Contents

Chapter 1

Basic concepts

1.1 Weak formulation of boundary value problems

Consider the following boundary value problem

$$\begin{cases} -u'' = f \text{ in }]0,1[\\ u(0) = u(1) = 0. \end{cases} \tag{1.1}$$

If u is the solution of (1.1) and v is a sufficiently regular function satisfies $v(0) = v(1) = 0$, then multiplying the first equation of (1.1) by v integrating we get:

$$\int_0^1 -u''(x)v(x)\ dx = \int_0^1 f(x)v(x)\ dx =: (f,v) \tag{1.2}$$

then integration by parts yields

$$(f,v) := \int_0^1 u'(x)v'(x)\ dx =: a(u,v). \tag{1.3}$$

Let us define (formally, for the moment since the notion of derivative to be used has not been made precise)

$$V = \{v \in L^2(]0,1[); a(v,v) < +\infty \text{ and } v(0) = v(1) = 0\}.$$

Definition 1.1

The space V is called the functional space

Then the solution of (1.1) is characterized by

$$\begin{cases} u \in V \text{ s.t} \\ a(u,v) = (f,v), \quad \forall v \in V. \end{cases} \tag{1.4}$$

Definition 1.2

Problem (1.4) is called *the weak formulation* or the variational problem of problem (1.1).

Remark 1.1

By construction it is clear that any solution of (1.1) is a solution of (1.4). Of course, the central issue is that (1.4) gives a solution of (1.1). The following theorem verifies this under some simplifying assumptions.

Theorem 1.1

Suppose that $f \in C^0([0,1])$ and $u \in C^2([0,1])$ satisfies (1.4). Then u is a solution of (1.1).

Proof

Let $v \in V \cap C^1([0,1])$, then

$$(f,v) = \int_\Omega u'v' \; dx = -\int_\Omega u''v \; dx + u'v|_0^1 = (-u'',v).$$

So,

$$(f+u'',v) = 0, \quad \forall v \in V \cap C^1([0,1]). \tag{1.5}$$

To finish the proof, it suffices to proof that (1.5) implies $-u'' = f$. Indeed, if not ($-u'' \neq f$) then (1.5) implies the existence of x_0 and x_1, $0 < x_0 < x_1 < 1$ such that $(f+u'')$ has the same sign on $[x_0,x_1]$ i.e $f+u'' > 0$ or $f+u'' < 0$.
take

$$v(x) = (x-x_0)^2(x-x_1)^2$$

this implies that $(f+u'',v) \neq 0$ which is in contradiction with (1.5).

Remark 1.2

The choice of the functional space is crucial. It seems to be natural that the space were we seek the solution is the same as the space of the test functions v Observe that if the space if with finite dimension then the weak formulation (1.4) can be reformulated as a linear system $Ax = b$. in that case, the uniqueness means that A is injective thus it is an isomorphisme this implies the existence also.But we work here in infinite dimensional case, and the previous theory does not work here. The existence of solution follows from the Lax-Milgram theorem (see Section 1.3) which suppose that the space V must be a Hilbert space.

1.1.0.1 Green's Formula

Let Ω be a bounded open subset of $\mathbb{R}^N$ with boundary $\partial\Omega$ of class C^1.

Theorem 1.2

Let $v \in C^1(\bar{\Omega})$ with compact support in $\bar{\Omega}$. Then we have

$$\int_\Omega \frac{\partial v}{\partial x_i}(x) \; dx = \int_{\partial\Omega} v(x) n_i(x) \; ds \tag{1.6}$$

where n_i is the i-th component of the unit outward normal to Ω.

The trace theorem (Theorem 1.6) allows us to generalize Green's formula to elements of the space $H^1(\Omega)$.

Theorem 1.3

Let Ω be a bounded open subset of $\mathbb{R}^N$ with boundary $\partial\Omega$ of class C^1. If u, v are two functions in $H^1(\omega)$. Then

$$\int_\Omega u\frac{\partial v}{\partial x_i}(x)\ dx = -\int_\Omega \frac{\partial u}{\partial x_i}(x)\ v(x)\ dx + \int_{\partial\Omega} u(x)v(x)n_i(x)\ ds \tag{1.7}$$

Theorem 1.4: (Green)

Let Ω be a bounded open subset of $\mathbb{R}^N$ with boundary $\partial\Omega$ of class C^1.
If $u \in H^2(\Omega)$ and $v \in H^1(\Omega)$. Then

$$-\int_\Omega \Delta u\ v\ dx = \int_\Omega \nabla u \cdot \nabla v\ dx - \int_{\partial\Omega} \frac{\partial u}{\partial n}\ v\ ds \tag{1.8}$$

1.2 Sobolev spaces $H^1(\Omega), H^m(\Omega)$

Let Ω be a bounded open subset of $\mathbb{R}^N$, $C_0^\infty(\Omega)$ is the space of class C^∞ with compact support included in Ω and $\mathcal{D}'(\Omega)$ is the space of distributions defined on Ω

Definition 1.3

Given a function $v \in L^2(\Omega)$. we say that v has a weak derivative if there exists $w_i \in L^2(\Omega)$, for all $i = 1, ..., N$ such that for all $\phi \in C_0^\infty(\Omega)$ we have

$$\int_\Omega v\frac{\partial \phi}{\partial x_i} = -\int_\Omega w_i\phi\ dx$$

the functions w_i are called the weak derivatives and they are denoted by $\dfrac{\partial v}{\partial x_i}$.

Definition 1.4

Let Ω be an open set of $\mathbb{R}^N$. The Sobolev space $H^1(\Omega)$ is given by :

$$H^1(\Omega) =: \{u \in L^2(\Omega); \frac{\partial u}{\partial x_i} \in L^2(\Omega), i = 1, 2, ..., N\} \tag{1.9}$$

where $\dfrac{\partial u}{\partial x_i}$ is the weak derivative of u

More, generally, Let

$$\alpha = (\alpha_1, \alpha_2, ..., \alpha_i, ..., \alpha_N), \alpha_i \in \mathbb{N}, \quad i = 1, 2, ..., N$$

a multi index We note

$$D^\alpha = \partial_1^{\alpha_1}\partial_2^{\alpha_2}...\partial_N^{\alpha_N}, \quad |\alpha| = \sum_{i=1}^N \alpha_i.$$

Definition 1.5

for $m \in \mathbb{N}$

$$H^m(\Omega) =: \{u \in \mathcal{D}'(\Omega); D^\alpha u \in L^2(\Omega) \quad |\alpha| \le m\} \tag{1.10}$$

For $m=0$ we have $H^0(\Omega) = L^2(\Omega)$ and for $m=1$ we find the definition given by (1.4). We equiped $H^m(\Omega)$ by the inner product:

$$(u,v)_m = \sum_{|\alpha|\le m} \int_\Omega D^\alpha u D^\alpha v \, dx. \tag{1.11}$$

The associated norm is

$$\|u\|_{m,\Omega} = \left(\sum_{|\alpha|\le m} \int_\Omega |D^\alpha u|^2 \right)^{1/2}. \tag{1.12}$$

Proposition 1.1

i) If $m \ge m'$, $H^m(\Omega)$ is continuously embedded in $H^{m'}(\Omega)$.
ii) $H^m(\Omega)$ equiped with the inner product (1.11) is a Hilbert spacce.

1.2.1 A Sobolev embedding theorem

Theorem 1.5

Let Ω be a bounded open subset of $\mathbb{R}^2$ with Lipschitz boundary m and k two integer numbers satisfying $k-m>1$. Then there exists a positive constant C such that for all $u \in H^k(\Omega)$ we have:

$$\|D^m u\|_{L^\infty(\Omega)} \le C\|u\|_{k,\Omega} \tag{1.13}$$

In addition there exists a function of class C^m equal to u almost every where.

1.2.2 Trace theorem

Suppose that Ω is suffiicetly regular (of classe C^1, for example), Then we define the trace γ_0 by :

$$\begin{aligned} \gamma_0 : H^1(\Omega) \cap C^0(\bar{\Omega}) &\to L^2(\partial\Omega) \cap C^0(\overline{\partial\Omega}) \\ u &\mapsto \gamma_0(u) = u_{/\partial\Omega} \end{aligned} \tag{1.14}$$

Theorem 1.6

The linear map γ_0 given in (1.14)
can be extended to a linear continuous map from $H^1(\Omega)$ to $L^2(\partial\Omega)$, i.e., there exists c such that:

$$\|u\|_{L^2(\partial\Omega)} \le c \, \|u\|_{H^1(\Omega)}, \forall u \in H^1(\Omega). \tag{1.15}$$

1.2.3 L'espace $H_0^1(\Omega)$

Definition 1.6

Let $\mathcal{D}(\Omega)$ the space of C^∞ with compact support included in Ω. We define $H_0^1(\Omega)$ as the aderence of $\mathcal{D}(\Omega)$ in $H^1(\Omega)$, i.e.,

$$\overline{\mathcal{D}(\Omega)}^{\|\cdot\|_{1,\Omega}} =: H_0^1(\Omega).$$

Theorem 1.7

The $H_0^1(\Omega)$ is the kernel of γ_0, i.e.,

$$H_0^1(\Omega) = \{u \in H^1(\Omega), u_{/\partial\Omega} = 0\}$$

Lemma 1.1: Poincaré's inequality

Let Ω a bounded set of $\mathbb{R}^N$. Then, there exists a positive C which depends only on Ω such that:

$$\forall v \in H_0^1(\Omega), \quad \|v\|_{L^2} \leq C\|\nabla v\|_{L^2}. \tag{1.16}$$

1.2.3.1 Green's formula

Let Ω be an open bounded subset of $\mathbb{R}^N$ with boundary $\partial\Omega$ of class C^1.

Theorem 1.8

Let $v \in C^1(\bar{\Omega})$ has a compact support in $\bar{\Omega}$.
Then we have

$$\int_\Omega \frac{\partial v}{\partial x_i}(x)\ dx = \int_{\partial\Omega} v(x) n_i(x)\ ds \tag{1.17}$$

where n_i is the i'th componenet of the exterior normal of Ω.

The trace theorem allows the generalisaton of the Green's formula to the elements of the space $H^1(\Omega)$.

Theorem 1.9

Let Ω be an open bounded subset of $\mathbb{R}^N$ with boundary $\partial\Omega$ of class C^1. Ii u, v two functions of $H^1(\omega)$. Then

$$\int_\Omega u \frac{\partial v}{\partial x_i}(x)\ dx = -\int_\Omega \frac{\partial u}{\partial x_i}(x)\ v(x)\ dx + \int_{\partial\Omega} u(x) v(x) n_i(x)\ ds \tag{1.18}$$

Theorem 1.10: (Green)

Let Ω a bounded open subset of $\mathbb{R}^N$ with boundary $\partial\Omega$ of class C^1.
If $u \in H^2(\Omega)$ and $v \in H^1(\Omega)$. Then

$$-\int_\Omega \Delta u\ v\ dx = \int_\Omega \nabla u \cdot \nabla v\ dx - \int_{\partial\Omega} \frac{\partial u}{\partial n}\ v\ ds \tag{1.19}$$

1.2.3.2 The spaces $H^{-1}(\Omega)$, $H^{1/2}(\partial\Omega)$ and $H^{-1/2}(\partial\Omega)$

The space $H^{-1}(\Omega)$ is by definition the dual space of $H_0^1(\Omega)$, i.e This space is equipped with the norm:

$$\|f\|_{H^{-1}(\Omega)} = \sup_{v\in H_0^1(\Omega)} \frac{\langle f, v\rangle}{|v|_{H^1(\Omega)}}$$

The space $H^{1/2}(\partial\Omega)$ is the trace space of the element of $H^1(\Omega)$ on the boundary of Ω. Thanks to the inequality (1.15) the space $H^{1/2}(\partial\Omega)$ is a sub space of $L^2(\partial\Omega)$. We denote by $H^{-1/2}(\partial\Omega)$ to the dual space of $H^{1/2}(\partial\Omega)$.

1.2.4 The Sobolev spaces $W^{m,p}(\Omega)$

Definition 1.7

The space

$$W^{m,p}(\Omega) = \{u \in L^p(\Omega); D^\alpha u \in L^p(\Omega) \quad \text{for} \quad |\alpha| \leq m\}$$

The norm and the semi norm are given by:

$$\|v\|_{W^{m,p}(\Omega)} = \left(\sum_{|\alpha|\leq m} \|D^\alpha v\|^p_{L^p(\Omega)}\right)^{1/p}$$

$$|v|_{W^{m,p}(\Omega)} = \left(\sum_{|\alpha|= m} \|D^\alpha v\|^p_{L^p(\Omega)}\right)^{1/p}$$

We observe that this space generalises the space H^m, since,

$$H^m(\Omega) = W^{m,2}(\Omega)$$

Definition 1.8

Let $\Omega \subset \mathbb{R}^d$. The Sobolev number of the space $W^{m,p}(\Omega)$, is given by:

$$sob(W^{m,p}) = m - d/p$$

1.2.5 Properties of Sobolev space $W^{m,p}(\Omega)$

We present here some properties of the Sobolev space $W^{m,p}(\Omega)$ without proof for more details see (voir [3]).

Definition 1.9

A map from the space X to the space Y is compact if it transforms any bounded sequence in X to a sequence where we can extract a convergente subsequence in Y If $X \subset Y$ we say that the embedding of X in Y is compact if the identity is compact and we denote it by $X \hookrightarrow Y$.

Theorem 1.11

Let $m > k$, $\text{sob}(W^{m,p}(\Omega)) > \text{sob}(W^{k,q}(\Omega))$, with Lipschitz boundary $\partial\Omega$. Then

$$W^{m,p}(\Omega) \hookrightarrow W^{k,q}(\Omega)$$

We Note that:

$$u(x) = \log\left|\log\left(\frac{|x|}{2}\right)\right| \in W^{1,d}(\Omega) \setminus L^\infty(\Omega) \text{ where } \Omega \text{ is the unit ball of } \mathbb{R}^d$$

and we have have also

$$sob(W^{1,d}(\Omega)) = 1 - d/d = 0 = 0 - d/\infty = sob(L^\infty).$$

Remark 1.3

We observe that in general, equal Sobolev numbers does not imply necessarily the embbeding between the corresponding spaces.

1.3 Well posed problems

In this section we consider problems of the form

$$\begin{cases} \text{Find } u \in W \quad \text{s.t} \\ a(u,v) = F(v), \quad \forall v \in V. \end{cases} \tag{1.20}$$

where the form $a(\cdot,\cdot)$ is bilinear on $W \times V$ and the form $F(\cdot)$ is continuous on V.

Definition 1.10

We say that the problem (1.20) is well posed (in Hadamard sense) if it has a unique solution depends continuously on the data (stable).

1.3.1 Lax-Milgram theorem

The Lax-Milgram theorem [13] is an interesting theorem it treats the case $V = W$. Consider a variational problem of the form:

$$\begin{cases} \text{Find } u \in V \text{ s.t} \\ a(u,v) = F(v), \quad \forall v \in V. \end{cases} \tag{1.21}$$

The proof is based on:

- The Riesz representation theorem, and
- the Banach fixed point theorem

Theorem 1.12: (Lax-Milgram)

Let V be a Hilbert space, equipped with the norm: $\|\cdot\|_V$. We suppose that :

i) the bilinear form $a(\cdot,\cdot)$ is continuous, i.e

$$\exists \beta < +\infty, \quad \forall (u,v) \in V \times V, \quad |a(u,v)| \leq \beta \|u\|_V \|v\|_V;$$

ii) the bilinear form $a(\cdot,\cdot)$ is coercive, i.e

$$\exists \alpha > 0, \quad \forall u \in V, \quad a(u,u) \geq \alpha \|u\|_V^2. \tag{1.22}$$

iii) The linear form $F(\cdot)$ is bounded, i.e

$$\exists \gamma < +\infty, \quad \forall v \in V, \quad |F(v)| \leq \gamma \|v\|_V;$$

Then the problem (1.21) has a unique solution. Moreover, we have the following a priori estimates:

$$\|u\|_V \leq \frac{\|F\|_{V'}}{\alpha}. \tag{1.23}$$

Proof

For all $u \in V$, we define the operator Au by

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

Then $Au \in V'$ and we have also the map $u \mapsto Au$ is linear and continuous from V into V', i.e., $A \in \mathcal{L}(V,V')$ since we have,

$$\|Au\|_{V'} = \sup_{v\neq 0} \frac{|Au(v)|}{\|v\|} \leq \beta\|u\|$$

We use the fact that $Au \in V'$, then the Riesz representation theorem claims that:

$$\exists! w_0 \in V, \text{ s.t } < Au, v >= (w_0, v), \quad \forall v \in V.$$

Hence, the proof of Lax-Milgram theorem is equivalent to:

$$\begin{cases} \text{Find } u \in V \text{ s.t} \\ \quad \tau Au = w_0 = \tau F \text{ in } V. \end{cases} \tag{1.24}$$

where, τ is the Riesz isomorphism from V' into V.
Or

$$\begin{cases} \text{Find } u \in V \text{ s.t} \\ \quad Au = F \text{ in } V'. \end{cases} \tag{1.25}$$

Then we define the operator

$$Tw = w - \lambda(\tau Aw - \tau F).$$

If the operator T is a contraction, i.e

$$\exists M < 1, \quad \|Tx - Ty\| \leq M\|x - y\|$$

then the equation

$$Tu = u \tag{1.26}$$

has a unique solution in V.
If it is the case then

$$\lambda(\tau Au - \tau F) = 0 \iff \tau Au = \tau F.$$

Hence, the problem reduces to prouve that such $\lambda \neq 0$ exists.
For all $v_1, v_2 \in V$, we put $v = v_1 - v_2$. Then,

$$\begin{aligned}
\|Tv_1 - Tv_2\|^2 &= \|v_1 - v_2 - \lambda(\tau Av_1 - \tau Av_2)\|^2 \\
&= \|v - \lambda(\tau Av)\|^2 && \tau \text{ and } A \text{ are linear} \\
&= \|v\|^2 - 2\lambda(\tau Av, v) + \lambda^2\|\tau Av\|^2 \\
&= \|v\|^2 - 2\lambda\tau Av(v) + \lambda^2 Av(\tau Av) && \text{definition of } \tau \\
&= \|v\|^2 - 2\lambda a(v,v) + \lambda^2 a(v, \tau Av) && \text{definition of } A \\
&\leq \|v\|^2 - 2\lambda\alpha\|v\|^2 + \lambda^2\beta\|v\|\|\tau Av\|, && \text{Continuity and the coercivity of } a \\
&\leq (1 - 2\lambda\alpha + \lambda^2\beta^2)\|v_1 - v_2\|^2 \\
&= M^2\|v_1 - v_2\|^2.
\end{aligned}$$

We recall that α and β are the coercivity and the continuity constants of the bilinear form a. So, we need to choose λ such that:

$$1 - 2\lambda\alpha + \lambda^2\beta^2 < 1, \text{ i.e., } \lambda(\lambda\beta^2 - 2\alpha) < 0.$$

If we take $\lambda \in (0, \frac{2\alpha}{\beta^2})$, then $M < 1$.
To prouve (1.23), we have:

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

Remark 1.4

The Lax-Milgram theroem gives a sufficient condition for the well-posedness of the variational problem (1.4).

1.4 The Galerkin Method

The fundamental idea on which the finite element method is built is the Galerkin method. In this section, we will study the principle of the method, its optimal nature in terms of approximation error and its reformulation using a linear system.

1.4.1 Principe of the method

Galerkin's method provides a simple and elegant way to approach the solution of the problem (1.21).

Subsequently, we will always assume that the assumptions of the Lax-Milgram theorem are satisfied for a problem of the form(1.21), so that this problem is well posed.

The principle of Galerkin's method consists of replacing the infinite dimensional space V (where the exacte solution exist) by a finite dimensional space V_h (where the approximate solution is computed) The space V_h is called approximation space. The index h refers to the fineness of the mesh which was used to construct the space V_h; its role will be clarified in the following sections. Subsequently, we will assume

$$V_h \subset V$$

When this condition is satisfied, we speak on conforming approximation. [1] The approximate version of the problem (1.21) consists of

$$\begin{cases} \text{find } u_h \in V_h \quad \text{s.t} \\ a(u_h, v_h) = F(v_h), \quad \forall v_h \in V_h. \end{cases} \tag{1.27}$$

Proposition 1.2

The approximate problem (1.27) admits one and only one solution.

Proof

The bilinear form $a(\cdot,\cdot)$ being coercive on V, and therefore on V_h since $V_h \subset V$. We conclude using the Lax-Milgram theorem

1.4.2 A priori error estimation

Our objective is now to estimate the approximation error $e_h = u - u_h$ in the norm $\|.\|_V$. Let us start by observing that since $V_h \subset V$, we have for all $v_h \in V_h$,

$$a(u, v_h) = F(v_h).$$

---

[1] We can also design Galerkin methods in a non-conforming framework, i.e.,

$$V_h \not\subseteq V,$$

Consequently, the approximate problem (1.27) is equivalent to searching for $v_h \in V_h$ such that

$$a(u-u_h, v_h) = a(e_h, v_h) = 0 \quad \forall v_h \in V_h \tag{1.28}$$

This equation is called the Galerkin orthogonality relation. When the bilinear form a is symmetric, it defines a scalar product $a(.,.)$ on V. The assumptions of coercivity and continuity on a mean that the norm induced by this scalar product is equivalent to the norm $\|.\|_V$ since

$$\alpha\|v\|_V^2 \le a(v,v) \le \beta\|v\|_V^2, \quad \forall v \in V.$$

The orthogonality relation (1.28) admits a simple geometric interpretation: u_h is the orthogonal projection on V_h of the exact solution u with respect to the scalar product $a(.,.)$. (see figure 1.1)

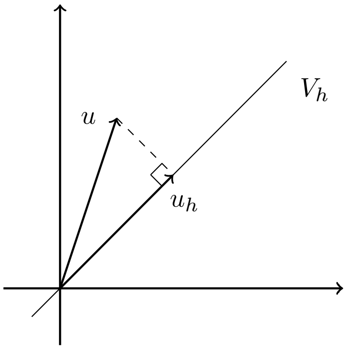

Fig. 1.1 Geometric interpretation $V = \mathbb{R}^2$

Lemma 1.2: Cea's Lemma

We have the following error estimate

$$\|u-u_h\|_V \le \frac{\beta}{\alpha} \inf_{v_h \in V_h} \|u-v_h\|_V. \tag{1.29}$$

Proof

We distinguish two cases

1. If $u \in V_h$, in this case, u is solution of the discrete problem (1.27), and by uniqueness of the solution,
$$u = u_h.$$
Moreover,
$$\inf_{v_h \in V_h} \|u-v_h\|_V = 0.$$
So if $u \in V_h$. The estimate (1.29) reduces to the tautology $0 \le 0$
2. If $u \notin V_h$.

we have for any $u_h \in V_h$,
$$a(u-u_h, u-u_h) = a(u-u_h, u-v_h).$$
Using coercivity and the form a, it comes

$$\alpha\|u-u_h\|_V^2 \le \beta\|u-u_h\|_V\|u-v_h\|_V$$

hence the estimate by dividing by $\alpha\|u-u_h\|_V$ and taking the infimum on $v_h \in V_h$.

1.5 The linear system

V_h being of finite dimension, the approximate problem (1.27) reduces to the resolution of a linear system. Indeed, let $N = dim(V_h)$ and let $(\varphi_1, ..., \varphi_N)$ be a base of V_h. Let's us put

$$u_h = \sum_{i=1}^{N} u_i \varphi_i$$

The problem (1.27) is equivalent to looking for $U = (u_1, ..., u_N) \in \mathbb{R}^N$ such that

$$\sum_{j=1}^{N} a(\varphi_j, \varphi_i) u_j = b(\varphi_i), \quad 1 \le i \le N$$

We put

$$A = (A_{ij})_{1 \le i,j \le N} \in \mathbb{R}^{N,N}, \quad A_{ij} = a(\varphi_j, \varphi_i), \tag{1.30}$$

and

$$B = (B_i)_{1 \le i \le N} \in \mathbb{R}^N, B_i = b(\varphi_i), \tag{1.31}$$

we obtain the following linear system

$$AU = B. \tag{1.32}$$

The A matrix is called the stiffness matrix in reference to the problems in mechanics where it was first introduced.
The properties of the matrix A are directly inherited from those of the bilinear form $a(\cdot, \cdot)$.
We have the following result

Proposition 1.3

If the bilinear form $a(\cdot, \cdot)$ is symmetric, the matrix A is symmetric. Furthermore, if the bilinear form $a(\cdot, \cdot)$ is coercive, the matrix A is positive definite.

Proof

The property about the symmetry of A is obvious. Let's show the one on the definite positivity. Let $\xi = (\xi_1, ..., \xi_N) \in \mathbb{R}^N$ and let $x = \sum_{i,j=1}^{N} \xi_i \varphi_i$. A direct calculation shows that

$$(\xi, A\xi)_{\mathbb{R}^N} = \sum_{i,j=1}^{N} \xi_i A_{ij} \xi_j$$

$$\begin{aligned} &= \sum_{i,j=1}^{N} \xi_i \xi_j a(\varphi_j, \varphi_i) \\ &= a(\sum_{j=1}^{N} \xi_i \varphi_j, \sum_{i=1}^{N} \xi_i \varphi_i) \\ &= a(x, x) \end{aligned}$$

so that $(\xi, A\xi)_{\mathbb{R}^N} = 0$ implies by coercivity $x = 0$, that is to say $\xi = 0$.

Chapter 2

Construction of finite element spaces

2.1 Ciarlet's finite element

The first part of the definition is formalised by Ciarlet's [8] definition of a finite element.

Definition 2.1

Let

1. the element domain $K \subset \mathbb{R}^n$ be a compact set, (with piecewise smooth boundary when $n > 1$),
2. the space of shape functions $\mathcal{P}$ be a finite dimensional space of functions on K, and
3. the set of nodal variables $\mathcal{N} = (N_0, \dots, N_k)$ be a basis for the dual space $\mathcal{P}'$.

The triple $(K, \mathcal{P}, \mathcal{N})$ is called a finite element.

For the cases considered in this course, K will be a polygon such as a triangle, square, tetrahedron or cube, and $\mathcal{P}$ will be a space of polynomials. Here, $\mathcal{P}'$ is the dual space to $\mathcal{P}$, defined as the space of linear functions from $\mathcal{P}$ to $\mathbb{R}$.
Examples of dual functions to $\mathcal{P}$ include:

1. The evaluation of $p \in \mathcal{P}$ at a point $x \in K$.
2. The integral of $p \in \mathcal{P}$ over a line $L \in K$.
3. The integral of $p \in \mathcal{P}$ over K.
4. The evaluation of a component of the derivative of $p \in \mathcal{P}$ at a point $x \in K$.

Example 2.1: (1-d Lagrange element)

The 1-dimensional Lagrange element $(K, \mathcal{P}, \mathcal{N})$ of degree k is defined by

1. the element K is an interval $I \subset \mathbb{R}$,
2. the space $\mathcal{P}$ is the $(k+1)$-dimensional space of all polynomials of degree k on K, and
3. the set of nodal variables $\mathcal{N} = (N_0, \dots, N_k)$ given by:

$$N_i(v) = v(x_i), \quad x_i = a + \frac{(b-a)i}{k}, \quad \forall v \in \mathcal{P}, \quad i = 0, \dots, k.$$

Ciarlet's finite element provides us with a standard way to define a basis for the $\mathcal{P}$, called the nodal basis.

Definition 2.2

Let $(K,\mathcal{P},\mathcal{N})$ be a finite element. The nodal basis is the basis $\{\phi_0,\phi_2,\ldots,\phi_k\}$ of $\mathcal{P}'$ that is dual to $\mathcal{N}$, i.e

$$N_i(\phi_j)=\delta_{ij}$$

Given a triple $(K,\mathcal{P},\mathcal{N})$

Lemma 2.1: (Dual condition)

Let K and $\mathcal{P}$ defined as above and let $\{N_0,N_1,\ldots,N_k\}\in\mathcal{P}'$. Let $\{\psi_0,\psi_1,\ldots,\psi_k\}$ be a basis for $\mathcal{P}$. Then the following statements are equivalent.

1. $\{N_0,N_1,\ldots,N_k\}$ is a basis of $\mathcal{P}'$
2. if $v\in\mathcal{P}$ satisfies $N_i(v)=0$ for $i=0,\ldots,k$ then $v=0$.

Definition 2.3: (Unisolvence)

We say that $\mathcal{N}$ determines $\mathcal{P}$ if it satisfies condition 2 of Lemma 2.1. If this is the case, we say that $(K,\mathcal{P},\mathcal{N})$ is unisolvent.

2.1.1 2D and 3D finite elements

We would like to construct some finite elements with 2D and 3D domains $\mathcal{K}$. The fundamental theorem of algebra does not directly help us there, but the following lemma is useful when checking that $\mathcal{N}$ determines $\mathcal{P}$ in those cases.

Lemma 2.2

et $p(x):\mathbb{R}^d\to\mathbb{R}$ be a polynomial of degree $k\geq 1$ that vanishes on a hyperplane $\mathit{\Pi}_L$ defined by:

$$\mathit{\Pi}_L:=\{x:L(x)=0\}$$

for a non-degenerate affine function $L(x):\mathbb{R}^d\to\mathbb{R}$. Then $p(x)=L(x)q(x)$ where $q(x)$ is a ploly-nomial of degree $k-1$.

Proof

Choose coordinates (by shifting the origin and applying a linear transformation) such that $x = (x_1,\ldots,x_d)$ with $L(x) = x_d$ so Π_L is defined by $x_d = 0$. Then the general form for a polynomial is

$$p(x_1,\ldots,x_d) = \sum_{i_d=0}^{k}\left(\sum_{|i_1+\ldots+i_{d-1}|\le k-i_d} c_{i_1,\ldots,i_{d-1},i_d} x_d^{i_d}\prod_{l=1}^{d-1} x_l^{i_l}\right),$$

then $p(x_1,\ldots,x_{d-1},0) = 0$ for all $(x_1,\ldots,x_{d-1})$, so

$$0 = \left(\sum_{|i_1+\ldots+i_{d-1}|\le k} c_{i_1,\ldots,i_{d-1},0}\prod_{l=1}^{d-1} x_l^{i_l}\right)$$

which means that

$$c_{i_1,\ldots,i_{d-1},0} = 0, \quad \forall |i_1+\ldots+i_{d-1}| \le k.$$

This means we may rewrite

$$\begin{aligned} P(x) &= L(x)\underbrace{\left(\sum_{i_d=1}^{k}\sum_{|i_1+\ldots+i_{d-1}|\le k-i_d} c_{i_1,\ldots,i_{d-1},i_d} x_d^{i_d-1}\prod_{l=1}^{d-1} x_l^{i_l}\right)}, \\ P(x) &= \underbrace{x_d}_{L(x)}\underbrace{\left(\sum_{i_d=0}^{k-1}\sum_{|i_1+\ldots+i_{d-1}|\le k-i_d} c_{i_1,\ldots,i_{d-1},i_d} x_d^{i_d-1}\prod_{l=1}^{d-1} x_l^{i_l}\right)}_{Q(x)}, \end{aligned} \tag{2.1}$$

with $\deg(Q) = k-1$.

Definition 2.4: (Lagrange elements on triangles)

The triangular Lagrange element of degree k, $(K,\mathcal{P},\mathcal{N})$ denoted $\mathbb{P}_k$, is defined as follows.

1. K is a (non-degenerate) triangle with vertices z_1, z_2, z_3
2. $\mathcal{P}$ is the space of degree k polynomials on K.
3. $\mathcal{N} = \{N_{i,j} : 0 \le i \le k, 0 \le j \le i\}$ defined by $N_{i,j}(v) = v(x_{i,j})$ where,

$$x_{i,j} = z_1 + (z_2 - z_1)\frac{i}{k} + (z_3 - z_1)\frac{j}{k}.$$

Example 2.2: (P1 elements on triangles)

The nodal basis for $P1$ elements is point evaluation at the three vertices, i.e

$$\mathcal{N}_i(v) = v(z_i), \quad i = 1,2,3$$

Lemma 2.3

The $P1$-Lagrange element on a triangle K is a finite element.

Proof

Let L_1,L_2 and L_3 be the three lines containing the vertices. Suppose that a polynomial $p \in \mathcal{P}$ vanishes at z_1, z_2 and z_3. Since $p_{|L_1}$ is a linear function of one variable that vanishes at two points, $p = 0$ on L_1. By Lemma 2.2 we can write $P = cL_1$, where c is a constant. But

$$0 = p(z_1) = cL_1(z_1) \implies c = 0$$

Example 2.3: (P2 elements on triangles)

The nodal basis for $P2$ elements is point evaluation at the three vertices, plus point evaluation at the three edge centres, i.e

$$\begin{aligned} \mathcal{N}_i(v) &= v(z_i), \qquad i = 1,2,3 \\ \mathcal{N}_{i+3}(v) &= v(z_{i+3}), \quad i = 1,2,3, \quad z_{i+3} \text{ are the three edge centres} \end{aligned}$$

Lemma 2.4

The $P2$-Lagrange element on a triangle K is a finite element.

Proof

We need to check that $\mathcal{N}$ determines $\mathcal{P}$. As before, let L_1, L_2 and L_3 be non-trivial linear functions that define the edges of the triangle. Suppose that the polynomial $p \in \mathcal{P}_2$ vanishes at $z_1, z_2, \ldots, z_6$. Since $p_{|L_1}$ is a quadratic function of one variable that vanishes at three points, $p = 0$ on L_1. By Lemma2.2 we can write $P = L_1 Q_1$ where $\deg Q_1 = 1$. But p also vanishes on L_2. Therefore, $(L_1 Q_1)_{|L_2} = 0$. Hence, on L_2, either $L_1 = 0$ or $Q_1 = 0$. But L_1 can equal zero only at one point of L_2 since we have a non-degenerate triangle. Therefore, $Q_1 = 0$ on L_2, except possibly at one point. By continuity, we have $Q_1 = 0$ on L_2. By Lemma 2.2, we can write $Q_1 = L_2 Q_2$, where deg $Q_2 = 0$. Then we can write $P = cL_1L_2$. But $p(z_6) = 0$ and z_6 does not lie on either L_1 or L_2. Therefore,

$$0 = p(z_6) = cL_1(z_6)L_2(z_6)$$

since $L_1(z_6) \neq 0$ and $L_2(z_6) \neq 0$ then $c = 0$. Thus $p = 0$.

Definition 2.5: (Cubic Hermite elements on triangles)

The cubic Hermite element is defined as follows:

i) $\mathcal{K}$ is a (nondegenerate) triangle,
ii) $\mathcal{P}$ is the space of cubic polynomials on K
iii) $\mathcal{N} = \{N_1, N_2, \ldots, N_{10}\}$ defined as follows:

- (N_1, N_2, N_3) evaluation of p at vertices
- $(N_4, \ldots, N_9)$ evaluation of the gradient of p at the 3 triangle vertices.
- N_{10} evaluation of p at the centre of the triangle.

Lemma 2.5

The cubic Hermite element is a finite element.

Proof

Let L_1, L_2 and L_3 again be non-trivial linear functions that define the edges of the triangle. Suppose that for a polynomial $p \in \mathcal{P}_3, N_i(p) = 0$ for $i = 1, 2, \ldots, 10$. Restricting p to L_1, we see that z_2 and z_3 are double roots of p since $p(z_2) = 0, p'(z_2) = 0$ and $p(z_3) = 0, p'(z_3) = 0$, where ' denotes differentiation along the straight line L_1. But the only third order polynomial in one variable with four roots is the zero polynomial, hence hence $p = 0$ along L_1. Similarly, $p = 0$ along L_2 and L_3. We can, therefore, write

$$p = cL_1L_2L_3$$

2.1.2 Global continuity

Next we need to know how to glue finite elements together to form spaces defined over a triangulation (mesh). To do this we need to develop a language for specifying connections between finite element functions between element domains

Definition 2.6

A finite element space V_h is a C^m finite element space if $u \in C^m$ for all $u \in V$.

2.2 The Interpolant

Now that we have examined a number of finite elements, we wish to piece them together to create subspaces of Sobolev spaces. We begin by defining the (local) interpolant

Definition 2.7

Given a finite element $(K,\mathcal{P},\mathcal{N})$. Let $\{\phi_i, \quad i = 1, \ldots, k\}$ be the basis dual to $\mathcal{N}$. If v is a function for which all $N_i, i = 1, \ldots, k$, are defined, then we define the local interpolant by

$$\mathcal{I}_\mathcal{K}(v)(x) := \sum_{i=1}^{k} \mathcal{N}_i(v)\phi_i(x)$$

Proposition 2.1

The local interpolant $\mathcal{I}_\mathcal{K}$ is linear.

Proposition 2.2

$$N_i(\mathcal{I}_\mathcal{K}(v)) = N_i(v)$$

Proof

$N_i(\phi_j) = \delta_{ij}$

References

1. G. Allaire. Analyse Numérique et Optimisation. Éditions de l'École Polytechnique 2012.
2. Babuska I. Courant element: before and after. Lecture Notes in Pure and Applied Mathematics, vol. 164. Marcel Dekker: New York, 1994; 37-51.
3. H. Brezis. Analyse fonctionnelle. Théorie et applications. Masson, Paris, 1991.
4. F. Brezzi. On the existence, uniqueness and approximation of saddle point problems arising from Lagrange multipliers. RAIRO Anal. Numer. (1974) 129-151.
5. S. Brenner and R. Scott. The Mathematical Theory of Finite Element Methods. Third Edition 2008 Springer.
6. Z. Chen and Ricardo H. Nochetto: Residual type a posteriori error estimates for elliptic obstacle problems. Numer. Math. (2000) 84: 527-548
7. A. Ern et J.L Guermond. Éléments finis: théorie, applications, mise en oeuvre. Springer 2002.
8. P. Ciarlet. The Finite Element Method for Elliptic Problems. NorthHolland, Amsterdam, 1978.
9. Courant R. Variational methods for the solution of problems of equilibrium and vibration. Bull. Am. Math. Soc. 1943; 49:1-23.
10. Dautray, R. et Lions, J.-L., Analyse mathématique et calcul numérique pour les sciences et les techniques, Vol. 5, Masson, Paris, 1985.
11. V. Girault et P-A. Raviart. Finite element methods for Navier-Stokes equations. Springer-Verlag, Berlin, 1986.
12. Gustafsson, T, Stenberg, R and Videman, J. A stabilised finite element method for the plate obstacle problem. BIT Numerical Mathematics. 59, pages: 97-124 (2019) Springer.
13. P. D Lax and A.N Miligram. Parabolic equations. In Contributions to the theory of partial differential equations, Annals of Mathematics Studies, no. 33, pp. 167-190. Princeton University Press, Princeton, NJ, 1954.
14. H. Le Dret. Équations aux dérivées partielles elliptiques non linéaires. Springer 2013.
15. J. Nečas. Les méthodes directes en théorie des équations elliptiques, Masson, Paris, 1967.
16. R. Verfurth. A Posteriori Error Estimation Techniques for Finite Element Methods
17. J.T. Oden and J.N Reddy. An Introduction to the Mathematical Theory of Finite Elements. Wiley-Interscience, New York, 1976
18. Wachspress, E.L. A Rational Finite Element Basis. Academic Press, New York, 1975
19. K. Yosida. Functional Analysis. Springer-Verlag,1980.