Chapter 2

MATHEMATICAL PRELIMINARIES

2.1 Linear function spaces, operators and functionals

A *set* $\mathcal{U}$ is a collection of objects, referred to as *elements* or *points*. If u is an element of the set $\mathcal{U}$, one writes $u \in \mathcal{U}$. If not, one writes $u \notin \mathcal{U}$. Let $\mathcal{U}$, $\mathcal{V}$ be two sets. The set $\mathcal{U}$ is a *subset* of the set $\mathcal{V}$ (denoted as $\mathcal{U} \subseteq \mathcal{V}$ or $\mathcal{V} \supseteq \mathcal{U}$) if every element of $\mathcal{U}$ is also an element of $\mathcal{V}$. The set $\mathcal{U}$ is a *proper subset* of the set $\mathcal{V}$ (denoted as $\mathcal{U} \subset \mathcal{V}$ or $\mathcal{V} \supset \mathcal{U}$) if every element of $\mathcal{U}$ is also an element of $\mathcal{V}$, but there exists at least one element of $\mathcal{V}$ that does not belong to $\mathcal{U}$.

Consider a set $\mathcal{V}$ whose members can be scalars, vectors or functions, as visualized in Figure 2.1. Assume that $\mathcal{V}$ is endowed with an addition operation $(+)$ and a scalar multiplication operation $(\cdot)$, which do not necessarily coincide with the classical addition and multiplication for real numbers.

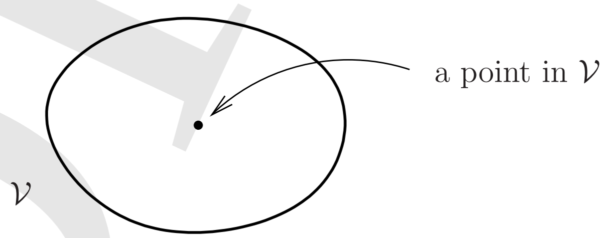

Figure 2.1: *Schematic depiction of a set* $\mathcal{V}$

A *linear* (or *vector*) *space* $\{\mathcal{V}, +; \mathbb{R}, \cdot\}$ is defined by the following properties for any

$u, v, w \in \mathcal{V}$ and $\alpha, \beta \in \mathbb{R}$:

(i) $\alpha \cdot u + \beta \cdot v \in \mathcal{V}$ (closure),

(ii) $(u+v)+w = u+(v+w)$ (associativity with respect to +),

(iii) $\exists\ 0 \in \mathcal{V} \quad | \quad u+0=u$ (existence of null element),

(iv) $\exists\ -u \in \mathcal{V} \quad | \quad u+(-u)=0$ (existence of negative element),

(v) $u+v=v+u$ (commutativity),

(vi) $(\alpha\beta)\cdot u = \alpha\cdot(\beta\cdot u)$ (associativity with respect to $\cdot$),

(vii) $(\alpha+\beta)\cdot u = \alpha\cdot u+\beta\cdot u$ (distributivity with respect to $\mathbb{R}$),

(viii) $\alpha\cdot(u+v)=\alpha\cdot u+\alpha\cdot v$ (distributivity with respect to $\mathcal{V}$),

(ix) $1\cdot u = u$ (existence of identity).

Examples:

(a) $\mathcal{V} = \mathbb{P}_2 = \{\text{all second degree polynomials } ax^2+bx+c\}$ with the standard polynomial addition and scalar multiplication.

It can be trivially verified that $\{\mathbb{P}_2, +; \mathbb{R}, \cdot\}$ is a linear function space. $\mathbb{P}_2$ is also "equivalent" to an <u>ordered</u> triad $(a,b,c)\in\mathbb{R}^3$.

(b) Define $\mathcal{V} = \{(x,y)\in\mathbb{R}^2 \mid x^2+y^2=1\}$ with the standard addition and scalar multiplication for vectors. Note that given $u=(x_1,y_1)$ and $v=(x_2,y_2)$ as in Figure 2.2, property (i) is

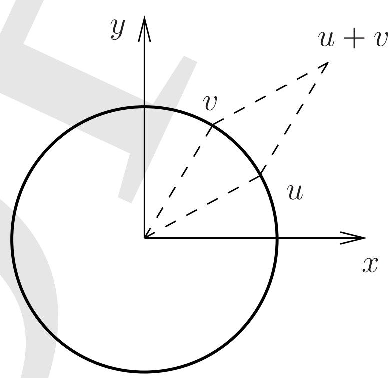

Figure 2.2: *Example of a set that does not form a linear space*

violated, i.e., since in general, for $\alpha = \beta = 1$

$$u + v = (x_1 + x_2 \ , \ y_1 + y_2) \ ,$$

and $(x_1 + x_2)^2 + (y_1 + y_2)^2 \neq 1$. Thus, $\{\mathcal{V}, +; \mathbb{R}, \cdot\}$ is not a linear space.

Consider a linear space $\{\mathcal{V}, +; \mathbb{R}, \cdot\}$ and a subset $\mathcal{U}$ of $\mathcal{V}$. Then $\mathcal{U}$ forms a *linear subspace* of $\mathcal{V}$ with respect to the same operations $(+)$ and $(\cdot)$, if, for any $u, v \in \mathcal{U}$ and $\alpha, \beta, \in \mathbb{R}$

$$\alpha \cdot u \ + \ \beta \cdot v \ \in \ \mathcal{U} \ ,$$

i.e., closure is maintained within $\mathcal{U}$.

Example:

(a) Define the set $\mathbb{P}_n$ of all algebraic polynomials of degree smaller or equal to $n > 2$ and consider the linear space $\{\mathbb{P}_n, +; \mathbb{R}, \cdot\}$ with the usual polynomial addition and scalar multiplication. Then, $\mathbb{P}_2$ is a linear subspace of $\{\mathbb{P}_n, +; \mathbb{R}, \cdot\}$.

Let $\mathcal{U}$, $\mathcal{V}$ be two sets and define a *mapping* f from $\mathcal{U}$ to $\mathcal{V}$ as a rule that assigns to each point $u \in \mathcal{U}$ a unique point $f(u) \in \mathcal{V}$, see Figure 2.3. The usual notation for a mapping is:

$$f \ : \ u \in \mathcal{U} \rightarrow f(u) \in \mathcal{V} \ .$$

With reference to the above setting, $\mathcal{U}$ is called the *domain* of f, whereas $\mathcal{V}$ is termed the *range* of f.

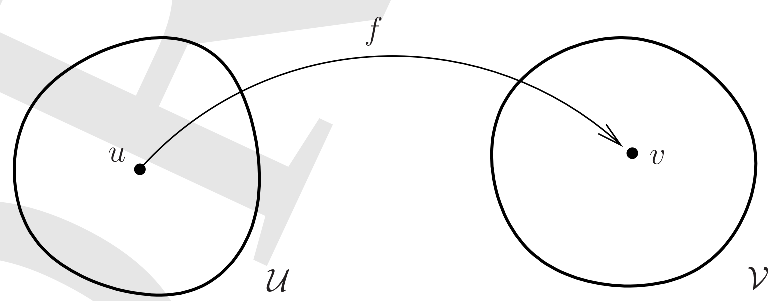

Figure 2.3: *Mapping between two sets*

The above definitions are general in that they apply to completely general types of sets $\mathcal{U}$ and $\mathcal{V}$. By convention, the following special classes of mappings are identified here:

(1) *function*: a mapping from a set with scalar or vector points to scalars or vectors, i.e.,

$$f \;:\; x \in \mathcal{U} \rightarrow f(x) \in \mathbb{R}^m \quad ; \quad \mathcal{U} = \mathbb{R}^n \quad , \quad n, m \in \mathbb{N} \ldots \ ,$$

(2) *functional*: a mapping from a set with function points (namely points that correspond to functions) to the real numbers, i.e.,

$$I \;:\; u \in \mathcal{U} \rightarrow I[u] \in \mathcal{V} \subset \mathbb{R} \quad ; \quad \mathcal{U} \text{ a function space} \ .$$

(3) *operator*: a mapping from a set of functions to another set of functions, i.e.

$$A \;:\; u \in \mathcal{U} \rightarrow A[u] \in \mathcal{V} \quad ; \quad \mathcal{U}, \mathcal{V} \text{ function spaces} \ .$$

The preceding distinction between functions, functionals and operators is largely arbitrary: all of the above mappings can be classified as operators by viewing $\mathbb{R}^n$ as a simple function space. However, the distinction will be observed for didactic purposes.

Examples:

(a) $f(\mathbf{x}) = \sqrt{x_1^2 + x_2^2}$ is a function, where $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$.

(b) $I[u] = \displaystyle\int_0^1 u(x)dx$ is a functional, where u belongs to a function space, say $u(x) \in \mathbb{P}_n$.

(c) $A[u] = \dfrac{d}{dx}u(x)$ is a (differential) operator where $u(x) \in \mathcal{U}$, where $\mathcal{U}$ is a function space.

Given a linear space $\mathcal{U}$, an operator $A \;:\; \mathcal{U} \rightarrow \mathcal{V}$ is called *linear*, provided that, for all $u_1, u_2 \in \mathcal{U}$ and $\alpha, \beta \in \mathbb{R}$,

$$A[\alpha \cdot u_1 + \beta \cdot u_2] = \alpha \cdot A[u_1] + \beta \cdot A[u_2] \ .$$

Linear partial differential equations can be formally obtained as mappings of an appropriate function space to another, induced by the action of linear differential operators. For example, consider a linear second-order partial differential equation of the form

$$a u_{,xx} + b u_{,x} \;=\; c \ ,$$

where a,b and c are functions of x and y only. The operational form of the above equation is written as

$$A[u] \;=\; c \ ,$$

where the linear differential operator A is defined as

$$A[\,\cdot\,] \;=\; a(\,\cdot\,)_{,xx} \;+\; b(\,\cdot\,)_{,x}$$

over a space of functions $u(x)$ that possess second derivatives in the domain of analysis.

2.2 Continuity and differentiability

Consider a real function $f : \mathcal{U} \rightarrow \mathbb{R}$, where $\mathcal{U} \subset \mathbb{R}$. The function f is *continuous at a point* $x = x_0$ if, given any scalar $\epsilon > 0$, there exists a scalar $\delta(\epsilon)$, such that

$$|\; f(x) \;-\; f(x_0)\;| < \epsilon \;, \tag{2.1}$$

provided that

$$|\; x - x_0 \;| < \delta \;. \tag{2.2}$$

The function f is called *continuous*, if it is continuous at all points of its domain. A function f is of class $C^k(\mathcal{U})$ (k integer ≥ 0) if it is k-times continuously differentiable (i.e., it possesses derivatives to k-th order and they are continuous functions).

Examples:

(a) The function $f :\; (0,2) \rightarrow \mathbb{R}$ defined as

$$f(x) \;=\; \begin{cases} x & \text{if } 0 < x < 1 \\ 2 - x & \text{if } 1 \leq x < 2 \end{cases}$$

is of class $C^0(\mathcal{U})$, but not of $C^1(\mathcal{U})$, see Figure 2.4.

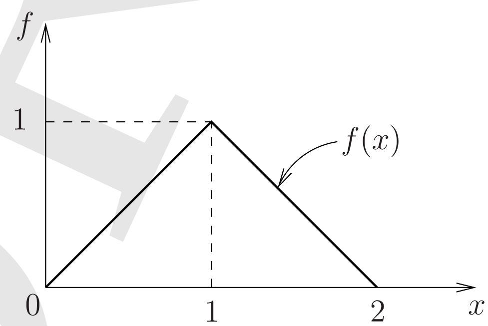

Figure 2.4: *A function of class* $C^0(0,2)$

(b) Any polynomial function $P(x) :\; \mathcal{U} \rightarrow \mathbb{R}$ is of class $C^\infty(\mathcal{U})$.

The above definition can be easily generalized to certain subsets of $\mathbb{R}^n$: a function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is of class $C^k(\mathcal{U})$ if all of its partial derivatives up to k-th order are continuous. Further generalizations to operators will be discussed later.

The "smoothness" (i.e., the degree of continuity) of functions plays a significant role in the proper construction of finite elements approximations.

2.3 Inner products, norms and completeness

2.3.1 Inner products

Consider a linear space $\{\mathcal{V}, + \,;\, \mathbb{R}, \cdot\}$ and define a mapping $\langle \cdot \,,\, \cdot \rangle : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$, such that for all u, v and $w \in \mathcal{V}$ and $\alpha \in \mathbb{R}$, the following properties hold:

(i) $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$,

(ii) $\langle u, v \rangle = \langle v, u \rangle$,

(iii) $\langle \alpha \cdot u, v \rangle = \alpha \langle u, v \rangle$,

(iv) $\langle u, u \rangle \geq 0$ and $\langle u, u \rangle = 0 \Leftrightarrow u = 0$.

A mapping with the above properties is called an *inner product* on $\mathcal{V} \times \mathcal{V}$. A linear space $\mathcal{V}$ endowed with an inner product is called an *inner product space*. If two elements u, v of $\mathcal{V}$ satisfy the condition $\langle u, v \rangle = 0$, then they are *orthogonal* relative to the inner product $\langle \cdot \,,\, \cdot \rangle$.

Examples:

(a) Set $\mathcal{V} = \mathbb{R}^n$ and for any vectors $\mathbf{x} = [x_1 \; x_2 \; \ldots \; x_n]^T$ and $\mathbf{y} = [y_1 \; y_2 \; \ldots \; y_n]^T$ in $\mathcal{V}$, define the mapping

$$\langle \mathbf{x}, \mathbf{y} \rangle \;=\; \mathbf{x}^T \mathbf{y} = \sum_{i=1}^{n} x_i y_i \;.$$

This is the conventional dot-product between vectors in $\mathbb{R}^n$. It is easy to show that the above mapping is an inner product on $\mathcal{V} \times \mathcal{V}$. This inner product-space in called the n-dimensional Euclidean vector space.

(b) The L_2-inner product for functions $u, v \in \mathcal{U}$ with domain Ω is defined as

$$\langle u, v \rangle \;=\; \int_\Omega uv \, d\Omega \;.$$

2.3.2 Norms

Recall the classical definition of distance (in the Euclidean sense) between two points in $\mathbb{R}^2$. Given any two points $\mathbf{x}_1 = (x_1, y_1)$ and $\mathbf{x}_2 = (x_2, y_2)$ as in Figure 2.5, define the "distance" function $d: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}_0^+$ as

$$d(\mathbf{x}_1, \mathbf{x}_2) \;=\; \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}\;. \tag{2.3}$$

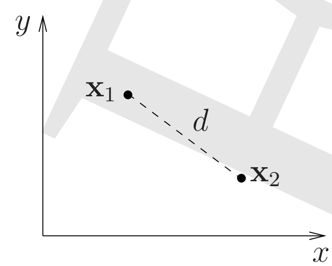

Figure 2.5: *Distance between two points in the classical Euclidean sense*

It is important to establish a similar notion of proximity ("closeness") between functions rather than merely between points in a Euclidean space. Moreover, we need to quantify the "largeness" of a function. The appropriate context for the above is provided by norms.

Consider a linear space $\{\mathcal{V}, + \,;\, \mathbb{R}, \cdot\}$ and define a mapping $\|\cdot\|: \mathcal{V} \to \mathbb{R}$ such that, for all $u, v \in \mathcal{V}$ and $\alpha \in \mathbb{R}$, the following properties hold:

(i) $\|u + v\| \;\leq\; \|u\| + \|v\|$ (triangular inequality),

(ii) $\|\alpha \cdot u\| \;=\; |a| \|u\|$,

(iii) $\|u\| \;\geq\; 0$ and $\|u\| \;=\; 0 \;\Leftrightarrow\; u \;=\; 0$.

A mapping with the above properties is called a *norm* on $\mathcal{V}$. A linear space $\mathcal{V}$ endowed with a norm is called a *normed linear space* (NLS).

Examples:

(a) Consider the n-dimensional Euclidean space $\mathbb{R}^n$ and let $\mathbf{x} = [x_1 \; x_2 \; \ldots \; x_n]^T \in \mathbb{R}^n$. Some standard norms in $\mathbb{R}^n$ are defined as follows:

- the 1-norm: $\|\mathbf{x}\|_1 \;=\; \sum_{i=1}^n |x_i|$,
- the 2-norm: $\|\mathbf{x}\|_2 \;=\; \left(\sum_{i=1}^n x_i^2\right)^{1/2}$,

- the ∞-norm: $\|\mathbf{x}\|_\infty \;=\; \max_{1\le i\le n}|x_i|$.

(b) The L_2-norm of a square-integrable function $u \in \mathcal{U}$ with domain Ω is defined as

$$\|u\|_2 \;=\; (\int_\Omega u^2\, d\Omega)^{1/2}\ .$$

Using norms, we can quantify convergence of a sequence of functions u_n to u in $\mathcal{U}$ by referring to the distance function d between u_n and u, defined as

$$d(u_n,u) \;=\; \|u_n - u\|. \tag{2.4}$$

In particular, we say that $u_n \to u \in \mathcal{U}$ if $\forall \epsilon > 0\;\; \exists N(\epsilon)$, so that

$$d(u_n,u) < \epsilon \quad \forall\, n > N\ . \tag{2.5}$$

Typically, the limit of a convergent sequence u_n of functions in $\mathcal{U}$ is not known in advance. Indeed, consider the case of a series of approximate function solutions to a partial differential equation having an unknown (and possibly unavailable in closed form) exact solution u. A sequence u_n is called *Cauchy convergent* if, for any $\epsilon > 0$, $\exists N(\epsilon)$ such that

$$d(u_m,u_n) \;=\; \|u_m - u_n\| \;<\; \epsilon \quad \forall\, m,n > N\ . \tag{2.6}$$

Although it will not be proved here, it is easy to verify that convergence of a sequence u_n implies Cauchy convergence, but the opposite is not necessary true.

Given any point u in a normed linear space $\mathcal{U}$, one may identify the neighborhood $\mathcal{N}_r(u)$ of u with radius $r > 0$ as the set of points v for which

$$d(u,v) \;<\; r\ . \tag{2.7}$$

Then, a subset $\mathcal{V}$ of $\mathcal{U}$ is termed *open* if, for each point $u \in \mathcal{V}$, there exists a neighborhood $\mathcal{N}_r(u)$ which is fully contained in $\mathcal{V}$. The complement $\widetilde{\mathcal{V}}$ of an open set $\mathcal{V}$ (defined as the set of all points in $\mathcal{U}$ that do not belong to $\mathcal{V}$) is, by definition, a *closed* set. The *closure* of a set $\mathcal{V}$, denoted by $\bar{\mathcal{V}}$ is defined as the smallest closed set that contains $\mathcal{V}$.

Example:

(a) Consider the set of real numbers $\mathbb{R}$ equipped with the usual norm (i.e., the absolute value). The set $\mathcal{V}$ defined as $\mathcal{V} = \{x \in \mathbb{R} \;\;|\;\; 0 < x < 1\} = (0,1)$ is open.

2.3.3 Banach spaces

A linear space $\mathcal{U}$ for which every Cauchy sequence converges to "point" $u \in \mathcal{U}$ is called a *complete* space. Complete normed linear spaces are also referred to as *Banach spaces*. Complete inner product spaces are called *Hilbert spaces*. Hilbert spaces form the proper functional context for the mathematical analysis of finite element methods. The basic goal of such mathematical analysis is to establish conditions under which specific finite element approximations lead to a sequence of solutions that converge to the exact solution of the differential equation under investigation.

Hilbert spaces are also Banach spaces, while the opposite is generally not true. Indeed, the inner product of a Hilbert space induces an associated norm (called the "natural norm") given by

$$\|u\| \;=\; \langle u, u\rangle^{1/2} \,. \tag{2.8}$$

To prove that $\langle u, u\rangle^{1/2}$ is actually a norm, it is sufficient to show that the three defining properties of a norm hold. Properties (ii) and (iii) are easily verified using the fact that $\langle \cdot, \cdot \rangle$ is an inner product, i.e., for (ii)

$$\|\alpha \cdot u\| \;=\; \langle \alpha \cdot u, \alpha \cdot u\rangle^{1/2} \;=\; \big(\alpha^2 \langle u, u\rangle\big)^{1/2} \;=\; |\alpha|\|u\| \,, \tag{2.9}$$

and for (iii)

$$\|u\| \;=\; \langle u, u\rangle^{1/2} \geq 0 \quad , \quad \|u\| \;=\; \langle u, u\rangle^{1/2} \;=\; 0 \Leftrightarrow u \;=\; 0 \,. \tag{2.10}$$

Property (i) (i.e., the triangular inequality) merits more attention: in order to prove that it holds, we make use of the *Cauchy-Schwartz inequality*, which states that for any $u, v \;\in \mathcal{V}$

$$|\langle u, v\rangle| \;\leq\; \|u\|\|v\| \,. \tag{2.11}$$

To prove (2.11), first note that it holds trivially as an equality if $u = 0$ or $v = 0$. Then, define a function $F:\; \mathbb{R} \to \mathbb{R}_0^+$ as

$$F(\lambda) \;=\; \|u + \lambda \cdot v\|^2 \;\;;\;\; \lambda \in \mathbb{R} \,, \tag{2.12}$$

where u, v are arbitrary (although fixed) non-zero points of $\mathcal{V}$ and λ is a scalar. Making use of the definition of the natural norm and the inner product properties, we have

$$\begin{aligned} F(\lambda) = \langle u + \lambda \cdot v, u + \lambda \cdot v\rangle \;&=\; \langle u, u\rangle \;+\; 2\lambda\langle u, v\rangle \;+\; \lambda^2\langle v, v\rangle \\ &= \|u\|^2 \;+\; 2\lambda\langle u, v\rangle \;+\; \lambda^2\|v\|^2 \,. \end{aligned} \tag{2.13}$$

Noting that $F(\lambda) = 0$ has at most one real non-zero root (i.e., if and when $u + \lambda \cdot v = 0$), it follows that, since

$$\langle v, v\rangle\lambda \;=\; -\langle u, v\rangle \pm \sqrt{\langle u, v\rangle^2 - \|u\|^2\|v\|^2}\;, \tag{2.14}$$

the inequality

$$\langle u, v\rangle^2 \;-\; \|u\|^2\|v\|^2 \;\leq\; 0 \tag{2.15}$$

must hold, thus yielding (2.11).

Using the Cauchy-Schwartz inequality, return to property (i) of a norm and note that

$$\begin{aligned}\|u+v\|^2 =&\langle u+v, u+v\rangle \;=\; \langle u, u\rangle \;+\; 2\langle u, v\rangle \;+\; \langle v, v\rangle \;=\; \|u\|^2 \;+2\langle u, v\rangle \;+\; \|v\|^2 \\ \leq&\|u\|^2 \;+\; 2\|u\|\|v\| \;+\; \|v\|^2 \;=\; (\|u\| + \|v\|)^2\;,\end{aligned} \tag{2.16}$$

which implies that the triangular inequality holds.

In the remainder of this section some of the commonly used finite element function spaces are introduced. First, define the L_2-space of functions with domain $\Omega \subset \mathbb{R}^n$ as

$$L_2(\Omega) \;=\; \left\{ u:\; \Omega \to \mathbb{R} \quad | \quad \int_\Omega u^2 d\Omega < \infty \right\}\;. \tag{2.17}$$

The above space contains all square-integrable functions defined on Ω.

Also, define the *Sobolev space* $H^m(\Omega)$ of order m (where m is a non-negative integer) as

$$H^m(\Omega) \;=\; \{u:\; \Omega \to \mathbb{R} \quad | \quad D^\alpha u \in L_2(\Omega)\;\; \forall\, \alpha \leq m\}\;, \tag{2.18}$$

where

$$D^\alpha u \;=\; \frac{\partial^\alpha u}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \ldots \partial x_n^{\alpha_n}} \qquad , \qquad \alpha = \alpha_1 + \alpha_2 + \ldots + \alpha_n\;, \tag{2.19}$$

is the generic partial derivative of order α, and α is an non-negative integer. Using the above definitions, it is clear that $L_2(\Omega) \equiv H^0(\Omega)$. An inner product can be defined for $H^m(\Omega)$ as

$$\langle u, v\rangle_{H^m(\Omega)} \;=\; \int_\Omega \Big\{ \sum_{\alpha=0}^{m} \sum_{\beta=\alpha} D^\beta u\, D^\beta v \Big\}\, d\Omega\;, \tag{2.20}$$

and the corresponding natural norm as

$$\|u\|_{H^m(\Omega)} \;=\; \langle u, u\rangle_{H^m(\Omega)}^{1/2} \;=\; \Big(\int_\Omega \Big\{ \sum_{\alpha=0}^{m} \sum_{\beta=\alpha} \big(D^\beta u\big)^2 \Big\}\, d\Omega \Big)^{1/2} \;=\; \Big(\sum_{\alpha=0}^{m} \sum_{\beta=\alpha} \|D^\beta u\|^2_{L_2(\Omega)} \Big)^{1/2}\;. \tag{2.21}$$

Example:

(a) Assume $\Omega \subset \mathbb{R}^2$ and $m = 1$. Then

$$\langle u, v\rangle_{H^1(\Omega)} = \int_\Omega \left(uv + \frac{\partial u}{\partial x_1}\frac{\partial v}{\partial x_1} + \frac{\partial u}{\partial x_2}\frac{\partial v}{\partial x_2} \right) dx_1 dx_2 \ ,$$

and

$$\|u\|_{H^1(\Omega)} = \left[\int_\Omega \left\{ u^2 + \left(\frac{\partial u}{\partial x_1}\right)^2 + \left(\frac{\partial u}{\partial x_2}\right)^2 \right\} dx_1 dx_2 \right]^{1/2} .$$

Clearly, for the above inner product to make sense (or, equivalently, for u to belong to $H^1(\Omega)$), it is necessary that u and both of its first derivatives be square-integrable.

Standard theorems from elementary calculus guarantee that continuous functions are always square-integrable in a domain where they remain bounded. Similarly, piecewise continuous functions are also square integrable, provided that they possess a "small" number of discontinuities. The Dirac-delta function, defined on $\mathbb{R}^n$ by the property

$$\int_\Omega \delta(x_1, x_2, \ldots, x_n) f(x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \ldots dx_n \ = \ f(0, 0, \ldots, 0) \ , \tag{2.22}$$

for any continuous function f on Ω, where Ω contains the origin $(0, 0, \ldots, 0)$, is the single example of a function which is not square-integrable and may be encountered in finite element approximations.

Example:

(a) Consider the piecewise linear function $u(x)$ in Figure 2.6. Clearly, the function is square-integrable. Its derivative $\dfrac{du}{dx}$ is a Heaviside function, and is also square-integrable. However, the second derivative $\dfrac{d^2u}{dx^2}$, which is a Dirac-delta function is not square-integrable.

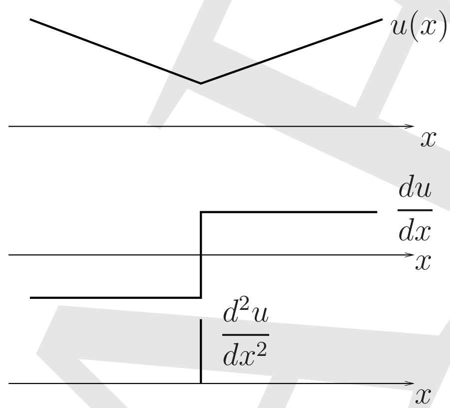

Figure 2.6: A piecewise linear function and its derivatives

Given that a function belongs to $H^m(\Omega)$, it is important to obtain an estimate of its smoothness on the sufficiently smooth boundary $\partial\Omega$ of the domain of analysis. Denoting the outward normal to the boundary by n, define (to within some technicalities) the fractional space $H^{m-j-1/2}(\partial\Omega)$ as

$$H^{m-j-1/2}(\partial\Omega) \;=\; \{\phi \text{ on } \partial\Omega \quad | \quad \exists\, u \in H^m(\Omega) \quad | \quad \gamma_j u \;=\; \phi \text{ on } \partial\Omega\} \;, \tag{2.23}$$

where the trace operator $\gamma_j \;:\; H^m(\Omega) \;\mapsto\; L_2(\partial\Omega)$ is given by

$$\gamma_j \;=\; \frac{\partial^j u}{\partial n^j} \quad , \quad 0 \le j \le m-1 \;. \tag{2.24}$$

Negative Sobolev spaces H^{-m} can also be defined and are of interest in the mathematical analysis of the finite element method. Bypassing the formal definition, one may simply note that a function u defined on Ω belongs to $H^{-1}(\Omega)$ if its anti-derivative belongs to $L_2(\Omega)$.

A formal connection between continuity and integrability of functions can be established by means of *Sobolev's lemma*. The simplest version of this theorem states that given an open set $\Omega \subset \mathbb{R}^n$ with sufficiently smooth boundary, and letting $C_b^k(\Omega)$ be the space of bounded functions of class $C^k(\Omega)$, then

$$H^m(\Omega) \;\subset\; C_b^k(\Omega) \;, \tag{2.25}$$

if, and only if, $m > k + n/2$. For example, setting $m = 2$, $k = 1$ and $n = 1$, one concludes from the preceding theorem that the space of H^2 functions on the real line is embedded in the space of bounded C^1-functions.

2.3.4 Linear operators and bilinear forms in Hilbert spaces

Consider a linear operator $A : \mathcal{U} \mapsto \mathcal{V}$, $v = A[u]$, where $\mathcal{U}$, $\mathcal{V}$ are Hilbert spaces, as in Figure 2.7. Some important definitions follow:

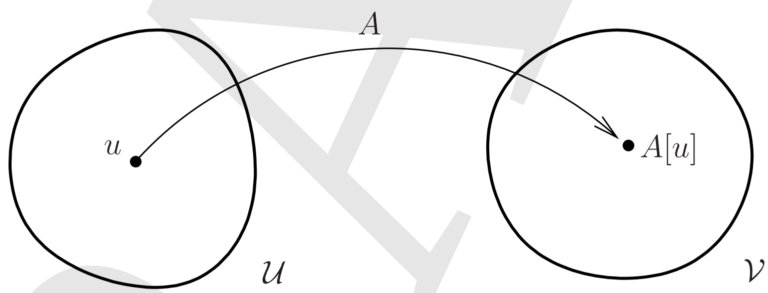

Figure 2.7: *A linear operator mapping* $\mathcal{U}$ *to* $\mathcal{V}$

A is *bounded* if there exists a constant $M > 0$, such that $\|A[u]\|_{\mathcal{V}} \leq M\|u\|_{\mathcal{U}}$, for all $u \in \mathcal{U}$. We say that M is a *bound* to the operator. Next, A is (uniformly) *continuous* if, for any $\epsilon > 0$, there is a $\delta = \delta(\epsilon)$ such that $\|A[u] - A[v]\|_{\mathcal{V}} < \epsilon$ for any $u, v \in \mathcal{U}$ that satisfy $\|u - v\|_{\mathcal{U}} < \delta$. It is easy to show that, in the context of linear operators, boundedness implies (uniform) continuity and vice-versa (i.e., the two properties are equivalent).

A linear operator $A : \mathcal{U} \mapsto \mathcal{V} \subset \mathcal{U}$ is *symmetric* relative to a given inner product $\langle \cdot, \cdot \rangle$ defined on $\mathcal{U} \times \mathcal{U}$, if

$$\langle A[u], v \rangle \;=\; \langle u, A[v] \rangle \;, \tag{2.26}$$

for all u, $v \in \mathcal{U}$.

Example:

(a) Let $\mathcal{U} = \mathbb{R}^n$ and A be an operator identified with the action of an $n \times n$ symmetric matrix $\mathbf{A}$ on an n-dimensional vector $\mathbf{x}$, so that $A[\mathbf{x}] = \mathbf{A}\mathbf{x}$. Also, define an associated inner product as

$$\langle \mathbf{x}, A[\mathbf{y}] \rangle \;=\; \mathbf{x} \cdot \mathbf{A}\mathbf{y} \;,$$

i.e., as the usual dot-product between vectors. Then

$$\begin{aligned} \langle \mathbf{x}, A[\mathbf{y}] \rangle \;&=\; \mathbf{x} \cdot \mathbf{A}\mathbf{y} = \mathbf{x} \cdot \mathbf{A}^T \mathbf{y} \\ &=\; (\mathbf{A}\mathbf{x}) \cdot \mathbf{y} \;=\; \langle A[\mathbf{x}], \mathbf{y} \rangle \end{aligned}$$

implies that A is a symmetric (algebraic) operator.

A symmetric operator A is termed *positive* if $\langle A[u], u \rangle \geq 0$, for all $u \in \mathcal{U}$.

The *adjoint* A^* of an operator A with reference to the inner product $\langle \cdot, \cdot \rangle$ on $\mathcal{U} \times \mathcal{U}$ is defined by

$$\langle A[u], v \rangle \;=\; \langle u, A^*[v] \rangle \;, \tag{2.27}$$

for all $u, v \in \mathcal{U}$. An operator A is termed *self-adjoint* if $A = A^*$. It is clear that every self-adjoint operator is symmetric, but the converse is not true.

Define an operator $B : \mathcal{U} \times \mathcal{V} \mapsto \mathbb{R}$ as in Figure 2.8, where $\mathcal{U}$ and $\mathcal{V}$ are Hilbert spaces, such that for all $u, u_1, u_2 \in \mathcal{U}$, $v, v_1, v_2 \in \mathcal{V}$ and $\alpha, \beta \in \mathbb{R}$,

(i) $B(\alpha \cdot u_1 + \beta \cdot u_2, v) = \alpha B(u_1, v) + \beta B(u_2, v)$,

(ii) $B(u, \alpha \cdot v_1 + \beta \cdot v_2) = \alpha B(u, v_1) + \beta B(u, v_2)$.

Then, B is called a *bilinear form* on $\mathcal{U} \times \mathcal{V}$. The bilinear form B is *continuous* if there is a constant $M > 0$, such that, for all $u \in \mathcal{U}$ and $v \in \mathcal{V}$,

$$|B(u, v)| \leq M \|u\|_{\mathcal{U}} \|v\|_{\mathcal{V}} \;. \tag{2.28}$$

Consider a bilinear form $B(u, v)$ and fix $u \in \mathcal{U}$. Then an operator $A_u : \mathcal{V} \mapsto \mathbb{R}$ is defined according to

$$A_u[v] \;=\; B(u, v) \quad ; \quad u \text{ fixed} \;. \tag{2.29}$$

Operator A_u is called the *formal operator* associated with the bilinear form B. Similarly, when $v \in \mathcal{V}$ is fixed in $B(u, v)$, then an operator $A_v : \mathcal{U} \mapsto \mathbb{R}$ is defined as

$$A_v[u] \;=\; B(u, v) \quad ; \quad v \text{ fixed} \;, \tag{2.30}$$

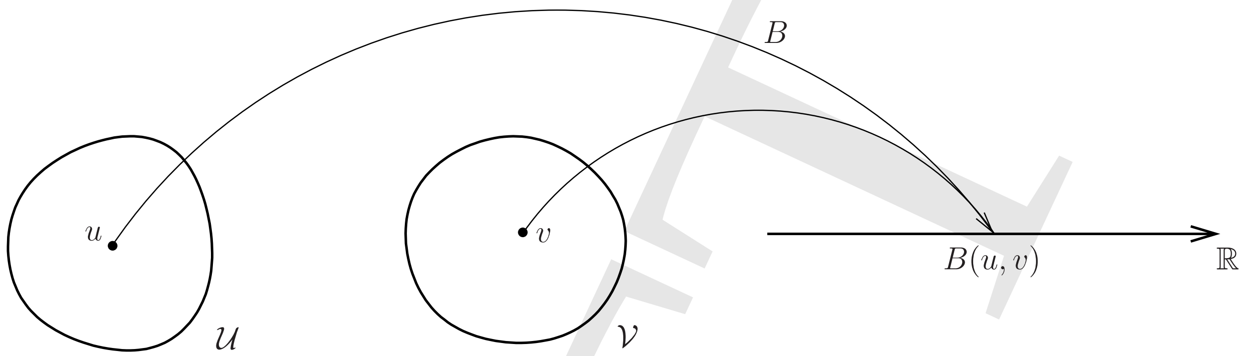

Figure 2.8: *A bilinear form on* $\mathcal{U} \times \mathcal{V}$

and is called the *formal adjoint* of A_u.

Clearly, both A_u and A_v are linear (since they emanate from a bilinear form) and are often referred to as linear forms or linear functionals.

2.4 Background on variational calculus

The solutions to partial differential equations are often associated with extremization of functionals over a properly defined space of admissible functions. This subject will be addressed in Chapter 4 of the notes. Some preliminary information on variational calculus is presented here as background to forthcoming developments.

Consider a functional $I : \mathcal{U} \mapsto \mathbb{R}$, where $\mathcal{U}$ consists of functions $u = u(x, y, \ldots)$ that can play the role of the dependent variable in a partial differential equation. The *variation* δu of u is an arbitrary function defined on the same domain as u and represents admissible changes to the function u. Thus, if $\Omega \subset \mathbb{R}^n$ is the domain of $u \in \mathcal{U}$ with boundary $\partial\Omega$, where

$$\mathcal{U} \; = \; \left\{ u \in H^1(\Omega) \quad | \quad u = \bar{u} \text{ on } \partial\Omega \right\} \; ,$$

then δu belongs to the set

$$\mathcal{U}_0 \; = \; \left\{ u \in H^1(\Omega) \quad | \quad u = 0 \text{ on } \partial\Omega \right\} \; .$$

The preceding example illustrates that the variation δu of a function u is essentially restricted only by conditions related to the definition of the function u itself.

As already mentioned, interest will be focused on the determination of functions u^*, which render the functional $I[u]$ stationary (i.e., minimum, maximum or a saddle point), as schematically indicated in Figure 2.9.

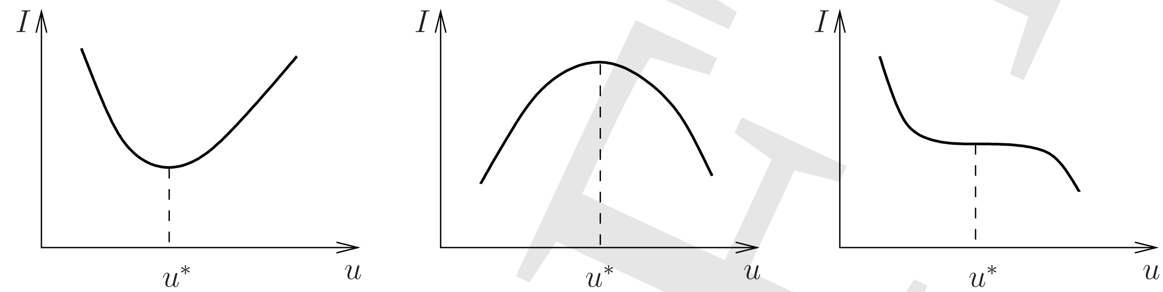

Figure 2.9: *A functional exhibiting a minimum, maximum or saddle point at* $u = u^*$

Define the (first) *variation* $\delta I[u]$ of $I[u]$ as

$$\delta I[u] = \lim_{w \to 0} \frac{I[u + w\delta u] \; - \; I[u]}{w} \; , \tag{2.31}$$

and, by induction, the k-th variation as

$$\delta^k I[u] \;=\; \delta(\delta^{k-1} I[u]) \; , \quad k = 2, 3, \ldots \; . \tag{2.32}$$

Alternatively, the variations of $I[u]$ can be determined by first expanding $I[u + \delta u]$ around u and then forming $\delta^k I[u]$, $k = 1, 2, \ldots$, from all terms that involve only the k-th power of δu, according to

$$I[u \; + \; \delta u] \;=\; I[u] \;+\; \delta I[u] \;+\; \frac{1}{2!}\delta^2 I[u] \;+\; \frac{1}{3!}\delta^3 I[u] \;+\; \ldots \; . \tag{2.33}$$

Examples:

(a) Let I be defined on $\mathbb{R}^n$ as

$$I[\mathbf{u}] \;=\; \frac{1}{2}\mathbf{u}\cdot\mathbf{A}\mathbf{u} \;-\; \mathbf{u}\cdot\mathbf{b}\;,$$

where $\mathbf{A}$ is an $n\times n$ symmetric positive-definite matrix and $\mathbf{b}$ belongs to $\mathbb{R}^n$. Using (2.31), it follows that

$$\delta I[\mathbf{u}] \;=\; \delta\mathbf{u}\cdot\mathbf{A}\mathbf{u} \;-\; \delta\mathbf{u}\cdot\mathbf{b}$$

and

$$\delta^2 I[\mathbf{u}] \;=\; \delta\mathbf{u}\cdot\mathbf{A}\delta\mathbf{u}\;.$$

Therefore, it is seen that minimization of the above functional yields a system of n linear algebraic equations with n unknowns. Since $\mathbf{A}$ is assumed positive-definite, the system has a unique solution

$$\mathbf{u} \;=\; \mathbf{A}^{-1}\mathbf{b}\;,$$

which coincides with the minimum of $I[\mathbf{u}]$. Several iterative methods for the solution of linear algebraic systems effectively exploit this minimization property.

(b) The variations of functional $I[u]$ defined as

$$I[u] \;=\; \int_0^1 u^2\,dx$$

can be determined by directly using (2.31). Thus,

$$\begin{aligned}\delta I[u] \;&=\; \lim_{\omega\to 0}\frac{\int_0^1\left[(u\;+\;\omega\delta u)^2\;-\;u^2\right]dx}{\omega}\\ &=\; \lim_{\omega\to 0}\int_0^1\left[2u\,\delta u\;+\;\omega(\delta u)^2\right]dx \;=\; \int_0^1 2u\,\delta u\,dx\;,\end{aligned}$$

$$\begin{aligned}\delta^2 I[u] \;&=\; \delta(\delta I[u]) \;=\; \lim_{\omega\to 0}\frac{\int_0^1\left[2(u\;+\;\omega\delta u)\,\delta u\;-\;2u\,\delta u\right]dx}{\omega}\\ &=\; 2\int_0^1(\delta u)^2\,dx\end{aligned}$$

and

$$\delta^k I[u] \;=\; 0\;,\qquad k=3,4,\ldots\;.$$

Using the alternative definition (2.33) for the variations of $I[u]$, write

$$\begin{aligned}I[u+\delta u] = \int_0^1 (u+\delta u)^2\,dx \;&=\; \int_0^1 u^2\,dx + 2\int_0^1 u\delta u\,dx + \int_0^1(\delta u)^2\,dx\\ &=\; I[u]+\delta I[u]+\frac{1}{2!}\delta^2 I[u]\;,\end{aligned}$$

leading again to the expressions for $\delta^k I[u]$ determined above.

Remarks:

- In the variation of $I[u]$, the independent variables $x, y, \ldots$ that are arguments of u remain "frozen", since the variation is taken over the functions u themselves and not over the variables of their domain.

- Standard operations from differential calculus also apply to variational calculus, e.g., for any two functionals I_1 and I_2 defined on the same function space and any scalar constants α and β,

$$\begin{aligned} \delta(\alpha I_1 + \beta I_2) &= \alpha\, \delta I_1 + \beta\, \delta I_2 \ , \\ \delta(I_1 I_2) &= \delta I_1\, I_2 + I_1\, \delta I_2 \ . \end{aligned}$$

- Differentiation/integration and variation are operations that generally commute, i.e., for $u = u(x)$,

$$\delta \frac{du}{dx} = \frac{d}{dx}(\delta u) \ ,$$

  assuming continuity of $\dfrac{du}{dx}$, and

$$\delta \int_\Omega u\, dx = \int_\Omega \delta u\, dx \ ,$$

  assuming that the domain of integration Ω is independent of u.

- If a functional I depends on functions $u, v, \ldots$, then the variation of I obviously depends on the variations of all $u, v, \ldots$, i.e.,

$$\delta I[u, v, \ldots] = \lim_{\omega \to 0} \frac{I[u + \omega\delta u, v + \omega\delta v, \ldots] - I[u, v, \ldots]}{\omega} \ ,$$

  and

$$\delta^k I[u, v, \ldots] = \delta(\delta^{k-1} I[u, v, \ldots]) \ , \quad k = 2, 3, \ldots$$

  or, alternatively,

$$I[u + \delta u, v + \delta v, \ldots] = I[u, v, \ldots] + \delta I[u, v, \ldots] + \frac{1}{2!}\delta^2 I[u, v, \ldots] + \ldots \ .$$

- If a functional I depends on both u and its derivatives u', u'', $\ldots$, then the variation of I also depends on the variation of all u', u'', $\ldots$, namely

$$\delta I[u, u', u'', \ldots] = \lim_{\omega \to 0} \frac{I[u + \omega\delta u, u' + \omega\delta u', u'' + \omega\delta u'', \ldots] - I[u, u', u'', \ldots]}{\omega} \ .$$

Let u^* be a function that extremizes $I[u]$, and write for any variation δu around u^*

$$I[u^* + \delta u] \;=\; I[u^*] \;+\; \delta I[u^*] \;+\; \frac{1}{2!}\delta^2 I[u^*] \;+\; \ldots \;. \tag{2.34}$$

Equation (2.34) implies that necessary and sufficient condition for extremization of I at $u = u^*$ is that

$$\delta I[u^*] \;=\; 0 \;. \tag{2.35}$$

A weaker definition of the variation of a functional is obtained using the notion of a *directional* (or *Gâteaux*) *differential* of $I[u]$ at point u in the direction v, denoted by $D_v I[u]$ (or $DI(u, v)$). This is defined as

$$D_v I[u] \;=\; \left[\frac{d}{dw} I[u + wv]\right]_{w=0} \;. \tag{2.36}$$

For a large class of functionals, the variation $\delta I[u]$ can be interpreted as the Gâteaux differential of $I[u]$ in the direction δu.

Example:

(a) Consider a functional $I[u]$ defined as

$$I[u] \;=\; \int_\Omega u^2 \, d\Omega \;.$$

The directional derivative of I at u in the direction v is given by

$$\begin{aligned}
D_v I[u] \;&=\; \left[\frac{d}{dw}\int_\Omega (u \;+\; wv)^2 \, d\Omega\right]_{w=0} \\
&=\; \left[\frac{d}{dw}\int_\Omega \left[u^2 \;+\; w2uv \;+\; w^2 v^2\right] \, d\Omega\right]_{w=0} \\
&=\; \left[\int_\Omega \frac{d}{dw}\left[u^2 \;+\; w2uv \;+\; w^2 v^2\right] \, d\Omega\right]_{w=0} \\
&=\; \left[\int_\Omega \left[2uv \;+\; 2wv^2\right] \, d\Omega\right]_{w=0} \\
&=\; \int_\Omega 2uv \, d\Omega \;.
\end{aligned}$$

This result can be compared with that of a previous exercise, where it has been deduced that

$$\delta I[u] \;=\; \int_\Omega 2u\delta u \, d\Omega \;.$$

2.5 Exercises

Problem 1

(a) Given $\mathbf{x} \in \mathbb{R}^n$ with components $(x_1, x_2, \ldots, x_n)$, show that the functions $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$, $\|\mathbf{x}\|_2 = (\sum_{i=1}^n x_i^2\,)^{1/2}$ and $\|\mathbf{x}\|_\infty = \max_{1\leq i\leq n} |x_i|$ are norms.

(b) Determine the points of $\mathbb{R}^2$ for which $\|\mathbf{x}\|_1 = 1$, $\|\mathbf{x}\|_2 = 1$ and $\|\mathbf{x}\|_\infty = 1$ separately. Sketch on a single plot the geometric curves corresponding to the above points.

Problem 2

Compute the inner product $< u, v >_{L_2(\Omega)}$ for $u = x{+}1$ and $v = 3x^2{+}1$, given that $\Omega = [-1, 1]$.

Problem 3

(a) Write the explicit form of the inner product $< u, v >_{H^2(\Omega)}$ and the associated norm on $H^2(\Omega)$ given that $\Omega \subset \mathbb{R}^2$.

(b) Using the result of part (a), find the “distance” in the H^2-norm between functions $u = \sin x + y$ and $v = x$ for $\Omega = \{(x, y) \;\;|\;\; |x| \leq \pi,\; |y| \leq \pi\}$.

Problem 4

Show the parallelogram law for inner product spaces:

$$\|u \;+\; v\|^2 \;+\; \|u \;-\; v\|^2 \;=\; 2\|u\|^2 \;+\; 2\|v\|^2 \;.$$

Problem 5

Show that the integral I given by

$$I \;=\; \int_a^b \delta^2(x)\, dx \quad , \quad a \;<\; 0 \;<\; b$$

is not well-defined.

Hint: Use the definition of the Dirac-delta function $\delta(x)$ and construct a sequence of integrals I_n converging to I.

Problem 6

Given $\Omega = \{\mathbf{x} \in \mathbb{R}^2 \;\;|\;\; \|\mathbf{x}\|_2 \leq 1\}$, show that the function $f(\mathbf{x}) = \|\mathbf{x}\|_2^\alpha$ belongs to the Sobolev space $H^1(\Omega)$ for $\alpha > 0$. In addition, determine the range of α so that f also belong to $H^2(\Omega)$.

Problem 7

Consider the operator $A : \mathbb{R}^3 \mapsto \mathbb{R}^3$ defined as

$$A[(x_1, x_2, x_3)] = (x_1 + x_2 , 2x_1 + x_3 , x_2 - 2x_3) .$$

(a) Show that A is linear.

(b) Using the $\| \cdot \|_2$-norm for vectors, show that A is bounded and find an appropriate bound M.

(c) Use the inner product $< \cdot, \cdot >$ defined according to $< \mathbf{x}, \mathbf{y} >= \mathbf{x}^T \mathbf{y}$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$, to check the operator A for symmetry.

Problem 8

Let an operator $A : C^\infty(\Omega) \mapsto \mathcal{V} \subset C^\infty(\Omega)$ be defined as

$$A[u] = \frac{d^2}{dx^2}\big[a(x)\frac{d^2 u}{dx^2}\big] + b(x)u ,$$

where $\Omega = (0, L)$, $a(x) > 0, b(x) > 0, \forall x \in \Omega$. Show that the operator is linear, symmetric and positive, assuming that $u(0) = u(L) = 0$ and $\frac{du}{dx}(0) = \frac{du}{dx}(L) = 0$. Use the L_2-inner product in ascertaining symmetry and positiveness.

Problem 9

Determine the degree of smoothness of the real functions u_1 and u_2 by identifying the classes $C^k(\Omega)$ and $H^m(\Omega)$ to which they belong:

(a) $u_1(x) = x^n$, n integer > 0 , $\Omega = (0, s)$, $s < \infty$.

(b) The function $u_2(x)$ is defined by means of its first derivative

$$\frac{du_2}{dx}(x) = \begin{cases} x - 0.25 , & 0 < x \leq 1 \\ 0.25 - x , & 1 < x < 2 \end{cases} ,$$

such that $u_2(0) = 0$ and $u_2(2) = -1$, where $\Omega = (0, 2)$.

Problem 10

Show that if u is a real-valued function of class $C^1(\Omega)$, where $\Omega \in \mathbb{R}$, then $\delta \frac{du}{dx} = \frac{d(\delta u)}{dx}$, i.e., the operations of variation and differentiation commute.

Problem 11

Let the functional $I[u, u']$ be defined as

$$I[u, u'] = \int_0^1 \big(1 + u^2 + u'^2\big)\, dx .$$

(a) Compute the variations $\delta I[u, u']$ and $\delta^2 I[u, u']$ using the respective definitions.

(b) What is the value of the differential δI for $u = x^2$ and $\delta u = x$?

2.6 Suggestions for further reading

Sections 2.1-2.3

[1] G. Strang and G.J. Fix. *An Analysis of the Finite Element Method.* Prentice-Hall, Englewood Cliffs, 1973. [*The index of notations (p. 297) offers an excellent, albeit brief, discussion of mathematical preliminaries*].

[2] J.N. Reddy. *Applied Functional Analysis and Variational Methods in Engineering.* McGraw-Hill, New York, 1986. [*This book contains a very comprehensive and readable introduction to Functional Analysis with emphasis to applications in continuum mechanics*].

[3] T.J.R. Hughes. *The Finite Element Method; Linear Static and Dynamic Finite Element Analysis.* Prentice-Hall, Englewood Cliffs, 1987. [*Appendices 1.I and 4.I discuss concisely the mathematical preliminaries to the analysis of the finite element method*].

Section 2.4

[1] O. Bolza. *Lectures on the Calculus of Variations.* Chelsea, New York, 3rd edition, 1973. [*A classic book on calculus of variations that can serve as a reference, but not as a didactic text*].

[2] H. Sagan. *Introduction to the Calculus of Variations.* Dover, New York, 1992. [*A modern text on calculus of variations – Chapter 1 is very readable and pertinent to the present discussion of mathematical concepts*].