Lax-Milgram theorem are satisfied for \( M = 1 \) and \( \alpha = 1 \), the problem has a unique solution and moreover,
\[
|u|_{1, \Omega} \leq \|f\|_{-1, \Omega}.
\]
(4.9)

If the function \( f \) has more regularity, the solution \( u \) will also have more regularity. For instance, if \( f \in L^2(\Omega) \), then \( u \in H^2(\Omega) \) since \( \Delta u \in L^2(\Omega) \) and all the second derivatives are in \( L^2(\Omega) \).

**Definition 4.3** An open set \( \Omega \) is said to be Lipschitz (or Lipschitz continuous) if it is bounded and if in the vicinity of any point of its boundary, the boundary can be locally parametrized by a Lipschitz function \( \varphi \), the domain being located on one side of the boundary.

An open set \( \Omega \) is said to be of class \( C^{m,1} \) if the Lipschitz function \( \varphi \) can be replaced by a \( C^m \)-continuous function, whose derivatives of order \( m \) are Lipschitz continuous.

**Theorem 4.3** 1. Suppose \( \Omega \) is of class \( C^{1,1} \) or \( \Omega \) is a convex polygon (polyhedron). Then, the operator \( -\Delta \) is an isomorphism of \( H^2(\Omega) \cap H^1_0(\Omega) \) on \( L^2(\Omega) \).

2. Suppose \( \Omega \) is an arbitrary polygon in \( \mathbb{R}^2 \). For each \( t \) such that \( 1 < t \leq 4/3 \), the operator \( -\Delta \) is an isomorphism of \( W^{2,t}(\Omega) \cap H^1_0(\Omega) \) on \( L^t(\Omega) \).

3. Suppose \( \Omega \) is a Lipschitz polyhedron in \( \mathbb{R}^3 \). The operator \( -\Delta \) is an isomorphism of \( H^{3/2}(\Omega) \cap H^1_0(\Omega) \) on \( L^{3/2}(\Omega) \).

Then, under the Lax-Milgram hypothesis, there exists a constant \( C \) such that
\[
\forall u \in H^2(\Omega) \cap H^1_0(\Omega), \quad \|u\|_{2, \Omega} \leq C \|\Delta u\|_{0, \Omega}.
\]

**Remark 4.3** This result is not true if \( \Omega \) is a polygon with a concave corner, i.e. there exists right-hand side terms \( f \) for which the solution of the Dirichlet problem is not in \( H^2(\Omega) \).

4.3.2 Non-homogeneous Dirichlet conditions

Now, we consider the non-homogeneous Dirichlet problem:

\( \text{Given } f \in H^{-1}(\Omega) \text{ and given } g \in H^{1/2}(\partial \Omega), \text{ find } u \in H^1(\Omega) \text{ such that:} \)
\[
-\Delta u = f \quad \text{in } \Omega, \\
u = g \quad \text{on } \partial \Omega.
\]
(4.10)
Notice that this problem has at most one solution. The aim is to retrieve a homogeneous problem that we can solve. The trace theory indicates that if $\partial\Omega$ is $C^1$-continuous and if $g \in H^{1/2}(\partial\Omega)$ then there exists a function $u_g \in H^1(\Omega)$ such that $\gamma_0(u_g) = g$. In other words, the trace of $u_g$ on $\partial\Omega$ is the function $g$. This function $u_g$ is called a Dirichlet lift of $g$ in $\Omega$. We pose $u_0 = u - u_g$ and our problem is equivalent to the homogeneous problem:

Given $f \in H^{-1}(\Omega)$ and $g \in H^{1/2}(\partial\Omega)$, find $u_0 \in H^1(\Omega)$ such that

$$
-\Delta u_0 = f + \Delta u_g \quad \text{in } \Omega \\
u_0 = 0 \quad \text{on } \partial\Omega.
$$

Notice first that if $u_g \in H^1(\Omega)$ then $\Delta u_g \in \mathcal{D}'(\Omega)$ but $\Delta u_g \notin L^2(\Omega)$. However, since $\Delta u_g \in H^{-1}(\Omega)$, this problem has a unique solution $u_0$. Hence, $u = u_0 + u_g$ is solution of the initial problem (4.10) that has at most one solution, then this solution is unique and is independent of the Dirichlet lift function $u_g$. Since,

$$
\|\Delta u_g\|_{-1,\Omega} \leq \|\nabla u_g\|_{0,\Omega} = |u_g|_{1,\Omega},
$$

the inequality (4.9) leads to write:

$$
|u_0|_{1,\Omega} \leq \|f\|_{-1,\Omega} + \|\Delta u_g\|_{-1,\Omega} \leq \|f\|_{-1,\Omega} + |u_g|_{1,\Omega}.
$$

Hence, for every lift $u_g$ of the function $g$,

$$
\|u\|_{1,\Omega} \leq \|u_0\|_{1,\Omega} + \|u_g\|_{1,\Omega} \leq (C^2 + 1)^{1/2}\|f\|_{-1,\Omega} + ((C^2 + 1)^{1/2} + 1)\|u_g\|_{1,\Omega},
$$

and then

$$
\|u\|_{1,\Omega} \leq (C^2 + 1)^{1/2}\|f\|_{-1,\Omega} + ((C^2 + 1)^{1/2} + 1) \inf_{\gamma_0v = g} \|v\|_{1,\Omega}
\leq (C^2 + 1)^{1/2}\|f\|_{-1,\Omega} + ((C^2 + 1)^{1/2} + 1)\|g\|_{H^{1/2}(\partial\Omega)}.
$$

Finally, we enounce several results that summarize and extend the previous remarks.

**Theorem 4.4 (interior regularity)** Let consider $\Omega \subset \mathbb{R}^d$ an open bounded domain such that the Dirichlet problem is well-posed. We note $u \in H^1_0(\Omega)$ the weak solution. Then, if $f \in H^m_{loc}(\Omega)$ for a certain $m \geq 0$, then $u \in H^{m+2}_{loc}(\Omega)$, where

$$
H^m_{loc}(\Omega) = \{ u \in \mathcal{D}'(\Omega), \varphi \in H^m(\mathbb{R}^d), \forall \varphi \in \mathcal{D}(\Omega) \}.
$$

**Theorem 4.5 (global regularity)** Let $u$ be the weak solution of the Dirichlet problem (4.10). If $\partial\Omega$ is bounded and $C^{m+2}$-continuous (or if $\Omega = \mathbb{R}^d_+$) and if $f \in H^m(\Omega)$ ($m \geq 0$), then $u \in H^{m+2}(\Omega)$. Furthermore, there exists a constant $C_m > 0$ (independent of $u$ and $f$) such that:

$$
\|u\|_{H^{m+2}(\Omega)} \leq C_m \|f\|_{H^m(\Omega)}.
$$

**Theorem 4.6** If $\Omega$ is $C^1$-continuous and bounded in one direction, the non-homogeneous Dirichlet problem (4.10) admits:
4.3. Abstract variational problems

1. A unique weak solution $u \in H^1(\Omega)$ if $f \in L^2(\Omega)$ and $g \in H^{1/2}(\partial \Omega)$. The solution $u$ is such that:

$$\|u\|_{H^1(\Omega)} \leq C (\|f\|_{L^2(\Omega)} + \|g\|_{H^{1/2}(\partial \Omega)}),$$

where the constant $C$ is independent from $u$, $f$ and $g$.

2. **interior regularity**: if $f \in H^m_{\text{loc}}(\Omega)$ then $u \in H^{m+2}_{\text{loc}}(\Omega)$ for every $m \in \mathbb{N}$.

3. **global regularity**: if $\partial \Omega$ is bounded and of class $C^{m+2}$ or if $\partial \Omega = \mathbb{R}^d_+$, $f \in H^m(\Omega)$ and $g \in H^{m+3/2}(\partial \Omega)$ ($m \in \mathbb{N}$) then $u \in H^{m+2}(\Omega)$. Furthermore, we have the following estimate:

$$\|u\|_{H^{m+2}(\Omega)} \leq C (\|f\|_{H^m(\Omega)} + \|g\|_{H^{m+3/2}(\partial \Omega)}),$$

where the constant $C$ is independent from $u$, $f$ and $g$.

**Proof.** The trace theory indicates that if $\partial \Omega$ is $C^1$-continuous and if $g \in H^{1/2}(\partial \Omega)$ then there exists a function $u_0 \in H^1(\Omega)$ such that $\gamma_0(u_0) = g$. We consider the problem (4.11) as being posed in $\mathcal{D}'(\Omega)$. Hence, we have:

$$\langle -\Delta u_0, \varphi \rangle_{\mathcal{D}' \times \mathcal{D}} = \langle f, \varphi \rangle_{\mathcal{D}' \times \mathcal{D}} + \langle \Delta u_g, \varphi \rangle_{\mathcal{D}' \times \mathcal{D}}, \quad \forall \varphi \in \mathcal{D}(\Omega).$$

By definition of the derivation of distributions, we have:

$$\sum_{i=1}^{d} \langle \partial_{x_i} u_0, \partial_{x_i} \varphi \rangle_{\mathcal{D}' \times \mathcal{D}} = \langle f, \varphi \rangle_{\mathcal{D}' \times \mathcal{D}} - \sum_{i=1}^{d} \langle \partial_{x_i} u_g, \partial_{x_i} \varphi \rangle_{\mathcal{D}' \times \mathcal{D}}, \quad \forall \varphi \in \mathcal{D}(\Omega).$$

All the terms being $L^2(\Omega)$ functions, we can rewrite the previous identity as:

$$\int_{\Omega} \nabla u_0 \cdot \nabla \varphi \, dx = \int_{\Omega} f \, dx - \int_{\Omega} \nabla u_g \cdot \nabla \varphi \, dx, \quad \forall \varphi \in \mathcal{D}(\Omega),$$

and the weak formulation of problem (4.11) becomes:

Find $u_0 \in H^1_0(\Omega)$ such that:

$$\int_{\Omega} \nabla u_0 \cdot \nabla w \, dx = \int_{\Omega} f \, dx - \int_{\Omega} \nabla u_g \cdot \nabla w \, dx, \quad \forall w \in H^1_0(\Omega).$$

In order to involve the Lax-Milgram theorem, we define the bilinear form $a(v, w) = \int_{\Omega} \nabla u_0 \cdot \nabla w \, dx$ and the linear form $l(w) = \int_{\Omega} f \, dx - \int_{\Omega} \nabla u_g \cdot \nabla w \, dx$. The linear form $l$ is continuous since we have the estimate:

$$|l(w)| \leq \|f\|_{L^2(\Omega)} \|w\|_{L^2(\Omega)} + \|
abla u_g\|_{L^2(\Omega)} \|
abla w\|_{L^2(\Omega)}.$$

If we pose $C = \max(\|f\|_{L^2(\Omega)}, \|
abla u_g\|_{L^2(\Omega)})$, we obtain:

$$|l(w)| \leq C \|w\|_{H^1(\Omega)}, \quad \forall w \in H^1_0(\Omega).$$

The Lax-Milgram theorem states that there exists a constant $C > 0$ such that:

$$\|u_0\|_{H^1(\Omega)} \leq C (\|f\|_{L^2(\Omega)} + \|
abla u_g\|_{L^2(\Omega)}).$$

Since $u = u_0 + u_g$, it is easy to obtain the same type of estimate for the function $u \in H^1(\Omega)$. The theorem 4.4 gives the second item. Finally, if $\Omega$ is $C^{m+2}$-continuous and if $g \in H^{m+3/2}(\partial \Omega)$, we can choose $u_g \in H^{m+2}(\Omega)$ such that $\|u_g\|_{H^{m+2}(\Omega)} \leq 1/2 \|g\|_{H^{m+3/2}(\Omega)}$. Then, $\Delta u_g \in L^2(\Omega)$ and theorem 4.5 can be invoked to conclude. \qed
Chapter 4. Weak formulation of elliptic problems

4.3.3 The Nečas theorem

Now, we introduce an important theoretical result due to the Czech mathematician Nečas (1929-).

**Theorem 4.7 (Nečas)** Let $V$ and $W$ be two Hilbert spaces. We consider the abstract problem

Given $l \in V'$, find $u \in W$ such that $a(u, v) = l(v)$, $\forall v \in V$, where $a$ is a bilinear continuous form on $W \times V$ and $f$ is a linear form on $V$. This problem is well-posed if and only if:

1. **(c1)** there exists a constant $\alpha > 0$ such that
   $$\inf_{w \in W} \sup_{v \in V} \frac{a(w, v)}{\|w\|_W \|v\|_V} \geq \alpha > 0,$$

2. **(c2)** we have:
   $$\forall v \in V, \quad (\forall w \in W, a(w, v) = 0) \Rightarrow (v = 0).$$

Under these assumptions, we have the stability estimate:

$$\forall l \in V', \quad \|u\|_W \leq \frac{1}{\alpha} \|l\|_{V'}. $$

**Proof.** We consider the operator $A \in \mathcal{L}(W, V')$, $u \in W \mapsto Au \in V'$, defined by $(Au, v')_V = a(u, v)$ for $v \in V$. The problem (P) can be written as: find $u \in W$ such that $Au = l$ in $V'$. Again, to ensure the existence and uniqueness of the solution, it is sufficient to show that $A$ is a one-to-one mapping. This can be established using the same approach than for the Lax-Milgram theorem. The stability inequality results from:

$$a\|u\|_W \leq \sup_{v \in V} \frac{a(u, v)}{\|v\|_V} \leq \sup_{v \in V} \frac{l(v)}{\|v\|_V} = \|l\|_{V'}. $$

\[\square\]

**Remark 4.4**

1. **The condition (c1) can be written in the following form:**

   $$\forall w \in W, \quad \alpha \|w\|_W \leq \sup_{v \in V} \frac{a(w, v)}{\|v\|_V}. $$

   This condition is equivalent to: $\text{Ker}(A) = \{0\}$ and $\text{Im}(A) = R(A)$ is closed.

2. **The condition (c2) is often found as follows:**

   $$\forall v \in V, \quad (v \neq 0) \Rightarrow \left(\sup_{w \in W} a(w, v) > 0\right). $$

   This condition is equivalent to: $\text{Ker}(A^\perp) = R(A)^\perp = \{0\}$.

3. **Both conditions (c1) and (c2) are optimal as they provide necessary and sufficient conditions for the problem (P) to be well-posed.**
4.3.4 Saddle-point problems

We turn to a different type of problem, called a mixed problem. Let $\Omega$ be a bounded open subset of $\mathbb{R}^d$, the Stokes problem reads: find $u, p$ in appropriate Hilbert spaces such that:

$$
\begin{align*}
-\Delta u + \nabla p &= f & \text{in } \Omega, \\
\text{div } u &= g & \text{in } \Omega, \\
\text{div } u &= 0 & \text{on } \partial \Omega.
\end{align*}
$$

(4.12)

where $f : \Omega \to \mathbb{R}^d$ and $g : \Omega \to \mathbb{R}$ are given functions and the unknowns $u$ and $p$ represent the velocity and the pressure in an (possibly incompressible) slow viscous flow confined in $\Omega$. The solution $(u, p)$ is a saddle point as will be explained hereafter. Notice that we have a constraint of zero velocity on the boundary: $\int_{\Omega} g = 0$ since $\int_{\Omega} \nabla \cdot u = \int_{\partial \Omega} u \cdot \nu = 0$, using the divergence theorem.

Suppose the solution $u, p$ is sufficiently smooth, we can take a test function $v$ sufficiently smooth with values in $\mathbb{R}^d$. Since the velocity vanishes on the boundary, we will consider a test function that vanishes on the boundary, just as we did for the homogeneous Dirichlet problem. We multiply each term of the first equation by $v$ and we integrate by parts on $\Omega$. We obtain successively the terms:

$$
- \int_{\Omega} v \cdot \Delta u = \sum_{i=1}^{d} \int_{\Omega} v_i \Delta u_i = \sum_{i=1}^{d} \int_{\Omega} \nabla u_i \cdot \nabla v_i
$$

$$
= \sum_{i,j=1}^{d} \int_{\Omega} \partial_{x_j} u_i \partial_{x_j} v_i = \int_{\Omega} \nabla u : \nabla v
$$

the term $\int_{\Omega} v \cdot \nabla p = - \int_{\Omega} p \text{ div } v$ and finally, the following equation:

$$
\int_{\Omega} \nabla u : \nabla v - \int_{\Omega} p \text{ div } v = \int_{\Omega} f \cdot v.
$$

(4.13)

In this equation, suppose $f \in L^2(\Omega)^d$, then all terms are meaningful if we consider functions $u \in H^1_0(\Omega)^d$ and $p \in L^2(\Omega)$. To test the second equation, we consider a sufficiently smooth test function $q$:

$$
\int_{\Omega} q \text{ div } u = \int_{\Omega} g q.
$$

Since $u \in H^1_0(\Omega)^d$, the left-hand side term is well defined if $q \in L^2(\Omega)$. However, since we notice that $\int_{\Omega} \text{ div } u = \int_{\Omega} g = 0$, It is not necessary to test this equation by the constants. Hence, we can restrict the test function space to the subspace of $L^2(\Omega)$ defined as:

$$
L^2_0(\Omega) = \{ q \in L^2(\Omega), \int_{\Omega} q = 0 \}.
$$

Moreover, if $p$ is a solution, $p + c$ is also a solution (where $c$ is a constant). To overcome this problem, we add a condition of zero mean pressure. The weak form of this problem reads:

Given $f \in L^2(\Omega)^d$ and $g \in L^2_0(\Omega)$, find $(u, p) \in H^1_0(\Omega)^d \times L^2_0(\Omega)$ such that:

$$
\begin{align*}
\int_{\Omega} \nabla u : \nabla v - \int_{\Omega} p \text{ div } v &= \int_{\Omega} f \cdot v, & \forall v \in H^1_0(\Omega)^d, \\
\int_{\Omega} q \text{ div } u &= \int_{\Omega} g q, & \forall q \in L^2_0(\Omega).
\end{align*}
$$

(4.14)
Chapter 4. Weak formulation of elliptic problems

Let $V$ and $W$ be two Hilbert spaces. We introduce a specific form of the problem:

Given $l \in V'$, find $u \in W$ such that: $a(u, v) = l(v)$, $\forall v \in V$

which leads under certain assumptions to a saddle-point problem. Let consider $f \in V'$, $g \in W'$ and two bilinear forms $a$, defined on $V \times V$, and $b$ defined on $W \times V$. The new abstract problem reads:

Given $f \in V'$, $g \in W'$, find $u \in V$ and $p \in W$ such that:

\[
\begin{aligned}
(S) \quad \left\{ \begin{array}{l}
a(u, v) + b(v, p) = f(v), \quad \forall v \in V, \\
b(u, q) = g(q), \quad \forall q \in W.
\end{array} \right.
\tag{4.15}
\end{aligned}
\]

Remark 4.5 It is easy to check that in the Stokes problem we considered, $V = H_0^1(\Omega)^d$, $W = L^2_0(\Omega)$, $a(u, v) = \int_\Omega \nabla u : \nabla v$, $b(v, p) = -\int_\Omega p \text{ div } v$, $f(v) = \int_\Omega f \cdot v$ and $g(q) = -\int_\Omega g \cdot q$.

The following result can be considered as the natural extension of the Nečas theorem for the abstract problem $(S)$.

Theorem 4.8 Let $V$ and $W$ be two Hilbert spaces, $f \in V'$, $g \in W'$, $a \in \mathcal{L}(V \times V, \mathbb{R})$ and $b \in \mathcal{L}(V \times W, \mathbb{R})$. The abstract problem $(S)$ is well-posed if and only if:

\[
\exists \alpha > 0, \quad \inf_{u \in \text{Ker}(b)} \sup_{v \in \text{Ker}(b)} \frac{a(u, v)}{\|u\|_V \|v\|_V} \geq \alpha,
\tag{4.16}
\]

\[
\forall v \in \text{Ker}(b), \quad (\forall u \in \text{Ker}(b), a(u, v) = 0) \Rightarrow (v = 0),
\]

where $\text{Ker}(b) = \{v \in V, \forall q \in W, b(v, q) = 0\}$ and

\[
\exists \beta > 0, \quad \inf_{q \in W} \sup_{v \in V} \frac{b(v, q)}{\|v\|_V \|q\|_W} \geq \beta.
\tag{4.17}
\]

Under these assumptions, we have the following estimates:

\[
\|u\|_V \leq c_1 \|f\|_{V'} + c_2 \|g\|_{W'},
\]

\[
\|p\|_W \leq c_3 \|f\|_{V'} + c_4 \|g\|_{W'}.
\tag{4.18}
\]

where the constants are given as: $c_1 = \frac{1}{\alpha}$, $c_2 = \frac{1}{\beta}(1 + \frac{\|a\|}{\alpha})$, $c_3 = \frac{1}{\beta}(1 + \frac{\|a\|}{\alpha})$ and $c_4 = \frac{\|a\|}{\beta^2}(1 + \frac{\|a\|}{\alpha})$.

Proof. We introduce the operators $A : X \to X$, $\langle Au, v \rangle_{X', X} = a(u, v)$ and $B : X \to M'$ (resp. $B^t : M \to X'$), $\langle Bu, q \rangle_{M', M} = b(v, q)$. Then, the problem (4.15) is equivalent to the problem of finding $u \in X$ and $p \in M'$ such that:

\[
\begin{aligned}
Au + B^t p &= f, \\
Bu &= g,
\end{aligned}
\]

and this problem is well-posed if and only if the two following conditions are satisfied (admitted here):

(i) $\text{Ker}(B) \to \text{Ker}(B')$ is an isomorphism,

(ii) $B : X \to M$ is surjective.

According to this result, we admit here that the problem is well-posed. There exists $u_g \in X$ such that $Bu_g = g$ and $\beta\|u_g\|_X \leq \|g\|_{M'}$. Introducing the variable $\phi = u - u_g$ leads to write:

\[
\forall v \in \text{Ker}(B), \quad a(\phi, v) = f(v) - a(u_g, v).
\]
By observing that:
\[ |f(v) - a(u, v)| \leq (\|f\|_{X'} + \|a\|_{\|u\|_{X}}) \|v\|_{X} \]
and by taking the supremum on \( v \in \text{Ker}(B) \), we deduce that:
\[ \alpha \|\phi\|_{X} \leq \|f\|_{X'} + \frac{\|a\|}{\beta} \|g\|_{M'}. \]

Using the triangle inequality \( \|u\|_{X} \leq \|u - u_{g}\|_{X} + \|u_{g}\|_{X} \) leads to the first inequality. The stability inequality on \( p \) results from \( \beta\|p\|_{M} \leq \|B^{1/2}p\|_{X'}, \) thus
\[ \beta\|p\|_{M} \leq \|a\| \|u\|_{X} + \|f\|_{X'}. \]

The results follows using the previous estimate on \( \|u\|_{X}. \)

\[\square\]

**Remark 4.6** If the bilinear form \( a(\cdot, \cdot) \) is \( V \)-elliptic then the conditions (4.16) are satisfied.

We noticed before that if the bilinear form \( a(\cdot, \cdot) \) is symmetric and positive, such abstract problem can be interpreted as a minimization problem. Here, if the bilinear form \( a(\cdot, \cdot) \) is symmetric positive, this minimization problem is transformed into a saddle-point problem.

**Definition 4.4** Given a linear mapping \( L : V \times W \to \mathbb{R} \), we say that \((u, p)\) is a saddle-point of \( L \) if:
\[ \forall (v, q) \in V \times W, \quad L(u, q) \leq L(u, p) \leq L(v, p). \]

**Proposition 4.1** \((u, p)\) is a saddle-point of the linear mapping \( L \) if and only if:
\[ \sup_{q \in W} L(u, q) = \inf_{v \in V} \sup_{q \in W} L(v, q) = L(u, p) = \sup_{q \in W} \inf_{v \in V} L(v, q) = \inf_{v \in V} L(v, p). \quad (4.19) \]

**Proof.** According to the previous definition, we have:
\[ \inf_{v \in V} \sup_{q \in W} L(v, q) \leq \sup_{q \in W} L(u, q) \leq L(u, p) \leq \inf_{v \in V} L(v, p) \leq \sup_{q \in W} \inf_{v \in V} L(v, q). \]
Moreover, for every \((v, q) \in V \times W\), we have
\[ \inf_{v' \in V} L(v', q) \leq L(v, q) \leq \sup_{q' \in W} L(v, q'), \]

and we deduce then
\[ \sup_{q \in W} \inf_{v \in V} L(v, q) \leq \inf_{v \in V} \sup_{q \in W} L(v, q). \]

In other words, we deduce
\[ \inf_{v \in V} \sup_{q \in W} L(v, q) = \sup_{q \in W} L(u, q) = L(u, p) = \inf_{v \in V} L(v, p) = \sup_{q \in W} \inf_{v \in V} L(v, q). \]

and the results follows. \(\square\)

**Theorem 4.9** Suppose the bilinear form \( a(\cdot, \cdot) \) is symmetric and positive. The pair \((u, p)\) is solution of the problem \((S)\) if and only if \((u, p)\) is a saddle-point of the functional
\[ L(v, q) = \frac{1}{2} a(v, v) + b(v, q) - f(v) - g(q). \quad (4.20) \]
4.4 Variational approximation of elliptic problems

This section is devoted to the approximation of an abstract problem using a numerical method.

4.4.1 Abstract theory

We return to the general situation for the elliptic problems and we suppose given

(i) a Hilbert space $V$ on $\mathbb{R}$, endowed with a norm denoted $\| \cdot \|,$

(ii) a bilinear form $a(\cdot, \cdot)$ continuous on $V \times V$ and $V$-elliptic, i.e. such that:

$$\forall u \in V, \forall v \in V, \quad a(u, v) \leq M \|u\| \|v\|$$

$$\forall v \in V, \quad a(v, v) \geq \alpha \|v\|^{2}.$$ 

(iii) a linear form $l(\cdot)$ continuous on $V$.

And, according to Lax-Milgram theorem, we know that the problem:

(\mathcal{P}) \quad \text{Find } u \in V \text{ such that } \forall v \in V, \quad a(u, v) = l(v),

has a unique solution. In this context, the Hilbert space $V$ is of infinite dimension; In order to obtain a numerical approximation of the solution $u$, we will consider replacing the problem (\mathcal{P}) by a "discrete problem" posed in a finite dimensional space, denoted $V_{h}$, where $h$ represents a discretization parameter intended to tend towards 0. For the sake of simplicity, we will only consider here a conforming approximation for which $V_{h} \subset V$.

Under these assumptions, the bilinear form $a(\cdot, \cdot)$ and the linear form $l(\cdot)$ are defined on $V_{h} \times V_{h}$ and $V_{h}$, respectively and we consider the discrete problem:

Find $u \in V$ such that:

\begin{equation}
(\mathcal{P}_{h}) \quad \forall v_{h} \in V_{h}, \quad a(u_{h}, v_{h}) = l(v_{h}).
\end{equation}

In this case, the discrete problem is called a Galerkin approximate problem, named after the Russian mathematician B.G. Galerkin (1871-1945). More generally, this approach is called a Galerkin method.

Under the three hypothesis (i) – (iii) given hereabove, we have the following result.
Theorem 4.10 (Lax-Milgram) The discrete problem \((P_h)\) defined by (4.21) has a unique solution \(u_h\) in \(V_h\) and we have the estimate:

\[
\forall l \in V', \quad \|u_h\|_V \leq \frac{1}{\alpha} \|l\|_{V'}.
\]

Proof. This result can be obtained by a direct application of the Lax-Milgram theorem, since \(V_h\) is a Hilbert space for the norm associated with \(V\).

Now, we analyze the error committed by replacing the solution \(u\) of the problem \((P)\) by the solution \(u_h\) of the problem \((P_h)\). The following result, due to the French mathematician Jean Céa (1933-), helps to understand why the variational approximation method is so interesting. It states that the discretization error of the problem \((P)\) is of the same order than the error committed by approximating \(V\) by \(V_h\).

Lemma 4.2 (Céa) Under the hypothesis of the Lax-Milgram theorem, we have the error estimate:

\[
\|u - u_h\| \leq \frac{M}{\alpha} \inf_{v_h \in V_h} \|u - v_h\|. \tag{4.22}
\]

Proof. Suppose \(v_h \in V_h\). We pose \(w_h = v_h - u_h\). Since \(V_h\) is a subspace of \(V\), \(w_h \in V_h\) and thus \(w_h \in V\). Hence, we have

\[
a(u - u_h, w_h) = 0,
\]

or

\[
a(u - u_h, u - u_h) = a(u - u_h, u - v_h).
\]

From the hypothesis \((i) - (iii)\), we deduce that

\[
\alpha \|u - u_h\|^2 \leq a(u - u_h, u - u_h) = a(u - u_h, u - v_h) \leq M \|u - u_h\| \|u - v_h\|,
\]

and the results follows.

Remark 4.7 When the bilinear form \(a(\cdot, \cdot)\) is symmetric, we have, for all \(v_h \in V_h\):

\[
a(u - v_h, u - v_h) = a(u - u_h, u - u_h) + a(u_h - v_h, u_h - v_h)
\]

and thus

\[
a(u - u_h, u - u_h) = \inf_{v_h \in V_h} a(u - v_h, u - v_h).
\]

and then, the constant in the error estimate is replaced by \(\sqrt{\frac{M}{\alpha}}\).

The theorem shows that the evaluation of the error is equivalent to the evaluation of the quantity \(\inf_{v_h \in V_h} \|u - v_h\|\), in other words, it consists in evaluating the distance in \(V\) between the solution \(u\) of the problem \((P)\) and the subspace \(V_h\) of \(V\).

Theorem 4.11 Suppose that there exists a subspace \(V \subset V\), dense in \(V\) and a linear mapping \(r_h : V \to V_h\) such that:

\[
\forall v \in V, \quad \lim_{h \to 0} \|v - r_h(v)\| = 0.
\]

Then, the approximation method converges, i.e.

\[
\lim_{h \to 0} \|u - u_h\| = 0.
\]
Chapter 4. Weak formulation of elliptic problems

Proof. Let $\varepsilon > 0$, since $V$ is dense in $V$, there exists an element $v \in V$ such that
\[ \|u - v\| \leq \frac{\varepsilon}{2C}, \quad C = \frac{M}{\alpha}. \]

For such element $v$, there exists $h(\varepsilon) > 0$ such that
\[ \forall h \leq h(\varepsilon), \quad \|u - r_h(v)\| \leq \frac{\varepsilon}{2C}. \]

For a sufficiently small $h \leq h(\varepsilon)$, we have
\[ \|u - u_h\| \leq C(\|u - v\| + \|v - r_h(v)\|) \leq \varepsilon. \]

and the result follows. \hfill \square

When the variational approximation converges, it is interesting to study its rate of convergence, when $h \to 0$. In general, we will look for a bound on the approximation error of the form
\[ \|u - u_h\| \leq C(u)h^k, \quad k > 0, \]
where $C(u)$ is a constant independent of $h$ and $k > 0$ is a constant. In such a case, the approximation is said to be of order $k$ and we write:
\[ \|u - u_h\| = O(h^k). \]

The approximation spaces $V_h$ of $V$ must be carefully chosen, and several criteria can be used to determine the appropriate spaces.

(i) $V_h$ must have a basis $(\varphi_1, \ldots, \varphi_N)$ such that the coefficients $a(\varphi_i, \varphi_j)$ and $l(\varphi_i)$ are easy to compute numerically, and such that the resulting linear system is not too difficult to solve with numerical algorithms,

(ii) $V_h$ must be chosen so that the approximation method converges and so that the numerical solution is an accurate approximation of the exact solution $u$.

Unfortunately, the requirements are often contradictory and a good compromise must be sought.

4.4.2 Linear system

The problem $(\mathcal{P}_h)$ is a linear system to solve. If $N = \dim V_h$, suppose $(\varphi_1, \ldots, \varphi_N)$ is a basis of $V_h$. The stiffness matrix $A \in \mathbb{R}^{N,N}$ of the system is the matrix of coefficients:
\[ a_{ij} = a(\varphi_j, \varphi_i), \quad 1 \leq i, j \leq N. \]

The solution $u_h$ can be decomposed in the basis of $V_h$ as:
\[ u_h = \sum_{i=1}^{N} u_i \varphi_i, \]

with the vector $U = (u_i)_{1 \leq i \leq N} \in \mathbb{R}^N$. Denoting the right-hand side term of equation (4.21) $F = (f_i) \in \mathbb{R}^N$, where
\[ f_i = l(\varphi_i) 1 \leq i \leq N, \]
leads us to write the problem $(\mathcal{P}_h)$ in the matrix form as:
\[ AU = F. \]

In the discrete forms of the Lax-Milgram and Nečas theorems, we face the following conditions:
4.4. Variational approximation of elliptic problems

(c0h) (Lax-Milgram) there exists a constant $\alpha_h > 0$ such that
\[ \forall u_h \in V_h, \quad a(u_h, v_h) \geq \alpha_h \| u_h \|_V^2. \]

(c1h) (Nečas) there exists a constant $\alpha_h > 0$ such that
\[ \inf_{w_h \in W_h} \sup_{v_h \in V_h} \frac{a(w_h, v_h)}{\| w_h \|_W \| v_h \|_V} \geq \alpha_h, \]

(c2h) (Nečas) we have:
\[ \forall v_h \in V_h, \quad (\sup_{w_h \in W_h} a(w_h, v_h) = 0) \Rightarrow (v_h = 0). \]

Proposition 4.2 We have the following properties:

1. (c0h) $\Leftrightarrow$ $A$ is positive definite,

2. (c1h) $\Leftrightarrow$ $\text{Ker}(A) = \{0\}$,

3. (c2h) $\Leftrightarrow$ $\text{rank} A = \dim V_h$.

Proof.

1. Suppose (c0h) holds. A direct calculation gives
\[ \forall X \in \mathbb{R}^N, \quad \langle AX, X \rangle_N = a(\Phi, \Phi) \geq \alpha \| \Phi \|_V^2 \geq c \| X \|_N^2, \]
where $\langle \cdot, \cdot \rangle_N$ denotes the Euclidean scalar product in $\mathbb{R}^N$, $\| \cdot \|_N$ the associated norm and for $X \in \mathbb{R}^N$, $X = (X_i)_{1 \leq i \leq N}$, $\Phi = \sum_{i=1}^N X_i \varphi_i \in V_h$. This result shows that $A$ is positive definite. Conversely, if the matrix $A$ is positive definite, for every $X \in \mathbb{R}^N$, with $\| X \|_N = 1$, we have $\langle AX, X \rangle_N > 0$, hence, we have
\[ \forall \Phi \in V_h, \Phi \neq 0, \quad a(\Phi, \Phi) = \langle AX, X \rangle_N = \left( A \frac{X}{\| X \|_N} \right) \| X \|_N^2 \geq c \| X \|_N^2 \geq c \| \Phi \|_V^2, \]
noticing that the unit sphere of $\mathbb{R}^N$ is compact and thus the continuous mapping $X \mapsto \langle AX, X \rangle_N$ is bounded by below by a constant $c > 0$.

2. We have the following equivalence:
\[ X \in \text{Ker}(A) \Leftrightarrow \forall i, \sum_{i=1}^N a_{ij} X_j = 0 \Leftrightarrow \forall i, a(\Phi, \varphi_i) = 0 \]
\[ \Leftrightarrow \sup_{v \in V_h} a(\Phi, v) = 0, \]
we deduce
\[ (c1h) \Rightarrow (\forall \Phi \in V_h, \ (\sup_{v \in V_h} a(\Phi, v) = 0) \Rightarrow (\Phi = 0)) \]
\[ \Rightarrow \text{Ker}(A) = \{0\}. \]
Conversely, suppose $\text{Ker}(A) = \{0\}$. We consider a sequence $(w_n)_{n \in \mathbb{N}} \in V_h$ with $\| w_n \|_W = 1$ and such that
\[ \sup_{v \in V_h} \frac{a(w_n, v)}{\| v \|_V} \leq \frac{1}{n}. \]
From the compacity of the unit sphere of $V_h$, we deduce that any subsequence also denoted $(w_n)$ tends towards $w$ when $n \to \infty$. The limit is such that $\| w \|_V = 1$ and $\sup_{v \in V_h} a(w, v) = 0$. This means that $W \in \text{Ker}(A)$, where $w = \sum_{i=1}^N W_i \varphi_i$. Hence, $W = 0$ which is in contradiction with the assumption that $\| w \|_V = 1$. 
3. proceed similarly with $A'$.

\[
\square
\]

**Remark 4.8** If the bilinear form $a(\cdot, \cdot)$ is symmetric so is the matrix $A$.

### 4.4.3 Approximation of a saddle-point problem

We consider here the approximation of the abstract problem $(\mathcal{S})$. Suppose $V_h$ is a subspace of $V$, $W_h$ is a subspace of $W$, both of finite dimensions and we consider the discrete problem:

\[
(S_h) \begin{cases}
    \text{find } u_h \in V_h \text{ and } p_h \in W_h \text{ such that } \\
    a(u_h, v_h) + b(v_h, p_h) = f(v_h), \quad \forall v_h \in V_h, \\
    b(u_h, q_h) = g(q_h), \quad \forall q_h \in W_h.
\end{cases}
\]

(4.23)

The following result establishes the existence and uniqueness conditions for a solution to the discrete problem.

**Theorem 4.12** The discrete problem $(S_h)$ is well-posed if and only if two conditions are satisfied:

\[
\exists \alpha_h > 0, \quad \inf_{u_h \in \text{Ker}(B_h)} \sup_{v_h \in \text{Ker}(B_h)} \frac{a(u_h, v_h)}{\|u_h\|_V \|v_h\|_V} \geq \alpha_h,
\]

(4.24)

where $B_h : V_h \to W'_h$ is the operator induced by the bilinear form $b$ on the discrete spaces: $(B_h v_h, q_h) = b(v_h, q_h)$ and $\text{Ker}(B_h)$ denotes the kernel of $B_h$:

\[
\text{Ker}(B_h) = \{ v_h \in V_h, \ b(v_h, q_h) = 0, \forall q_h \in W_h \},
\]

and if

\[
\exists \beta_h > 0, \quad \inf_{q_h \in W_h} \sup_{v_h \in V_h} \frac{b(v_h, q_h)}{\|v_h\|_V \|q_h\|_W} \geq \beta_h.
\]

(4.25)

and the solution $(u_h, p_h)$ verifies the following estimates:

\[
\|u - u_h\|_V \leq c_{1h} \inf_{v_h \in V_h} \|u - v_h\|_V + c_{2h} \inf_{q_h \in W_h} \|p - q_h\|_W,
\]

(4.26)

with $c_{1h} = \left(1 + \frac{\|a\|}{\alpha_h}\right) \left(1 + \frac{\|b\|}{\beta_h}\right)$, $c_{2h} = \frac{\|b\|}{\alpha_h}$ if $\text{Ker}(B_h) \not\subset \text{Ker}(B)$ and $c_{2h} = 0$ if $\text{Ker}(B_h) \subset \text{Ker}(B)$ and

\[
\|p - p_h\|_W \leq c_{3h} \inf_{v_h \in V_h} \|u - v_h\|_V + c_{4h} \inf_{q_h \in W_h} \|p - q_h\|_W,
\]

(4.27)

with $c_{3h} = c_{1h} \frac{\|a\|}{\beta_h}$ and $c_{4h} = 1 + \frac{\|b\|}{\beta_h} + c_{2h} \frac{\|a\|}{\alpha_h}$.

**Proof.** It consists in using Theorem 4.8 while noticing that in finite dimension, the first condition on $\alpha_h$ implies the first two conditions of the theorem (the injectivity and surjectivity are equivalent notions for an endomorphism in finite dimension).

We introduce the notation

\[
Z_h(g) = \{ w_h \in X_h, \ b(w_h, q_h) = g(q_h), \forall q_h \in M_h \}.
\]

Let consider $v_h \in X_h$. Since $B_h$ satisfies (4.25), there exists $r_h \in X_h$ such that

\[
\forall q_h \in M_h, \quad b(r_h, q_h) = b(u - v_h, q_h) \quad \text{and} \quad \beta_h \|r_h\|_X \leq \|b\|_V \|u - v_h\|_X.
\]
4.5. Application to boundary-value problems in \( \mathbb{R} \)

Since \( b(r_h + v_h, q_h) = g(q_h) \), then \( r_h + v_h \) belongs to the space \( Z_h(g) \). We pose \( w_h = r_h + v_h \) and we notice that \( u_h - w_h \in \text{Ker}(B_h) \) and we have:

\[
\alpha_h \| u_h - w_h \|_X \leq \sup_{y_h \in \text{Ker}(B_h)} \frac{a(u_h - w_h, y_h)}{\| y_h \|_X} \leq \sup_{y_h \in \text{Ker}(B_h)} \frac{a(u_h - u, y_h) + a(u - w_h, y_h)}{\| y_h \|_X} \leq \sup_{y_h \in \text{Ker}(B_h)} \frac{b(y_h, p - p_h) + a(u - w_h, y_h)}{\| y_h \|_X}
\]

If \( \text{Ker}(B_h) \subset \text{Ker}(B) \), then \( b(y_h, p - p_h) = 0 \) for \( y_h \in \text{Ker}(B_h) \) and thus

\[
\alpha_h \| u_h - w_h \|_X \leq \| a \| \| u - w_h \|_X.
\]

The triangle inequality yields

\[
\| u - u_h \|_X \leq \left(1 + \frac{\| a \|}{\alpha_h}\right) \| u - w_h \|_X.
\]

In the general case, we have \( b(y_h, p - p_h) = b(y_h, q_h) = 0 \) for every \( q_h \in M_h \) since \( y_h \in \text{Ker}(B_h) \). Hence

\[
\alpha_h \| u_h - w_h \|_X \leq \| a \| \| u - w_h \|_X + \| b \| \| p - q_h \|_M
\]

The triangle inequality yields

\[
\| u - u_h \|_X \leq \left(1 + \frac{\| a \|}{\alpha_h}\right) \| u - w_h \|_X + \frac{\| b \|}{\beta_h} \| p - q_h \|_M.
\]

By noticing that the inequality \( \| u - w_h \|_X \leq \| u - v_h \|_X + \| r_h \|_X \) can be written as follows:

\[
\| u - w_h \|_X \leq \left(1 + \frac{\| b \|}{\beta_h}\right) \| u - v_h \|_X,
\]

we obtain the desired result.

Since for every \( v_h \in X_h \), \( b(v_h, p - p_h) = a(u_h - u, v_h) \), we obtain by considering \( q_h \in M_h \):

\[
\forall v_h \in X_h, \quad b(v_h, q_h - p_h) = a(u_h - u, v_h) + b(v_h, q_h - p).
\]

And by taking into account the condition (4.25), we deduce:

\[
\beta_h \| q_h - p_h \|_M \leq \| a \| \| u - u_h \|_X + \| b \| \| p - q_h \|_M.
\]

The final result is obtained using the triangle inequality. \( \square \)

4.5 Application to boundary-value problems in \( \mathbb{R} \)

We will briefly present the weak formulation and the variational approximation concepts on two basic boundary-value problems in one dimension.

4.5.1 Dirichlet problem

Let consider the homogeneous Dirichlet boundary-value problem in one dimension:

\[
\begin{align*}
-\frac{d}{dx} \left( \mu \frac{du}{dx} \right) + \nu u &= f, & \forall x \in \Omega = ]0, 1[, \\
u u(0) &= u(1) = 0.
\end{align*}
\tag{4.28}
\]
for \( \nu \) and \( \mu \) given functions of \( L^\infty(\Omega) \) such that
\[
\nu(x) \geq \alpha > 0, \quad \mu(x) \geq 0, \quad \text{a.e. in } \Omega.
\]

We pose
\[
a(u, v) = \int_0^1 (\nu u' v' + \mu u v) \, dx,
\]
and the weak formulation of this problem consists in finding \( u \in V = H_0^1(\Omega) \) solving:
\[
\forall v \in H_0^1(\Omega), \quad a(u, v) = \int_0^1 f v \, dx.
\] (4.29)

The bilinear form \( a(\cdot, \cdot) \) is continuous on \( V \times V \) and \( V \)-elliptic, thus, for any \( f \in L^2(\Omega) \), the variational problem (4.29) is well-posed and has a unique solution \( u \in V \).

To construct an approximation \( u_h \) of \( u \), we consider the subspace \( V_h \subset V \) composed of piecewise affine continuous functions on intervals. Let introduce the step \( h = \frac{1}{N+1} \), for \( N \in \mathbb{N}^* \) given, and the points \( x_n = nh, \ 0 \leq n \leq N + 1 \). The latter subdivide the interval \( \Omega \) into \( N + 1 \) subintervals \( K_n = [x_n, x_{n+1}] \) of length \( h \). The subspace \( V_h \) is chosen as:
\[
V_h = \{ v \in V, \ v|_{K_n} \in \mathbb{P}_1, \ 0 \leq n \leq N \},
\]
where \( \mathbb{P}_1 \) denotes the space of polynomial functions of degree lesser than or equal to one. A function \( v \in V = H_0^1(\Omega) \) is (almost everywhere equal to) a continuous function on \([0, 1]\) and satisfies the conditions \( v(0) = v(1) = 0 \). Furthermore, any continuous function on \( \Omega \), piecewise \( C^1 \) on intervals and such that \( v(0) = v(1) = 0 \) is a function of \( H_0^1(\Omega) \). This leads to consider \( V_h \) as:
\[
V_h = \{ v \in C^0(\Omega), \ v(0) = v(1) = 0, \ v|_{K_n} \in \mathbb{P}_1, \ 0 \leq n \leq N \}.
\]
The dimension of \( V_h \) is exactly \( N \) and the sequence of functions \( (\varphi_n)_{1 \leq n \leq N} \in V_h \) defined by:
\[
\varphi_i(x_j) = \delta_{i,j}, \quad 1 \leq j \leq N,
\]
forms a basis of \( V_h \). Each function \( \varphi_i \) can be written as:
\[
\varphi_i = \begin{cases} 
1 - \frac{|x - x_i|}{h} & \forall x \in [x_{i-1}, x_{i+1}], \\
0 & \text{otherwise}
\end{cases}
\]
Hence, a function \( u_h \in V_h \) can be uniquely determined using its coordinates \( u_n = u_h(x_n) \) in the basis \( (\varphi_n)_{1 \leq n \leq N} \). The variational method consists in approaching the solution \( u \in V \) by the function \( u_h = \sum_{j=1}^N u_j \varphi_j \in V_h \), where the \((u_j)_{1 \leq j \leq N} \) are solution of the linear system:
\[
\sum_{j=1}^N a(\varphi_j, \varphi_i) u_j = \int_0^1 f \varphi_j \, dx, \quad 1 \leq i \leq N.
\]
Actually, the linear system is relatively simple since the support of a function \( \varphi_j \) is the interval \([x_{j-1}, x_{j+1}]\). This leads to consider the system:
\[
\begin{align*}
    a(\varphi_1, \varphi_1) u_1 + a(\varphi_2, \varphi_1) u_2 &= \int_0^1 f \varphi_1 \, dx \\
    \vdots \\
    a(\varphi_{i-1}, \varphi_i) u_{i-1} + a(\varphi_i, \varphi_i) u_i + a(\varphi_{i+1}, \varphi_i) u_{i+1} &= \int_0^1 f \varphi_i \, dx \\
    \vdots \\
    a(\varphi_{N-1}, \varphi_N) u_{N-1} + a(\varphi_N, \varphi_N) u_N &= \int_0^1 f \varphi_N \, dx
\end{align*}
\]
The next step consists in evaluating the coefficients of this linear system and specifically all integrals. To this end, we introduce the notations:

\[
\nu_{i+1/2} = \frac{1}{h} \int_{x_i}^{x_{i+1}} \nu(x) \, dx ,
\]

\[
\mu_{i+1/2}^- = \frac{3}{h^3} \int_{x_i}^{x_{i+1}} (x_{i+1} - x)^2 \mu(x) \, dx ,
\]

\[
\mu_{i+1/2}^+ = \frac{6}{h^3} \int_{x_i}^{x_{i+1}} (x_{i+1} - x)(x - x_i) \mu(x) \, dx ,
\]

\[
\mu_{i-1/2}^- = \frac{3}{h^3} \int_{x_i}^{x_{i+1}} (x - x_{i+1})^2 \mu(x) \, dx ,
\]

and

\[
\begin{align*}
&f^-_{i+1/2} = \frac{2}{h^2} \int_{x_i}^{x_{i+1}} (x_{i+1} - x)f(x) \, dx , \\
&f^+_{i+1/2} = \frac{2}{h^2} \int_{x_i}^{x_{i+1}} (x - x_{i+1})f(x) \, dx ,
\end{align*}
\]

yielding to the following linear system:

\[
\begin{align*}
\begin{cases}
\left( \frac{1}{h} (\nu_1/2 + \nu_3/2) + \frac{h}{3} (\mu_{1/2}^- + \mu_{3/2}^-) \right) u_1 + \left( -\frac{1}{h} \nu_{3/2} + \frac{h}{6} \mu_{3/2}^- \right) u_2 = \frac{h}{2} (f^+_{1/2} + f^-_{3/2}) , \\
\vdots \\
\left( -\frac{1}{h} \nu_{i-1/2} + \frac{h}{6} \mu_{i-1/2} \right) u_{i-1} + \left( \frac{1}{h} (\nu_{i-1/2} + \nu_{i+1/2}) + \frac{h}{3} (\mu_{i-1/2}^- + \mu_{i+1/2}^-) \right) u_i + \\
\left( -\frac{1}{h} \nu_{i+1/2} + \frac{h}{6} \mu_{i+1/2}^- \right) u_{i+1} = \frac{h}{2} (f^+_{i-1/2} + f^-_{i+1/2}) , \\
\vdots \\
\left( -\frac{1}{h} \nu_{N-1/2} + \frac{h}{6} \mu_{N-1/2} \right) u_{N-1} + \left( \frac{1}{h} (\nu_{N-1/2} + \nu_{N+1/2}) + \frac{h}{3} (\mu_{N-1/2}^- + \mu_{N+1/2}^-) \right) u_N = \frac{h}{2} (f^+_{N-1/2} + f^-_{N+1/2})
\end{cases}
\end{align*}
\]

The last issue that remains to be addressed concerns the right-hand side term. If the function \(f\) is an arbitrary function, the quantities \(f^+_{i+1/2}\) and \(f^-_{i+1/2}\) cannot be evaluated directly, they must be evaluated using a numerical integration scheme. Given a continuous function \(g\) defined on a bounded interval \([a, b]\), the mid-point formula allows to compute (exactly if \(g\) is affine) the integral of \(g\) on \([a, b]\) as:

\[
\int_a^b g(x) \, dx \approx (b - a) g \left( \frac{a + b}{2} \right).
\]

We proceed similarly for evaluating the quantities involving the functions \(\nu\) and \(\mu\). It is important to carefully evaluate such quantities as it affects the stiffness matrix \(A\) of the system.

The choice of the basis \((\varphi_n)_{1 \leq n \leq N}\) as described above leads to a tridiagonal matrix \(A\). Thus, efficient methods exist to resolve the system numerically (cf. Chapter 5).

### 4.5.2 Neumann problem

Let consider the Neumann boundary-value problem in one dimension:

\[
\begin{align*}
-\frac{d}{dx} \left( \nu \frac{du}{dx} \right) + \mu u &= f , & \quad \forall x \in \Omega = ]0, 1[, \\
u'(0) = u'(1) &= 0.
\end{align*}
\]
for \( \nu \) and \( \mu \) given functions of \( L^\infty(\Omega) \) such that
\[
\nu(x) \geq \alpha > 0, \quad \mu(x) \geq 0, \quad \text{a.e. in } \Omega.
\]

We have already seen that for any given \( f \in L^2(\Omega) \), this problem is well-posed and the solution \( u \in H^1(\Omega) \).

Considering the subdivision of \( \Omega \) into \( N+1 \) subintervals \( K_n = [x_n, x_{n+1}] \) of length \( h \), we introduce the space \( V_h \subset H^1(\Omega) \) defined as:
\[
V_h = \{ v \in C^0(\Omega), \quad v|_{K_n} \in \mathbb{P}_1, \quad 1 \leq n \leq N \}.
\]

It is a subspace of \( H^1(\Omega) \) of dimension \( N+2 \). A basis of this space is given by the functions \( (\varphi_n)_{0 \leq n \leq N+1} \):
\[
\varphi_0(x) = \begin{cases} 
1 - \frac{x}{h} & \text{if } 1 \leq x \leq h \\
0 & \text{otherwise}
\end{cases}
\]
\[
\varphi_i(x) = \begin{cases} 
1 - \frac{|x-x_i|}{h} & \text{if } x_{i-1} \leq x \leq x_{i+1} \\
0 & \text{otherwise}
\end{cases}
\]
\[
\varphi_{N+1}(x) = \begin{cases} 
1 - \frac{x}{h} & \text{if } 1-h \leq x \leq 1 \\
0 & \text{otherwise}
\end{cases}
\]

The variational approximation method consists in searching for the solution \( u_h = \sum_{j=0}^{N+1} u_j \varphi_j \in V_h \) solving the linear system:
\[
\begin{bmatrix}
\left( \frac{1}{h}(\nu_{1/2} + \frac{h}{3}(\mu_{1/2}) \right) u_0 + \left( -\frac{1}{h}\nu_{1/2} + \frac{h}{6}\mu_{1/2} \right) u_1 = \frac{h}{2}(f_{1/2}) \\
\vdots \\
\left( -\frac{1}{h}\nu_{i-1/2} + \frac{h}{6}\mu_{i-1/2} \right) u_{i-1} + \left( \frac{1}{h}(\nu_{i-1/2} + \nu_{i+1/2}) + \frac{h}{3}(\mu_{i-1/2} + \mu_{i+1/2}) \right) u_i + \\
\vdots \\
\left( -\frac{1}{h}\nu_{N+1/2} + \frac{h}{6}\mu_{N+1/2} \right) u_N + \left( \frac{1}{h}\nu_{N+1/2} + \frac{h}{3}\mu_{N+1/2} \right) u_{N+1} = \frac{h}{2}(f_{N+1/2})
\end{bmatrix}
\]

As for the Dirichlet boundary-value problem, the stiffness matrix \( A \) is tridiagonal and symmetric positive definite.

**Remark 4.9** In the Neumann problem, the numerical solution \( u_h \) does not verify the boundary conditions \( u'_h(0) = u'_h(1) = 0 \), neither a priori nor a posteriori.

### 4.6 Maximum principles

We conclude this chapter by a few words on maximum principles.
4.7 Exercises

Exercise 4.1 Suppose $\Omega \subset \mathbb{R}^d$ is an open set and $f \in L^2(\Omega)$ is a given function. We consider the following Neumann problem:

$$
\begin{aligned}
-\Delta u + u &= f, & \text{on } \Omega \\
\frac{\partial u}{\partial n} &= 0, & \text{on } \partial \Omega.
\end{aligned}
$$

(4.31)

1. Show that the weak formulation of this problem is:

$$
\forall v \in H^1(\Omega), \quad \int_\Omega \nabla u \cdot \nabla v + uv\,dx = \int_\Omega fv\,dx,
$$

and that there is a unique solution to this problem in $H^1(\Omega)$.

2. Show what the solution $u \in H^1(\Omega)$ of this problem verifies $-\Delta u + u = f$ in $\mathcal{D}'(\Omega)$.

3. Suppose $\Omega$ is of class $C^2$ and thus $u \in H^2(\Omega)$. In what sense is $u$ a solution of the initial problem?