BASIC FUNCTIONAL ANALYSIS FOR THE OPTIMIZATION OF PARTIAL DIFFERENTIAL EQUATIONS

S. VOLKWEIN

ABSTRACT. Infinite-dimensional optimization requires – among other things – many results from functional analysis. In this script basics from functional analytic theory is reviewed. The purpose of this work is to give a summary of important facts needed to work in our research group.

1. Functional Analysis – Results and Definitions

If M is a set and $M_1 \subset M$, the symbol $M \setminus M_1$ represents the *complement* of M_1 in M, i.e. $M \setminus M_1 = \{x \in M : x \notin M_1\}$. $\overline{M}$ will always denote the *closure* of the set M, which is the smallest closed set containing in M. The *interior* of the set M, M°, is the largest open set containing in M. The *boundary* of M is the set $\partial M = \overline{M} \setminus M^\circ$. The set of ordered pairs $\{(x, y) : x \in M_1,\ y \in M_2\}$ is called the *Cartesian product* of the sets M_1 and M_2 and it is denoted $M_1 \times M_2$.

Let $f : M \to M_1$ be a function (or mapping). $f(M)$ will usually called the range of f and will denoted ran (f). The set $\{x \in M : f(x) = 0\}$ is said to be the *kernel* of f and is denoted ker (f). A function f will be called *injective* if for each $y \in$ ran (f) there is at most one $x \in M$ such that $f(x) = y$; f is called *surjective* if ran $(f) = M_1$. If f is both injective and surjective, we will say it is *bijective*.
Let $f : M \to M_1$ and $g : M_1 \to M_2$ be two functions. The *composite mapping* $r = g \circ f$ is defined by $r : M \to M_2$, $x \mapsto r(x) = g(f(x))$.

Definition 1.1. *A* (real) linear space *is a set, V, over $\mathbb{R}$, whose elements satisfy the following properties*

1) $v + w = w + v$ *for all* $v, w \in V$,
2) $v + (w + u) = (v + w) + u$ *for all* $v, w, u \in V$,
3) *There is in V a unique element, denoted by* 0 *and called the zero element, such that* $v + 0 = v$ *for each* v,
4) *To each v in V corresponds a unique element, denoted by* $-v$*, such that* $v + (-v) = 0$,
5) $\alpha\,(v + w) = \alpha\, v + \alpha\, w$ *for all* $v, w \in V$ *and* $\alpha \in \mathbb{R}$,
6) $(\alpha + \beta)\, v = \alpha\, v + \beta\, v$ *for all* $v \in V$ *and* $\alpha, \beta \in \mathbb{R}$,
7) $\alpha\,(\beta\, v) = (\alpha\beta)\, v$ *for all* $v \in V$ *and* $\alpha, \beta \in \mathbb{R}$,
8) $1 \cdot v = v$ *for all* $v \in V$,
9) $0 \cdot v = 0$ *for all* $v \in V$.

Date: January 8, 2003.
1991 *Mathematics Subject Classification.* 35Kxx, 46Axx, 46Bxx, 46Cxx, 46Exx, 49Kxx.
Key words and phrases. Functional analysis, optimality conditions, function spaces, evolution problems.

A (real) normed linear space *is a linear space, V, over $\mathbb{R}$ and a function, $\|\cdot\|_V$, from V to $\mathbb{R}$ which satisfies:*

1) $\|v\|_V \ge 0$ *for all v in V,*
2) $\|v\|_V = 0$ *if and only if $v = 0$,*
3) $\|\alpha v\|_V = |\alpha|\,\|v\|_V$ *for all v in V and α in $\mathbb{R}$,*
4) $\|v + w\|_V \le \|v\|_V + \|w\|_V$ *for all v and w in V.*

A system of sets M_α, $\alpha \in I$, is called a *covering* of the set M if M is contained as a subset of the union $\bigcup_{\alpha \in I} M_\alpha$. A subset M of a linear space V is called *compact* if every system of open sets of V which covers M contains a finite system also covering M. A subset M in a linear space V is *precompact*, if $\overline{M}$ is compact in V. Further we call $M \subset V$ *bounded*, if there exists a constant $K > 0$ such that $\|v\|_V \le K$ for all $v \in M$.

Definition 1.2. *A* linear operator *from a normed linear space $(V_1, \|\cdot\|_{V_1})$ to a normed linear space $(V_2,\ \|\cdot\|_{V_2})$ is a mapping, $\mathcal{A}$, from V_1 to V_2 which has the following property:*

$$\mathcal{A}(\alpha v + \beta w) = \alpha\,\mathcal{A}(v) + \beta\,\mathcal{A}(w) \quad \textit{for all } v, w \in V_1 \textit{ and } \alpha, \beta \in \mathbb{R}.$$

$\mathcal{A}$ is called a bounded (linear) operator *if $\mathcal{A}$ is linear and there is some $K > 0$ such that $\|\mathcal{A}(v)\|_{V_2} \le K\,\|v\|_{V_1}$ for all $v \in V_1$.*

The smallest such K is called the *norm of $\mathcal{A}$*. By (1.1) we will introduce a notation for the norm of a bounded linear operator. A sequence of elements $\{v_n\}_{n\in\mathbb{N}}$ of a normed linear space V is said to *converge (strongly)* to an element $v \in V$, $v_n \to v$, $n \to \infty$, if $\lim_{n\to\infty} \|v - v_n\|_V = 0$. The sequence $\{v_n\}_{n\in\mathbb{N}}$ is called a *Cauchy sequence* if for all $\varepsilon > 0$ there exists one $N \in \mathbb{N}$ such that

$$\|v_m - v_n\|_V < \varepsilon \quad \text{for all } m, n \ge N.$$

A normed linear space in which all Cauchy sequences converge is called *complete*. A set M in a normed linear space V is called *dense* if every $v \in V$ is a limit of elements in M. A function f from a normed linear space $(V_1, \|\cdot\|_{V_1})$ to a normed space $(V_2, \|\cdot\|_{V_2})$ is called *continuous at v* if $\|f(v_n) - f(v)\|_{V_2} \to 0$ as $n \to \infty$ whenever $\|v_n - v\|_{V_1} \to 0$ as n tends to zero. We say f is *Lipschitz-continuous* if there exists a constant $\gamma_f > 0$ such that

$$\|f(v) - f(w)\|_{V_2} \le \gamma_f\,\|v - w\|_{V_1} \quad \text{for all } v, w \in V_1.$$

f is called *locally Lipschitz continuous* if for all open and bounded $O \subset V_1$ with $\overline{O} \subset V_1$ there exists $\gamma_f = \gamma_f(O) > 0$ such that

$$\|f(v) - f(w)\|_{V_2} \le \gamma_f\,\|v - w\|_{V_1} \quad \text{for all } v, w \in O.$$

Let V_1 and V_2 be normed linear spaces. A bijection f from V_1 to V_2 which preserves the norm, i.e.,

$$\|f(v) - f(w)\|_{V_2} = \|v - w\|_{V_1} \quad \text{for all } v, w \in V_1$$

is called an *isometry*. It is automatically continuous. V_1 and V_2 are said to be *isometric* if such an isometry exists.

Let V be a normed linear space. The set $\{w \in V : \|v - w\|_V < \rho\}$ is called the *open ball*, $B(v;\rho)$, of radius ρ about the point v. A set $U(v) \subset X$ is called a *neighborhood* of $v \in U(v)$ if $B(v;\rho) \subset U(v)$ for some $\rho > 0$. Let $M \subset V$. A point v is called a *limit point* of M, if for all $\rho > 0$ $B(v;\rho) \cap (M \setminus \{v\}) \ne \emptyset$, i.e., x is a limit point of M if M contains points other than v arbitrarily near v.

Lemma 1.3. *Let $\mathcal{A}$ be linear operator between two normed linear spaces. The following properties are equivalent:*

1) *$\mathcal{A}$ is continuous at one point.*
2) *$\mathcal{A}$ is continuous at all points.*
3) *$\mathcal{A}$ is bounded.*

Proof. For the proof we refer the reader to Theorem 6.1.1 on page 97 in [15]. □

V_1 and V_2 are normed linear spaces. We define

$$L(V_1, V_2) = \{\mathcal{A} : V_1 \to V_2,\ \mathcal{A} \text{ is linear and continuous}\}.$$

Due to Lemma 1.3 linear operators in $L(V_1, V_2)$ are bounded linear operators, which we also call *continuous operators*. Let us introduce the following norm on $L(V_1, V_2)$:

$$\|\mathcal{A}\|_{L(V_1,V_2)} = \sup_{\|v\|_{V_1} \le 1} \|\mathcal{A}(v)\|_{V_2} = \sup_{\|v\|_{V_1} = 1} \|\mathcal{A}(v)\|_{V_2} \quad \text{for all } \mathcal{A} \in L(V_1, V_2). \tag{1.1}$$

Definition 1.4. *A complete normed linear space is called a* Banach space.

We mention the *inverse mapping theorem*.

Theorem 1.5. *A continuous bijection of one Banach space onto another has a continuous inverse.*

Proof. We refer the reader to Theorem III.11 on page 83 in [13]. □

Lemma 1.6. *Let B_1, B_2 be two Banach spaces and $\mathcal{A}$ belong to $L(B_1, B_2)$. The passage to the inverse $\mathcal{A} \to \mathcal{A}^{-1}$ is continuous (non-linear) mapping of $L(B_1, B_2)$ into $L(B_2, B_1)$ for the norm.*

Proof. For a proof we refer to Theorem 3 on page 321 in [5]. □

Remark 1.7. *If the perturbation $\mathcal{B}$ of $\mathcal{A}$ is sufficiently small, i.e., $\|\mathcal{B}\|_{L(B_1,B_2)} < \|\mathcal{A}\|^{-1}_{L(B_1,B_2)}$ holds, then $\mathcal{A}-\mathcal{B}$ is invertible. Let $\{\mathcal{A}_n\}_{n\in\mathbb{N}}$ be a sequence in $L(B_1, B_2)$ and $\mathcal{A} \in L(B_1, B_2)$ such that $\mathcal{A}^{-1}$ exists and*

$$\lim_{n\to\infty} \|\mathcal{A}_n - \mathcal{A}\|_{L(B_1,B_2)} = 0.$$

Thus, there exists $N \in \mathbb{N}$ with $\|\mathcal{A}_n - \mathcal{A}\|_{L(B_1,B_2)} < \|\mathcal{A}\|^{-1}_{L(B_1,B_2)}$ for all $n \ge N$. This leads to $\mathcal{A}_n^{-1}$ exists for all $n \ge N$.

Lemma 1.8. *Let V, V_1 and V_2 be normed linear spaces. Then $L(V_1, V_2)$ with the norm $\|\cdot\|_{L(V_1,V_2)}$ is a normed linear space, and a Banach space, if V_2 is complete. If $\mathcal{A} \in L(V_1, V_2)$ and $\mathcal{B} \in L(V_2, V)$ we have $\mathcal{B} \circ \mathcal{A} \in L(V_1, V)$ and*

$$\|\mathcal{B} \circ \mathcal{A}\|_{L(V_1,V)} \le \|\mathcal{B}\|_{L(V_2,V)} \|\mathcal{A}\|_{L(V_1,V_2)}.$$

Proof. Let us refer to Satz 3.3 on page 102 in [2]. □

We set $L(V) = L(V, V)$. The *identity* on V is the continuous operator $\mathcal{I}_V : V \to V$ given by $\mathcal{I}_V(v) = v$ for all v in V.

Definition 1.9. *A (real) vector space X is called* (real) inner product space *if there is a real-valued function $\langle \cdot, \cdot \rangle_X$ on $X \times X$ that satisfies the following four conditions for all $x, y, z \in X$ and $\alpha \in \mathbb{R}$:*

1) *$\langle x, x\rangle_X \ge 0$ and $\langle x, x\rangle_X = 0$ if and only if $x = 0$,*
2) *$\langle x, y + z\rangle_X = \langle x, y\rangle_X + \langle x, z\rangle_X$,*

3) $\langle x, \alpha y\rangle_X = \alpha \langle x, y\rangle_X$,
4) $\langle x, y\rangle_X = \langle y, x\rangle_X$.

The function $\langle \cdot\,, \cdot\rangle_X : X \times X \to \mathbb{R}$ *is called* (real) inner product.

Example 1.10. *Let* $\mathbb{R}^n$ *denote the set of all* n*-tupels of real numbers. We define the inner product*

$$\langle \mathbf{x}, \mathbf{y}\rangle_{\mathbb{R}^n} = \mathbf{x}^T\mathbf{y} = \sum_{j=1}^{n} x_j y_j.$$

for all $\mathbf{x} = (x_1, \dots, x_n)^T$ *and* $\mathbf{y} = (y_1, \dots, y_n)^T$ *in* $\mathbb{R}^n$.

Let X and Y be inner product spaces. The mapping $a : X \times X \to Y$ with the properties

1) $a(\alpha x + \beta y, z) = \alpha\, a(x, z) + \beta\, a(y, z)$ for all $x, y, z \in X$ and $\alpha, \beta \in \mathbb{R}$,
2) $a(x, y) = a(y, x)$ for all $x, y \in X$,
3) $|a(x, y)| \le K\|x\|_X\, \|y\|_X$ for some $K > 0$ and for all $x, y \in X$

is said to be a *(real) continuous bilinear form*. Two vectors, x and y, in an inner product space X are said to be *orthogonal* if $\langle x, y\rangle_X = 0$. A collection $\{x_i\}_{i\in\mathbb{N}}$ of vectors in X is called an *orthonormal set* if $\langle x_i, x_i\rangle_X = 1$ for all i, and $\langle x_i, x_j\rangle_X = 0$ if $i \ne j$.

Definition 1.11. *A family* $\{x_\lambda\}_{\lambda\in\Lambda}$ *(Λ an index set) is said to be total (or complete) in the Hilbert space* X *if*

$$\langle x, x_\lambda\rangle_X = 0 \text{ for all } \lambda \in \Lambda \quad\Longrightarrow\quad x = 0.$$

A total orthonormal family is called an orthonormal base.

Lemma 1.12. *Every inner product space* X *is a normed linear space with the norm* $\|x\|_X = \sqrt{\langle x, x\rangle_X}$.

Proof. Let us refer to Theorem II.2 on page 38 in [13]. □

Definition 1.13. *A complete (real) inner product space is called a* (real) Hilbert space.

Suppose X and Y are Hilbert spaces. Then the set of pairs (x, y) with $x \in X$, $y \in Y$ is a Hilbert space called the *direct sum* of the spaces X and Y and denoted by $X \oplus Y$. The natural inner product on $X \oplus Y$ is given by

$$\langle (x_1, y_1), (x_2, y_2)\rangle_{X\oplus Y} = \langle x_1, x_2\rangle_X + \langle y_1, y_2\rangle_Y \tag{1.2}$$

for all $(x_1, y_1), (x_2, y_2) \in X \oplus Y$.

Let M be a closed subspace of a given Hilbert space X. Under the natural inner product that it inherits as a subspace of X, M is a Hilbert space. We denote by $M^\perp$ the set of vectors in X which are orthogonal to M; $M^\perp$ is called the *orthogonal complement* of M. It follows from the linearity of the inner product that $M^\perp$ is a linear subspace of X. Further, we can prove, that $M^\perp$ is closed. Thus $M^\perp$ is also a Hilbert space. M and $M^\perp$ have only the zero element in common. The next Theorem 1.14 is usually called the *projection theorem*.

Theorem 1.14. *Let* X *be a Hilbert space,* M *a closed subspace. Then every* $x \in X$ *can be uniquely written* $x = z + w$ *where* $z \in M$ *and* $w \in M^\perp$.

Proof. We refer the reader to Theorem II.3 on page 42 in [13]. □

Remark 1.15. *Theorem* 1.14 *sets up a natural isomorphism between* $M \oplus M^{\perp}$ *and* X *given by*

$$(z, w) \mapsto z + w. \tag{1.3}$$

We will suppress the isomorphism and simply write $X = M \oplus M^{\perp}$. *Let us choose* $x_1, x_2 \in X$, $z_1, z_2 \in M$ *and* $w_1, w_2 \in M^{\perp}$ *such that* $x_1 = z_1 + w_1$ *and* $x_2 = z_2 + w_2$. *By (1.3) we identify* x_1 *and* (z_1, w_1) *respectively* x_2 *and* (z_2, w_2). *On the Hilbert spaces* M *and* $M^{\perp}$ *we use the inner product* $\langle \cdot, \cdot \rangle_X$ *that* M *and* $M^{\perp}$ *inherit as a subspace of* X. *Then it follows*

$$\begin{aligned} \langle x_1, x_2 \rangle_X &= \langle z_1 + w_1, z_2 + w_2 \rangle_X \\ &= \langle z_1, z_2 \rangle_X + \underbrace{\langle w_1, z_2 \rangle_X + \langle z_1, w_2 \rangle_X}_{=0} + \langle w_1, w_2 \rangle_X \\ &\overset{(1.2)}{=} \langle (z_1, w_1), (z_2, w_2) \rangle_{M \oplus M^{\perp}} . \end{aligned}$$

Thus, the inner product on $M \oplus M^{\perp}$ *given by* (1.2) *coincides with* $\langle \cdot, \cdot \rangle_X$.

An important class of bounded operators on Hilbert spaces is that of the projections.

Definition 1.16. *Let* X *be a Hilbert space. A bounded operator* $\mathcal{P}$ *into itself is said to be a* projection *if* $\mathcal{P}^2 \equiv \mathcal{P}$ *holds.* $\mathcal{P}$ *is called* orthogonal *if* $\langle x - \mathcal{P}(x), \mathcal{P}(x) \rangle_X = 0$ *for all* $x \in X$.

The following result is known as the *principle of uniform boundedness* or the *Banach-Steinhaus theorem*.

Theorem 1.17. *Let* $\{\mathcal{A}\}$ *be a set in* $L(B_1, B_2)$ *for two Banach spaces* B_1 *and* B_2. *If* $\|\mathcal{A}(x)\|_{B_2}$ *is bounded for each fixed* $x \in B_1$, *as* $\mathcal{A}$ *ranges over* $\{\mathcal{A}\}$, *then there exists* $K > 0$ *such that* $\|\mathcal{A}\|_{L(B_1, B_2)} \leq K$ *for all of* $\{\mathcal{A}\}$.

Proof. Let us refer to Theorem 6.3.1 on page 112 in [15]. □

Now we introduce the dual space of a given Banach space.

Definition 1.18. *Let* B *be a Banach space. The space* $L(B, \mathbb{R})$ *is called the* dual space *of* B *and it is denoted by* B'. *The elements of* B' *are* continuous linear functionals. *We write* $\langle f, x \rangle_{B', B}$ *for the duality pairing of* $f \in B'$ *with an element* $x \in B$.

Since $\mathbb{R}$ is complete, B' is a Banach space (Lemma 1.8). We define $\| \cdot \|_{B'} = \| \cdot \|_{L(B, \mathbb{R})}$. In the applications we will often consider the dual X' of a Hilbert space X. Let us recall the *Riesz representation theorem*. For a proof we refer to Theorem II.4 on page 43 in [13].

Theorem 1.19. *Let* X *be a Hilbert space with dual* X'. *For each* $f \in X'$, *there is a unique* $y_f \in X$ *such that* $\langle f, x \rangle_{X', X} = \langle y_f, x \rangle_X$ *for all* $x \in X$. *In addition* $\|y_f\|_X = \|f\|_{X'}$.

A bounded linear operator from a normed linear space V_1 to a normed linear space V_2 is called an *isomorphism* if it is bijective and continuous and if it possesses a continuous inverse. If it preserves the norm, it is called *isometric isomorphism*. Due to Theorem 1.5 a bounded linear operator possesses a continuous inverse if it is bijective and both V_1 and V_2 are Banach spaces. Obviously, an isometric isomorphism has norm one.

Remark 1.20. *By Theorem* 1.19 *we define the* Riesz isomorphism $\mathcal{J}_X$ *which maps the Hilbert space* X *onto its dual* X' *by* $y_f \mapsto f$ *and*

$$\langle f, x \rangle_{X',X} = \langle \mathcal{J}_X(y_f), x \rangle_{X',X} = \langle y_f, x \rangle_X$$

for all $x \in X$. *Often there is made no difference between an element* $y_f \in X$ *and the corresponding element (the* riesz representant*)* $f \in X'$. *We point out that we have* $\|\mathcal{J}_X\|_{L(X,X')} = 1$. *Thus,* $\mathcal{J}_X$ *is an isometric isomorphism. The dual space* X' *is also a Hilbert space: Due to Theorem* 1.19 *the natural inner product on* X' *is given by*

$$\langle f, g \rangle_{X'} = \langle \mathcal{J}_X^{-1}(f), \mathcal{J}_X^{-1}(g) \rangle_X \tag{1.4}$$

for all $f, g \in X'$.

If the normed linear space V has a countable dense subset it is called to be *separable*.

Proposition 1.21. *Let* X *and* Y *be Hilbert spaces. We define*

$$\begin{aligned}
&\mathcal{A} : X' \oplus Y' \to (X \oplus Y)' \\
&\langle \mathcal{A}(f,g), (x,y) \rangle_{(X\oplus Y)',X\oplus Y} = \langle f, x \rangle_{X',X} + \langle g, y \rangle_{Y',Y}
\end{aligned}$$

for all $(f,g) \in X' \oplus Y'$ *and* $(x,y) \in X \oplus Y$. *Then* $\mathcal{A}$ *is an isometric isomorphism.*

Proof. We have mentioned that $X \oplus Y$ is a Hilbert space with the inner product (1.2). Due to Remark 1.20 the dual space $(X \oplus Y)'$ is also a Hilbert space. Since elements of X' and Y' are continuous, $\mathcal{A}$ is also continuous. $\mathcal{J}_Y$ and $\mathcal{J}_{X\oplus Y}$ are the Riesz isomorphisms which map Y onto Y' respectively $X \oplus Y$ onto $(X \oplus Y)'$. For $(f,g), (\tilde{f}, \tilde{g}) \in X' \oplus Y'$ and $\alpha, \beta \in \mathbb{R}$ we obtain

$$\begin{aligned}
&\langle \mathcal{A}(\alpha(f,g) + \beta(\tilde{f},\tilde{g}), (x,y) \rangle_{(X\oplus Y)',X\oplus Y} \\
&\qquad = \langle \alpha f + \beta \tilde{f}, x \rangle_{X',X} + \langle \alpha g + \beta \tilde{g}, y \rangle_{Y',Y} \\
&\qquad = \langle \alpha \mathcal{A}(f,g) + \beta \mathcal{A}(\tilde{f},\tilde{g}), (x,y) \rangle_{(X\oplus Y)',X\oplus Y}
\end{aligned}$$

for all $(x,y) \in X \oplus Y$. Therefore, $\mathcal{A}$ is linear and belongs to $L(X' \oplus Y', (X \oplus Y)')$. Let us assume $\mathcal{A}(f,g) = 0$ for f in X' and g in Y'. Then it follows

$$\langle \mathcal{A}(f,g), (x,y) \rangle_{(X\oplus Y)',X\oplus Y} = 0$$

for all $(x,y) \in X \oplus Y$. So, $\langle f, x \rangle_{X',X} = -\langle g, y \rangle_{Y',Y}$ for all $(x,y) \in X \oplus Y$. This is only true if $(f,g) = (0,0)$. Thus, $\mathcal{A}$ in injective. Let us choose $r \in (X \oplus Y)'$. $\mathcal{A}$ is surjective if and only if there exists $(f,g) \in X' \oplus Y'$ such that $\mathcal{A}(f,g) = r$. $\tilde{x} \in X$ and $\tilde{y} \in Y$ are defined by

$$(\tilde{x}, \tilde{y}) = \mathcal{J}_{X\oplus Y}^{-1}(r)\,. \tag{1.5}$$

If we set $f = \mathcal{J}_X(\tilde{x})$ and $g = \mathcal{J}_Y(\tilde{y})$. We achieve

$$\begin{aligned}
\langle \mathcal{A}(f,g), (x,y) \rangle_{(X\oplus Y)',X\oplus Y} &= \langle \mathcal{J}_X(\tilde{x}), x \rangle_{X',X} + \langle \mathcal{J}_Y(\tilde{y}), y \rangle_{Y',Y} \\
&= \langle \tilde{x}, x \rangle_X + \langle \tilde{y}, y \rangle_Y \\
&\overset{(1.2)}{=} \langle (\tilde{x}, \tilde{y}), (x,y) \rangle_{X\oplus Y} \\
&\overset{(1.5)}{=} \langle r, (x,y) \rangle_{(X\oplus Y)',X\oplus Y}
\end{aligned}$$

for all $(x,y) \in X \oplus Y$. This imply the surjectivity of $\mathcal{A}$. Hence, $\mathcal{A}$ is a bijection. By applying Theorem 1.5 $\mathcal{A}$ is an isomorphism. Finally, we prove the isometry of

$\mathcal{A}$: Let $(f,g) \in X' \oplus Y'$. For $(x,y) = \mathcal{J}^{-1}_{X\oplus Y}(\mathcal{A}(f,g))$ the proof of the surjectivity has shown $f = \mathcal{J}_X(x)$ and $g = \mathcal{J}_Y(y)$. Further we have $\|\mathcal{J}_X(x)\|_{X'} = \|x\|_X$ and $\|\mathcal{J}_Y(y)\|_{Y'} = \|y\|_Y$ (Theorem 1.19). This yields

$$\begin{aligned}
\|\mathcal{A}(f,g)\|^2_{(X\oplus Y)'} &\overset{\text{Lem. 1.12}}{=} \langle \mathcal{A}(f,g), \mathcal{A}(f,g)\rangle_{(X\oplus Y)'} \\
&\overset{(1.4)}{=} \langle \mathcal{J}^{-1}_{X\oplus Y}\mathcal{A}(f,g), \mathcal{J}^{-1}_{X\oplus Y}\mathcal{A}(f,g)\rangle_{X\oplus Y} \\
&= \langle (x,y),(x,y)\rangle_{X\oplus Y} \\
&\overset{(1.2)}{=} \langle x,x\rangle_X + \langle y,y\rangle_Y \\
&\overset{(1.4)}{=} \langle \mathcal{J}_X(x), \mathcal{J}_X(x)\rangle_{X'} + \langle \mathcal{J}_Y(y), \mathcal{J}_Y(y)\rangle_{Y'} \\
&= \langle f,f\rangle_{X'} + \langle g,g\rangle_{Y'} \\
&\overset{(1.2)}{=} \langle (f,g),(f,g)\rangle_{X'\oplus Y'} \\
&\overset{\text{Lem. 1.12}}{=} \|(f,g)\|^2_{X'\oplus Y'}.
\end{aligned}$$

Therefore, $\mathcal{A}$ is isometric, and the proof is complete. □

Remark 1.22. *Due to Proposition 1.21 we identify $X' \oplus Y'$ with $(X \oplus Y)'$, so that we use $X' \oplus Y'$ as the dual space of $X \oplus Y$. The natural inner product on $X' \oplus Y'$ is*

$$\begin{aligned}
\langle (f,g),(\tilde{f},\tilde{g})\rangle_{X'\oplus Y'} &= \langle f,\tilde{f}\rangle_{X'} + \langle g,\tilde{g}\rangle_{Y'} \\
&\overset{(1.4)}{=} \langle \mathcal{J}_X^{-1}(f), \mathcal{J}_X^{-1}(\tilde{f})\rangle_X + \langle \mathcal{J}_Y^{-1}(g), \mathcal{J}_Y^{-1}(\tilde{g})\rangle_Y
\end{aligned}$$

for all $(f,g),(\tilde{f},\tilde{g}) \in X' \oplus Y'$.

Definition 1.23. *A sequence $\{x_n\}_{n\in\mathbb{N}}$* converges weakly *to an element x of the Hilbert space X, $x_n \rightharpoonup x$, $n \to \infty$, if we have*

$$\lim_{n\to\infty} \langle f, x_n\rangle_{X',X} = \langle f, x\rangle_{X',X}$$

for all $f \in X'$.

By the Theorem 1.19, a sequence $\{x_n\}_{n\in\mathbb{N}}$ converges weakly to an element x of the Hilbert space X if $\lim_{n\to\infty}\langle y, x_n\rangle_X = \langle y, x\rangle_X$ for all $y \in X$. The following results are useful in the Hilbert space approach to differential equations.

Lemma 1.24. *A bounded sequence in a Hilbert space contains a weakly convergent subsequence.*

Proof. Let us refer to Theorem 5.12 on page 80 in [9]. □

Lemma 1.25. *Let X be a Hilbert space. Then the weak convergence of $\{x_n\}_{n\in\mathbb{N}}$ in X implies the boundedness of $\|x_n\|_X$.*

Proof. We refer the reader to Korollar 13.3 on page 61 in [10]. □

Lemma 1.26. *Let X be a finite dimensional Hilbert space. Then every weak convergent sequence converge strongly in X.*

Proof. We choose a sequence $\{x_n\}_{n\in\mathbb{N}}$ such that $x_n \rightharpoonup x$, $n \to \infty$, for some $x \in X$. Further let $\{\varphi_1, \ldots, \varphi_N\}$ be a orthonormal base for the N-dimensional Hilbert space X and $x = \sum_{i=1}^N x^i \varphi_i$, $x_n = \sum_{i=1}^N x_n^i \varphi_i$. Hence the weak convergence and the linearity of the inner product imply

$$0 = \lim_{n\to\infty} \big|\langle x_n - x, \varphi_j\rangle_X\big| \le \lim_{n\to\infty} \sum_{i=1}^N \big|x_n^i - x^i \langle \varphi_i, \varphi_j\rangle_X\big| = \lim_{n\to\infty} \big|x_n^j - x^j\big|$$

for $j = 1, \ldots, N$. Thus,

$$0 \le \lim_{n\to\infty} \|x_n - x\|_X = \lim_{n\to\infty} \Big\| \sum_{i=1}^N (x_n^i - x^i)\varphi_i \Big\|_X \le \sum_{i=1}^N \lim_{n\to\infty} |x_n^i - x^i| = 0.$$

□

Definition 1.27. *Let V_1 and V_2 be two normed linear spaces. A linear operator $\mathcal{K} \in L(V_1, V_2)$ is called* compact *if $\mathcal{K}$ takes bounded sets in V_1 into precompact sets in V_2.*

An important property of compact operators is given by:

Lemma 1.28. *A compact operator maps weakly convergent sequences into norm convergent sequences.*

Proof. Let us refer to Theorem VI.11 on page 199 in [13]. □

The following lemma is important since one can use it to prove that an operator is compact. For a proof we refer to Theorem VI.12 on page 200 in [13].

Lemma 1.29. *Let B_1, B_2 and B be Banach spaces and $\mathcal{A} \in L(B_1, B_2)$. If $\mathcal{B} \in L(B_2, B)$ and if $\mathcal{A}$ or $\mathcal{B}$ is compact, then $\mathcal{B} \circ \mathcal{A}$ is compact.*

Let us recall the *Fredholm alternative*. For a proof we refer to the corollary on page 203 in [13].

Theorem 1.30. *If $\mathcal{K}$ is a compact operator from a Hilbert space X into itself, then either $(\mathcal{I}_X - \mathcal{K})^{-1}$ exists or $\mathcal{K}(v) = v$ has a solution.*

A further important theorem about compact operators is given by *Riesz-Schauder theorem*. Therefore we need some more definitions.

Definition 1.31. *Let B be a Banach space and $\mathcal{A} \in L(B)$. A complex number λ is said to be in the resolvent set $\rho(\mathcal{A})$ of $\mathcal{A}$ if $\lambda\mathcal{I}_B - \mathcal{A}$ is a bijection with a bounded inverse. If $\lambda \notin \sigma(\mathcal{A})$, then λ is said to be in the spectrum $\sigma(\mathcal{A})$ of $\mathcal{A}$.*

We note that by Theorem 1.5, $\lambda\mathcal{I}_B - \mathcal{A}$ automatically has a bounded inverse if it is bijective. We consider a subset of the spectrum.

Definition 1.32. *Let B be a Banach space and $\mathcal{A} \in L(B)$. An element $v \neq 0$ which satisfies $\mathcal{A}(v) = \lambda v$ for some $\lambda \in \mathbb{C}$ is called an* eigenvector *(or* eigenfunction*) of $\mathcal{A}$; λ is called the corresponding* eigenvalue*.*

Theorem 1.33. *Let $\mathcal{K}$ be a compact operator on a Hilbert space X, then $\sigma(\mathcal{K})$ is a discrete set having no limit points except perhaps $\lambda = 0$. Further any nonzero $\lambda \in \sigma(\mathcal{K})$ is an eigenvalue of finite multiplicity (i.e., the corresponding space of eigenvectors is finite dimensional).*

Proof. We refer to Theorem VI.15 on page 203 in [13]. □

If $x_n \rightharpoonup x$, $n \to \infty$, we do not have $\lim_{n\to\infty} \|x_n\|_X = \|x\|_X$ in general. But we get the following result.

Lemma 1.34. *Let X be a Hilbert space. If $x_n \rightharpoonup x$, $n \to \infty$, then $\|x\|_X \leq \lim\limits_{n\to\infty} \inf \|x_n\|_X$.*

Proof. Let us refer to Lemma 13.2.1 on page 351 in [15]. □

Definition 1.35. *Let X be a Hilbert space. $f : X \to \mathbb{R}$ is called* weakly lower semicontinuous *if*

$$f(x) \leq \lim_{n\to\infty} \inf f(x_n)$$

as $x_n \rightharpoonup x$, $n \to \infty$.

Remark 1.36. *From Lemma 1.34 we infer that the norm is weakly lower semicontinuous.*

Definition 1.37. *Let X and Y be two Hilbert spaces with duals X' and Y' respectively. Then we associate with every bounded operator $\mathcal{A} : X \to Y'$ defined on all of X, an* adjoint*, denoted by $\mathcal{A}^\star$ and defined by*

$$\langle \mathcal{A}^\star(y), x\rangle_{X',X} = \langle \mathcal{A}(x), y\rangle_{Y',Y} \quad \text{for all } x \in X \text{ and } y \in Y.$$

Lemma 1.38. *Let X and Y be two Hilbert spaces with duals X' and Y', respectively, and $\mathcal{A} \in L(X, Y')$ be a bounded operator; its adjoint $\mathcal{A}^\star$ has the following properties:*

1) *$\mathcal{A}^\star \in L(Y, X')$;*
2) *$\|\mathcal{A}^\star\|_{L(Y,X')} = \|\mathcal{A}\|_{L(X,Y')}$: the mapping $\mathcal{A} \mapsto \mathcal{A}^\star$ is thus an isometry of $L(X, Y')$ into $L(Y, X')$;*
3) *We have*

$$\begin{aligned}
\overline{\text{ran } (\mathcal{J}_X^{-1}\mathcal{A}^\star)} &= (\ker\,(\mathcal{J}_Y^{-1}\mathcal{A})^\perp, \\
\overline{\text{ran } (\mathcal{J}_Y^{-1}\mathcal{A})} &= \ker\,(\mathcal{J}_X^{-1}\mathcal{A}^\star)^\perp, \\
\ker\,(\mathcal{J}_X^{-1}\mathcal{A}^\star) &= \ker\,(\mathcal{J}_Y^{-1}\mathcal{A}\mathcal{J}_X^{-1}\mathcal{A}^\star), \\
\overline{\text{ran } (\mathcal{J}_Y^{-1}\mathcal{A})} &= \overline{\text{ran } \mathcal{J}_Y^{-1}\mathcal{A}\mathcal{J}_X^{-1}\mathcal{A}^\star)}.
\end{aligned}$$

4) *If any of the two subspaces $\text{ran}\,(\mathcal{J}_Y^{-1}\mathcal{A})$, $\text{ran}\,(\mathcal{J}_Y^{-1}\mathcal{A}\mathcal{J}_X^{-1}\mathcal{A}^\star)$ is closed, than so the other.*

Proof. We refer the reader to Theorem 4 on page 322 in [5], Theorem 8.4 on page 232 and Theorem 11.2 on page 244 in [14]. □

Lemma 1.39. *Let X and Y be two Hilbert spaces with duals X' and Y' respectively. Then we have for every $\mathcal{A} \in L(X, Y')$:*

$$\mathcal{A} \text{ is surjective} \quad \Longrightarrow \quad \mathcal{A}^\star \text{ is injective.}$$

Proof. Since $\mathcal{A}$ is surjective, we get $\text{ran}\,(\mathcal{A}) = Y'$ and $\text{ran}\,(\mathcal{J}_Y^{-1}\mathcal{A}) = Y$. Due to Lemma 1.38 we obtain $\ker\,(\mathcal{J}_X^{-1}\mathcal{A}^\star)^\perp = Y$. This leads to $\ker\,(\mathcal{J}_X^{-1}\mathcal{A}^\star) = \{0\}$, and therefore we have $\ker\,(\mathcal{A}^\star) = \{0\}$. Hence, $\mathcal{A}^\star$ is injective. □

Remark 1.40. *If $\mathcal{A}^\star$ is injective then $\overline{\text{ran}\,(\mathcal{J}_Y^{-1}\mathcal{A})} = Y$ by Lemma 1.38. From this we derive $\overline{\text{ran}\,(\mathcal{A})} = Y'$. Therefore, $\text{ran}\,(\mathcal{A})$ is only dense in Y'. If $\text{ran}\,(\mathcal{A})$ is closed, $\mathcal{A}$ is surjective. In Theorem 1.42 we will give equivalent assertions for surjectivity of $\mathcal{A}$.*

Example 1.41. *We shall give an example for an injective operator that is not surjective. Let us choose* $\Omega \subset \mathbb{R}^n$ *and* $\mathcal{A}^\star = (-\Delta)^{-1} : L^2(\Omega) \to L^2(\Omega)$ [i], $u = \mathcal{A}^\star(f)$ *is a weak solution of*

$$\begin{cases} -\Delta u = f & \text{in } \Omega,^{\text{ii}} \\ u = 0 & \text{on } \partial\Omega, \end{cases}$$

i.e., $u \in H_0^1(\Omega)$ *solves*

$$\sum_{i=1}^n \int_\Omega u_{x_i}\, \varphi_{x_i}\, dx = \int_\Omega f\varphi\, dx \qquad \text{for all } \varphi \in H_0^1(\Omega)\,. \tag{1.6}$$

It can be proved that (1.6) *has a unique solution* $u \in H_0^1(\Omega)$*. It is clear, that* $\mathcal{A}^\star$ *is linear. Further, if* $u = \mathcal{A}^\star(f)$ *and* $v = \mathcal{A}^\star(g)$ *then* $w = u - v \in H_0^1(\Omega)$ *is a solution of*

$$\sum_{i=1}^n \int_\Omega w_{x_i}\, \varphi_{x_i}\, dx = \int_\Omega (f-g)\varphi\, dx \qquad \text{for all } \varphi \in H_0^1(\Omega).$$

$f \not\equiv g$ *leads to* $w \not\equiv 0$*. Thus,* $\mathcal{A}^\star$ *is injective. Further the linear operator* $-\Delta : L^2(\Omega) \to L^2(\Omega)$ *is selfadjoint. This yields*

$$\begin{aligned} \langle \mathcal{A}^\star(f), g\rangle_{L^2(\Omega)} &\overset{(4.1)}{=} -\int_\Omega \mathcal{A}^\star(f)\, g\, dx = -\int_\Omega \mathcal{A}^\star(f)\, \underbrace{\Delta\left(\mathcal{A}^\star(g)\right)}_{=g}\, dx \\ &\overset{\text{Int. by parts}}{=} -\int_\Omega \underbrace{\Delta\left(\mathcal{A}^\star(f)\right)}_{=f} \mathcal{A}^\star(g)\, dx = \int_\Omega \mathcal{A}^\star(g)\, f\, dx \\ &= \langle \mathcal{A}^\star(g), f\rangle_{L^2(\Omega)} \end{aligned}$$

for all $f, g \in L^2(\Omega)$*. Therefore,* $\mathcal{A}^\star \equiv \mathcal{A}$*. Since* $\operatorname{ran}(\mathcal{A}) \subset H_0^1(\Omega) \overset{\subset}{\neq} L^2(\Omega)$ *we conclude that* $\mathcal{A} = (-\Delta)^{-1}$ *is not surjective.*

Let V_1 and V_2 be two Hilbert spaces and $f : M \subset V_1 \to V_2$. f is said to be *closed* if its graph, $\{(x, f(x) : x \in M\}$, is a closed set in $V_1 \times V_2$. By $D(f)$ we denote the domain of f.

Theorem 1.42. *Let* $\mathcal{A}$ *be a closed densely defined linear operator from a Hilbert space* X *into the dual* Y' *of a Hilbert space* Y*. Then the following assertions are equivalent:*

1) $\mathcal{A}$ *is surjective, i.e.,* $\operatorname{ran}(\mathcal{A}) = Y'$.
2) *There exists a constant* $K \geq 0$ *such that*
$$\|v\|_Y \leq K\, \|\mathcal{A}^\star(v)\|_{X'} \quad \text{for all } v \in D(\mathcal{A}^\star) \subset Y.$$
3) $\ker(\mathcal{A}^\star) = \{0\}$ *and* $\operatorname{ran}(\mathcal{A}^\star)$ *is closed.*

Proof. Let us refer to Theorem II.19 on page 29 in [3]. □

Lemma 1.43. *Let* X *and* Y *be two Hilbert spaces with duals* X' *and* Y'*, respectively, and* $\mathcal{A} \in L(X, Y')$ *be surjective. Then the linear operator* $\mathcal{B} = \mathcal{A} \circ \mathcal{J}_X^{-1} \circ \mathcal{A}^\star : Y \to Y'$ *is bounded and invertible, and* $\mathcal{B}^{-1} \in L(Y', Y)$ *holds.*

---

[i] We introduce the Hilbert spaces $L^2(\Omega)$ and $H_0^1(\Omega)$ in Section 4.

[ii] The Laplacian is given by $\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}$.

Proof. Since $\mathcal{A}$ is surjective we have $\ker(\mathcal{A}^\star) = \{0\}$ by Lemma 1.39. Due to Lemma 1.38 we get $\ker(\mathcal{A}\mathcal{J}_X^{-1}\mathcal{A}^\star) = \{0\}$. Thus, $\mathcal{B}$ is injective. The surjectivity of $\mathcal{A}$ leads to the surjectivity of $\mathcal{J}_Y^{-1}\mathcal{A}$ and the closedness of $\operatorname{ran}(\mathcal{J}_Y^{-1}\mathcal{A})$. Applying Lemma 1.38 again we obtain $\operatorname{ran}(\mathcal{J}_Y^{-1}\mathcal{B})$ is closed, too. This leads to

$$Y = \operatorname{ran}(\mathcal{J}_Y^{-1}\mathcal{A}) = \overline{\operatorname{ran}(\mathcal{J}_Y^{-1}\mathcal{A})} \overset{\text{Lem. 1.38}}{=} \overline{\operatorname{ran}(\mathcal{J}_Y^{-1}\mathcal{B})} = \operatorname{ran}(\mathcal{J}_Y^{-1}\mathcal{B}).$$

Thus, $\mathcal{B}$ is surjective and therefore bijective. Since $\mathcal{A} \in L(X, Y')$ holds $\mathcal{A}^\star$ is bounded (Lemma 1.38). By Lemma 1.8 the operator $\mathcal{B}$ is continuous. Due to Theorem 1.5 we get $\mathcal{B}^{-1}$ exists and is continuous. □

2. Local theory of optimization

In this section we recall optimality conditions for infinite dimensional optimization problems.

Definition 2.1. *Let B_1 and B_2 be Banach spaces and $f : B_1 \to B_2$. If there exists $\mathcal{A} \in L(B_1, B_2)$, such that at some point $x \in B_1$*

$$\lim_{\|y\|_{B_1} \to 0} \frac{\|f(x+y) - f(x) - \mathcal{A}(y)\|_{B_2}}{\|y\|_{B_1}} = 0,$$

then $\mathcal{A}(y)$ is called the Fréchet-differential *of $f(x)$ at x, written $\delta f(x; y)$. The operator $\mathcal{A}$ is called the* Fréchet-derivative *of $f(x)$ at x, and we write $\mathcal{A} = f'(x)$ and $\delta f(x; y) = f'(x)y$.*

If $f : B_1 \to B_2$ has a Fréchet-derivative at x, it is unique (see Proposition 1 on page 172 in [11]) and f is continuous at x (see Proposition 3 on page 173 in [11]). Further we have

$$\langle f'(x), y\rangle_{B_2', B_2} = \lim_{t \to 0} \frac{1}{t} \left(f(x + ty) - f(x)\right) \tag{2.1}$$

for each $y \in B_2$ (see Proposition 2 on page 173 in [11]). If the correspondence $x \mapsto f'(x)$ is continuous at the point x_0, we say that the Fréchet derivative of f is *continuous* at x_0. If the derivative of f is continuous on some open set O, we say that f is *continuously Fréchet-differentiable* on O.

Let X be a Hilbert space and $J : X \to \mathbb{R}$ be a *cost functional*. We consider the constrained minimizing problem

$$\text{minimize } J(x) \quad \text{subject to} \quad e(x) = 0, \tag{2.2}$$

where e is a *constraint function* from X into the dual space Y' of a Hilbert space Y. The *Lagrange functional* associated with (2.2) is denoted

$$\mathcal{L}(x, \lambda) = J(x) + \langle e(x), \lambda\rangle_{Y', Y}$$

and the Lagrange multiplier is some specific $\lambda \in Y$. Partial derivatives with respect to the variable $x \in X$ will be denoted by primes.

Definition 2.2. *We consider the constrained minimizing problem* (2.2)*. If $x_0 \in X$ is such that $e'(x_0)$ maps X onto Y', the point x_0 is said to be a* regular point *of e.*

Theorem 2.3. *Let x_0 be a regular point of the continuously Fréchet-differentiable function e mapping the Hilbert space X into the dual space Y' of a Hilbert space Y. Then there is a neighborhood $U(x_0) \subset X$ of the point x_0, such that $e'(x)$ is surjective for all $x \in U(x_0)$.*

Proof. We define the linear mapping

$$\mathcal{B}(x) = e'(x) \circ \mathcal{J}_X^{-1} \circ e'(x)^\star \tag{2.3}$$

for all $x \in X$. From Lemma 1.43 and Definition 2.2 it follows that $\mathcal{B}(x_0)$ is invertible. Since the mapping $x \mapsto e(x)$ is continuously Fréchet-differentiable the function $x \mapsto \mathcal{B}(x)$ is continuous (Lemma 1.8). Due to Lemma 1.6 and Remark 1.7 there exists a neighborhood $U(x_0)$ of x_0 such that $\mathcal{B}(x)$ is invertible. By Lemma 1.38 this yields $e'(x)$ is surjective for all $x \in U(x_0)$. □

Remark 2.4. *Due to Lemma* 1.39 *the operator $e'(x)^\star$ is injective in the same neighborhood $U(x_0)$ of x_0.*

Our aim is to give necessary conditions for an extremum of J subject to $e(x) = 0$ where J is a real-valued functional on a Hilbert space X and e is a mapping from X into the dual space Y' of a Hilbert space Y.

Theorem 2.5. *If the continuously Fréchet-differentiable functional J has a local extremum under the constraint $e(x) = 0$ at a regular point x^*, then there exists an element $\lambda^* \in Y$ such that the Lagrangian functional is stationary at x^*, i.e.,*

$$\mathcal{L}'(x^*, \lambda^*) = J'(x^*) + e'(x^*)^\star \lambda^* = 0. \tag{2.4}$$

Proof. Let us refer to Theorem 1 on page 243 in [11]. □

Remark 2.6. *The equations (2.4) and $e(x^*) = 0$ are called the* first-order necessary optimality condition *for a local extremum of J at the point x^* under the constraint $e(x) = 0$.*

Now we mention the *second-order sufficient optimality condition* for a local minimum. We refer for a proof to Theorem 5.6 in [12].

Theorem 2.7. *The twice continuously Fréchet-differentiable functional J has a local minimum at the point x^* under the constraint $e(x) = 0$, if there exists $\lambda^* \in Y$ and $\kappa > 0$ such that*

$$\langle \mathcal{L}''(x^*, \lambda^*)v, v\rangle_{X',X} \geq \kappa \|v\|_X^2 \quad \textit{for all } v \in \ker\left(e'(x^*)\right).$$

3. Analysis of an abstract variational problem

In optimization theory saddle-point problems arise very often. For that purpose we recall basic results for an abstract saddlepoint problem.

Let X and Y be two (real) Hilbert spaces with dual spaces X' and Y' respectively. The following bilinear forms are given:

$$a : X \times X \to \mathbb{R}, \quad b : X \times Y \to \mathbb{R}$$

with norms

$$\|a\| = \sup_{\substack{\varphi,\phi \in X \\ \varphi,\phi \neq 0}} \frac{a(\varphi, \phi)}{\|\varphi\|_X \|\phi\|_X}, \quad \|b\| = \sup_{\substack{\varphi \in X, \psi \in Y \\ \varphi \neq 0, \psi \neq 0}} \frac{b(\varphi, \psi)}{\|\varphi\|_X \|\psi\|_Y}.$$

Let $l \in X'$ and $g \in Y'$ be given. We seek $(x, \lambda) \in X \times Y$ such that

$$\left\{ \begin{array}{rcll} a(x, \varphi) + b(\varphi, \lambda) & = & \langle l, \varphi\rangle_{X',X} & \text{for all } \varphi \in X, \\ b(x, \psi) & = & \langle g, \psi\rangle_{Y',Y} & \text{for all } \psi \in Y. \end{array} \right. \tag{3.1}$$

With the two bilinear forms a and b we define linear operators $\mathcal{A} \in L(X, X')$ and $\mathcal{B} \in L(X, Y')$:

$$\begin{aligned} \langle \mathcal{A}(\varphi), \phi \rangle_{X',X} &= a(\varphi, \phi) \qquad \text{for all } \varphi, \phi \in X\,, \\ \langle \mathcal{B}(\varphi), \psi \rangle_{Y',Y} &= b(\varphi, \psi) \qquad \text{for all } \varphi \in X,\ \psi \in Y\,. \end{aligned}$$

The adjoint operator $\mathcal{B}^\star$ of $\mathcal{B}$ is defined by Definition 1.37. It can be shown that

$$\|\mathcal{A}\|_{L(X,X')} = \|a\|, \quad \|\mathcal{B}\|_{L(X,Y')} = \|b\|.$$

Using these operators (3.1) yields to

$$\left\{ \begin{aligned} \mathcal{A}(x) + \mathcal{B}^\star(\lambda) &= l \quad \text{in } X' \\ \mathcal{B}(x) &= g \quad \text{in } Y'. \end{aligned} \right. \tag{3.2}$$

Let $V = \ker(\mathcal{B})$ and $V(g) = \{\varphi \in X : \mathcal{B}(\varphi) = g\}$. It follows

$$\left\{ \begin{array}{l} V(g) = \{\varphi \in X : b(\varphi, \psi) = \langle g, \psi \rangle_{Y',Y} \text{ for all } \psi \in Y\}, \\ V = V(0). \end{array} \right.$$

Since $\mathcal{B}$ is continuous, V is a closed subspace of X.

Now, we associate with (3.1) the following problem: Find $u \in V(g)$ such that

$$a(x, \varphi) = \langle l, \varphi \rangle_{X',X} \quad \text{for all } \varphi \in V. \tag{3.3}$$

It is clear, if $(x, \lambda) \in X \times Y$ is a solution to (3.1), then $u \in V(g)$ is a solution to (3.3). We want to find conditions, which ensure that the converse is true. Therefore, we define the *polar set* V^0 of V by

$$V^0 = \{f \in X' : \langle f, \varphi \rangle_{X',X} = 0 \text{ for all } \varphi \in V\}.$$

Lemma 3.1. *The following properties are equivalent:*

1) *there exists a constant $\beta > 0$ such that*

$$\inf_{\psi \in Y} \sup_{\varphi \in X} \frac{b(\varphi, \psi)}{\|\varphi\|_X \|\psi\|_Y} \geq \beta; \tag{3.4}$$

2) *the operator $\mathcal{B}^\star$ is an isomorphism from Y onto V^0 and*

$$\|\mathcal{B}^\star(\psi)\|_{X'} \geq \beta \|\psi\|_Y \quad \textit{for all } \psi \in Y;$$

3) *the operator $\mathcal{B}$ is an isomorphism from $V^\perp$ onto Y' and*

$$\|\mathcal{B}(\varphi)\|_{Y'} \geq \beta \|\varphi\|_X \quad \textit{for all } \varphi \in V^\perp.$$

Proof. Let us refer the reader to Lemma 4.1 on page 58 in [8]. □

The condition (3.4) is called *inf-sup condition* or *Babuška-Brezzi condition*. To formulate the next theorem, we introduce the linear continuous (restriction) operator $\mathcal{F} \in L(X', V')$:

$$\langle \mathcal{F}(f), \varphi \rangle_{V',V} = \langle f, \varphi \rangle_{X',X} \quad \text{for all } f \in X',\ \varphi \in V.$$

Obviously, we derive

$$\|\mathcal{F}(h)\|_{V'} \leq \|f\|_{X'}.$$

Theorem 3.2. (3.1)*) and* (3.2)*) are well-posed (i.e., there is a unique solution) if and only if the following conditions hold:*

1) *the operator $\mathcal{F} \circ \mathcal{A}$ is an isomorphism from V onto V',*
2) *the bilinear form $b(\cdot, \cdot)$ satisfies the inf-sup condition* (3.4).

Proof. We refer to Theorem 4.1 on page 59 in [8]. $\square$

The next corollary is an important application of Theorem 3.2.

Corollary 3.3. *Let a be V-elliptic, i.e., there exists a constant $\kappa_0 > 0$ such that*

$$a(\varphi,\varphi) \geq \kappa_0 \|\varphi\|_X^2 \quad \text{for all } \varphi \in V.$$

Then, (3.1) *respectively* (3.2) *is well-posed if and only if the bilinear form $b(\cdot,\cdot)$ satisfies the inf-sup condition* (3.4)*. Let $(x,\lambda) \in X \times Y$ be the unique solution to* (3.1)*. Then we have the estimates*

$$\begin{aligned} \|x\|_X &\leq \frac{1}{\kappa_0}\|l\|_{X'} + \Big(\frac{1}{\beta} + \frac{\|a\|}{\kappa_0\beta}\Big)\|g\|_{Y'}, \\ \|\lambda\|_Y &\leq \Big(\frac{1}{\beta} + \frac{\|a\|}{\kappa_0\beta}\Big)\|l\|_{X'} + \Big(\frac{\|a\|}{\beta^2} + \frac{\|a\|^2}{\kappa_0\beta^2}\Big)\|g\|_{Y'}. \end{aligned}$$

Proof. Let us refer to Corollary 4.1 on page 61 in [8] and Theorem 1.1 on page 42 in [4]. $\square$

Now we turn to the discretization of (3.1). Let X_N and Y_M be two finite dimensional spaces such that

$$X_N \subset X, \quad \dim X_N = N, \qquad Y_M \subset Y, \quad \dim Y_M = M.$$

Let X_N' and Y_M' denote their dual spaces with the dual norms:

$$\|l_N\|_{X_N'} = \sup_{\varphi_N \in X_N} \frac{\langle l_N, \varphi_N\rangle_{X',X}}{\|\varphi_N\|_X}, \qquad \|g_M\|_{Y_M'} = \sup_{\psi_M \in Y_M} \frac{\langle g_M, \psi_M\rangle_{Y',Y}}{\|\psi_M\|_Y}. \tag{3.5}$$

Clearly,

$$\|l\|_{X_N'} \leq \|l\|_{X'}, \qquad \|g\|_{Y_M'} \leq \|g\|_{Y'} \quad \text{for all } (l,g) \in X' \times Y'.$$

Like in the continuous case, we associate with $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ the operators $\mathcal{A}_N \in L(X,X_N')$, $\mathcal{B}_M \in L(X,Y_M')$ and $\mathcal{B}_N^\star \in L(Y,X_N')$ defined by

$$\begin{aligned} \langle \mathcal{A}_N(\phi), \varphi_N\rangle_{X',X} &= a(\phi,\varphi_N) \quad \text{for all } \varphi_N \in X_N, \quad \text{for all } \phi \in X, \\ \langle \mathcal{B}_M(\varphi), \psi_M\rangle_{Y',Y} &= b(\varphi,\psi_M) \quad \text{for all } \psi_M \in Y_M, \quad \text{for all } \varphi \in X, \\ \langle \mathcal{B}_N^\star(\psi), \varphi_N\rangle_{X',X} &= b(\varphi_N,\psi) \quad \text{for all } \varphi_N \in X_N, \quad \text{for all } \psi \in Y. \end{aligned}$$

$\mathcal{B}_N^\star$ is not the dual operator of $\mathcal{B}_M$ but if $\mathcal{B}_M$ is restricted to X_N and $\mathcal{B}_N^\star$ to Y_M, then $\mathcal{B}_M$ and $\mathcal{B}_N^\star$ are indeed dual operators. Moreover, we have:

$$\|\mathcal{B}_M(\varphi)\|_{Y_M'} \leq \|\mathcal{B}(\varphi)\|_{Y'} \quad \text{for all } \varphi \in X$$

with similar inequalities for $\|\mathcal{A}_N\|_{X_N'}$ and $\|\mathcal{B}_N^\star(\psi)\|_{X_N'}$. For each $g \in Y'$, we define the finite–dimensional analogue of $V(g)$:

$$V_{NM}(g) = \{\varphi_N \in X_N : b(\varphi_N,\psi_M) = \langle g,\psi_M\rangle_{Y',Y} \text{ for all } \psi_M \in Y_M\}$$

and we set

$$V_{NM} = V_{NM}(0) = \ker(\mathcal{B}_M) \cap X_N = \{\varphi_N \in X_N : b(\varphi_N,\psi_M) = 0 \text{ for all } \psi_M \in Y_M\}.$$

Right away we remark that generally $V_{NM} \not\subset V$ and $V_{NM}(g) \not\subset V(g)$ because Y_M is a proper subspace of Y. Now we approximate (3.1) by

$$\left\{ \begin{aligned} a(x_N,\varphi_N) + b(\varphi_N,\lambda_M) &= \langle l,\varphi_N\rangle_{X',X} \quad \text{for all } \varphi_N \in X_N, \\ b(x_N,\psi_M) &= \langle g,\psi_M\rangle_{Y',Y} \quad \text{for all } \psi_M \in Y_M \end{aligned} \right. \tag{3.6}$$

and we associate with (3.6) the following problem:

$$a(x_N,\varphi_N)=\langle l,\varphi_N\rangle_{X',X}\qquad \text{for all } \varphi_N\in V_{NM}. \tag{3.7}$$

Here again, the first component x_N of any solution (x_N,λ_M) of (3.6) is also a solution of (3.7). The converse is true due to the next theorem.

Theorem 3.4. *Let* (x,λ) *be a solution to* (3.1).

1) *Assume that the following conditions hold:*
   (a) $V_{NM}(g)\neq\emptyset$*;*
   (b) *there exists a constant* $\kappa_{NM}>0$ *such that:*

$$a(\varphi_N,\varphi_N)\geq \kappa_{NM}\,\|\varphi_N\|_X^2\quad \textit{for all } \varphi_N\in V_{NM}. \tag{3.8}$$

Then (3.7) *has a unique solution* $x_N\in V_{NM}(g)$ *and the "error bound" holds*

$$\begin{aligned}&\|x-x_N\|_X\\ &\quad\leq\Big(1+\frac{\|a\|}{\kappa_{NM}}\Big)\inf_{\varphi_N\in V_{NM}(g)}\|x-\varphi_N\|_X+\frac{\|b\|}{\kappa_{NM}}\inf_{\psi_M\in Y_M}\|\lambda-\psi_M\|_Y.\end{aligned} \tag{3.9}$$

2) *Assume that hypothesis holds and, in addition, that:*
   (c) *there exists a constant* $\beta_{NM}>0$ *such that*

$$\sup_{\varphi_N\in X}\frac{b(\varphi_N,\psi_M)}{\|\varphi_N\|_X}\geq\beta_{NM}\,\|\psi_M\|_Y\quad \textit{for all } \psi_M\in Y_M. \tag{3.10}$$

Then $V_{NM}(g)\neq\emptyset$ *and there exists a unique* λ_M *in* Y_M *such that* (x_N,λ_M) *is the only solution of* (3.6)*. Moreover, We get the estimate*

$$\begin{aligned}&\|x-x_N\|_X+\|\lambda-\lambda_M\|_Y\\ &\quad\leq K_{NM}\inf_{\varphi_N\in X_N}\|x-\varphi_N\|_X+C_{NM}\inf_{\psi_M\in Y_M}\|\lambda-\psi_M\|_Y,\end{aligned} \tag{3.11}$$

where the constants are given by

$$K_{NM}=\Big(1+\frac{1}{\beta_{NM}}+\frac{\|a\|}{\kappa_{NM}}\Big)\Big(1+\frac{\|b\|}{\beta_{NM}}\Big),\quad C_{NM}^2=\beta_{NM}+\|b\|+\frac{\|b\|}{\kappa_{NM}}.$$

Proof. For the proof we refer the reader to Theorem 1.1 on page 114 in [8] and Proposition 2.4 on page 54 in [4]. ☐

Remark 3.5. 1) *It can be shown that*

$$\inf_{\varphi_N\in V_{NM}(g)}\|x-\varphi_N\|_X\leq\Big(1+\frac{1}{\beta_{NM}}\Big)\inf_{\phi_N\in X_N}\|x-\phi_N\|_X$$

holds. The condition

$$\sup_{\varphi_N\in X}\frac{b(\varphi_N,\psi_M)}{\|\varphi_N\|_X}\geq\beta^*\,\|\psi_M\|_Y\quad \text{for all } \psi_M\in Y_M. \tag{3.12}$$

is called the inf-sup condition*.*

2) *If there exists two positive constants* $\kappa^*>0$ *and* $\beta^*>0$ *such that* $\kappa_{NM}\geq\kappa^*$ *and* $\beta_{NM}\geq\beta^*$*, then both* K_{NM} *and* C_{NM} *are independent of* N *and* M*:*

$$K=\Big(1+\frac{1}{\beta^*}+\frac{\|a\|}{\kappa^*}\Big)\Big(1+\frac{\|b\|}{\beta^*}\Big),\quad C=\beta^*+\|b\|+\frac{\|b\|}{\kappa^*}.$$

3) *The bound* (3.9) *can be slightly improved without making use of condition* (3.10). *Indeed, it can be proved (see Remark* 1.1 *on page 116 in* [8]*)*

$$\|x - x_N\|_X \le \Big(1 + \frac{\|a\|}{\kappa^*}\Big) \inf_{\varphi_N \in V_{NM}(g)} \|x - \varphi_N\|_X + \frac{1}{\kappa^*} \inf_{\psi_M \in Y_M} \sup_{\varphi_N \in V_{NM}} \frac{b(\varphi_N, \lambda - \psi_M)}{\|\varphi_N\|_X}.$$

Note that the expression

$$\inf_{\psi_M \in Y_M} \sup_{\varphi_N \in V_{NM}} \frac{b(\varphi_N, \lambda - \psi_M)}{\|\varphi_N\|_X}$$

takes into account the fact that $V_{NM} \not\subset V$. *It vanishes when* $V_{NM} \subset V$.

(1) *Observe that the bilinear form* a *is* V_{NM}*-elliptic as soon as* $a(\varphi_N, \varphi_N) > 0$ *for all* $\varphi_N \neq 0$. *Analogously, the bilinear form* b *satisfies the discrete inf-sup condition (3.8) provided* $\ker(\mathcal{B}_M) \cap Y_M = \{0\}$. *But in the general case both assumptions have to be checked.*

The following lemma established a useful criterion for (3.12)

Lemma 3.6. *The inf-sup condition* (3.12) *holds with a constant* $\beta^* > 0$ *independent of* N, M *if and only if there is a restriction operator* $r^N \in L(X, X_N)$ *satisfying:*

$$b(\varphi - r^N(\varphi), \psi_M) = 0 \quad \text{for all } (\psi_M, \varphi) \in Y_M \times X$$

and

$$\|r^N(\varphi)\|_X \le K \|\varphi\|_X \quad \text{for all } \varphi \in X$$

with a constant $K > 0$ *independent of* N.

Proof. Let us refer to Lemma 1.1 on page 117 in [8]. □

4. Function spaces

Since we are interested in optimal control of partial differential equations, we require basic definitions of function spaces and associated results.

The term *domain* and the symbol Ω shall be reserved for an open set in the n–dimensional, real Eucledian space $\mathbb{R}^n$. A typical point of $\mathbb{R}^n$ is denoted by $x = (x_1, \ldots, x_n)$; its norm $|x|_2 = (\sum_{j=1}^n x_j^2)^{\frac{1}{2}}$.

Definition 4.1. *Let* Ω *be an open set of* $\mathbb{R}^n$ *with boundary* Γ. *If boundary* Γ *is a* $(N-1)$*-dimensional manifold of class* C^r *(*$r \ge 1$ *which must be specified) and* Ω *is locally located on one side of* Γ, *we will say,* Ω is of class C^r. *The boundary* Γ *is* locally Lipschitz *if for any* $x \in \Gamma$, *there is a neighborhood such that* Γ *admits a representation as a hypersurface* $x_n = \theta(x_1, \ldots, x_{n-1})$, *where* θ *is Lipschitz continuous and* $x_1, \ldots, x_{n-1}$ *are rectangular coordinates in* $\mathbb{R}^n$ *in a basis that may be different from the canonical basis.*

Remark 4.2. *If* Ω *is of class* C^1, *then* Ω *is locally Lipschitz.*

Definition 4.3. $\Omega \subset \mathbb{R}^n$ *is said to be* disconnected *if there exists two nonempty subsets* $\Omega_1, \Omega_2 \subset \Omega$ *such that*

$$\Omega_1 \cap \Omega_2 = \emptyset, \quad \Omega = \Omega_1 \cup \Omega_2.$$

If Ω *is not disconnected,* Ω *is called* connected.

If $\mathbf{a} = (a_1, \ldots, a_n)$ is an n-tupel of nonnegative integers a_j, we call $\mathbf{a}$ a *multi-index* and denote by $x^{\mathbf{a}}$ the monomial $x_1^{a_1} \cdots x_n^{a_n}$, which has degree $|\mathbf{a}| = \sum_{j=1}^{n} a_j$. Similarly

$$D^{\mathbf{a}} = \frac{\partial^{a_1}}{\partial x_1^{a_1}} \cdots \frac{\partial^{a_n}}{\partial x_n^{a_n}}$$

denotes a differential operator of order $|\mathbf{a}|$. $D^{(0,\ldots,0)}\varphi = \varphi$ for a function φ on Ω. We shall write $G \subset\subset \Omega$ provided $\overline{G} \subset \Omega$ and $\overline{G}$ is a compact subset of $\mathbb{R}^n$. If φ is a function defined on G, we define the *support* of φ as

$$\operatorname{supp}\varphi = \overline{\{x \in G : \varphi(x) \neq 0\}}.$$

We say that φ has *compact support* in Ω if $\operatorname{supp}\varphi \subset\subset \Omega$. We shall denote by ∂G the boundary of G in $\mathbb{R}^n$. For any nonnegative integer m let $C^m(\Omega)$ be the vector space consisting of all functions φ which, together with all their partial derivatives $D^{\mathbf{a}}\varphi$ of order $|\mathbf{a}| \leq m$, are continuous on Ω. We abbreviate $C^0(\Omega) = C(\Omega)$. Let $C^\infty(\Omega) = \bigcap_{m=1}^{\infty} C^m(\Omega)$. The subspaces $C_0(\Omega)$ and $C_0^\infty(\Omega)$ consist of all those functions in $C(\Omega)$ and $C^\infty(\Omega)$, respectively, which have compact support in Ω.

Since Ω is open, functions in $C^m(\Omega)$ need not be bounded on Ω. If $\varphi \in C(\Omega)$ is bounded and uniformly continuous on Ω, then it possesses a unique, bounded, continuous extension to the closure $\overline{\Omega}$ of Ω. Accordingly, we define the vector space $C^m(\overline{\Omega})$ to consist of all those functions $\varphi \in C^m(\Omega)$ for which $D^{\mathbf{a}}\varphi$ is bounded and uniformly continuous on Ω for $0 \leq |\mathbf{a}| \leq m$. $C^m(\overline{\Omega})$ is a Banach space with norm given by

$$\|\varphi\|_{C^m(\overline{\Omega})} = \max_{0\leq|\mathbf{a}|\leq m} \sup_{x\in\Omega} |D^{\mathbf{a}}\varphi(x)|.$$

If $0 < s \leq 1$, we define $C^{m,s}(\overline{\Omega})$ to be the subspace of $C^m(\overline{\Omega})$ consisting of those functions φ, for which, $0 \leq |\mathbf{a}| \leq m$, $D^{\mathbf{a}}\varphi$ satisfies in Ω a *Hölder condition of exponent* s, that is, there exists a constant $K > 0$ such that

$$|D^{\mathbf{a}}\varphi(x) - D^{\mathbf{a}}\varphi(y)| \leq K\,|x-y|^s\,, \quad x, y \in \Omega.$$

$C^{m,s}(\overline{\Omega})$ is a Banach space with norm given by

$$\|\varphi\|_{C^{m,s}(\overline{\Omega})} = \|\varphi\|_{C^m(\overline{\Omega})} + \max_{0\leq|\mathbf{a}|\leq m} \sup_{x,y\in\Omega} \frac{|D^{\mathbf{a}}\varphi(x) - D^{\mathbf{a}}\varphi(y)|}{|x-y|^s}.$$

Where no confusion of domains may occur we will write $\|\cdot\|_{C^m}$ in place of $\|\cdot\|_{C^m(\overline{\Omega})}$ and $\|\cdot\|_{C^{m,s}}$ instead of $\|\cdot\|_{C^{m,s}(\overline{\Omega})}$. It should be noted that for $0 < r < s \leq 1$,

$$C^{m+1}(\overline{\Omega}) \subsetneq C^{m,s}(\overline{\Omega}) \subsetneq C^{m,r}(\overline{\Omega}) \subsetneq C^m(\overline{\Omega}).$$

We denote by $L^p(\Omega)$ the class of all measurable functions φ, defined on Ω, for which

$$\int_\Omega |\varphi|^p \, dx < \infty.$$

The function $\|\cdot\|_{L^p(\Omega)}$ defined by

$$\|\varphi\|_{L^p(\Omega)} = \Big(\int_\Omega |\varphi|^p \, dx \Big)^{\frac{1}{p}}$$

is a norm on $L^p(\Omega)$ provided $1 \leq p < \infty$. $L^2(\Omega)$ is a separable Hilbert space with the inner product

$$\langle \varphi, \psi \rangle_{L^2(\Omega)} = \int_\Omega \varphi\psi \, dx\,. \tag{4.1}$$

$$\|\varphi\|_{L^\infty(\Omega)} = \operatorname{ess\,sup}_{x\in\Omega} |\varphi(x)|$$

is a norm on $L^\infty(\Omega)$. In situations where no confusion of domains may occur we shall write $\|\cdot\|_{L^p}$ in place of $\|\cdot\|_{L^p(\Omega)}$ and $\langle\cdot\,,\cdot\rangle_{L^2}$ instead of $\langle\cdot\,,\cdot\rangle_{L^2(\Omega)}$. The next theorem is known as *Fischer-Riesz theorem*. For a proof we refer the reader to Theorem 2.10 on page 26 and Corollary 2.11 on page 27 in [1].

Theorem 4.4. *$L^p(\Omega)$ is a Banach space if $1 \le p \le \infty$. Every convergent sequence in $L^p(\Omega)$ has a subsequence converging pointwise a.e. on Ω.*

We shall have occasion to use a generalization of Hölders inequality.

Proposition 4.5. *Let Ω be a bounded domain in $\mathbb{R}^n$. Assume $p_1,\dots,p_m \in [1,\infty]$ and $\sum_{j=1}^m p_i^{-1} = 1$ (with $\infty^{-1} = 0$). If $\varphi_i \in L^{p_i}(\Omega)$ for $i = 1,\dots,m$ then the function $\prod_{j=1}^m \varphi_j$ belongs to $L^1(\Omega)$ and we have the estimate*

$$\int_\Omega |\prod_{j=1}^m \varphi_j|\,dx \le \prod_{j=1}^m \|\varphi_j\|_{L^{p_i}}. \tag{4.2}$$

Proof. The proof follows by an induction argument.

1) The case $m = 2$ follows directly from Hölders inequality (Theorem 2.3 on page 23 and Remark 2.5 on page 24 in [1]).
2) Induction hypothesis: We assume that the statement is proved for $m \ge 2$.
3) Now let $p_1,\dots,p_{m+1} \in [1,\infty]$, $m \ge 2$ and $\sum_{j=1}^{m+1} p_j^{-1} = 1$. For $\varphi_i \in L^{p_i}(\Omega)$ we know

$$\int_\Omega |\prod_{j=1}^{m+1} \varphi_j|\,dx \overset{\text{Hölder's ineq.}}{\le} \|\prod_{j=1}^m \varphi_j\|_{L^q}\,\|\varphi_{m+1}\|_{L^{p_{m+1}}}$$

with $q^{-1} + p_{m+1}^{-1} = 1$. Since $\varphi_i \in L^{p_i}(\Omega)$ we have $\varphi_j^q \in L^{\frac{p_i}{q}}(\Omega)$ and $\sum_{j=1}^m (\frac{p_i}{q})^{-1} = 1$. By applying the induction hypothesis we obtain:

$$\|\prod_{j=1}^m \varphi_j\|_{L^q} = \Big(\int_\Omega |\prod_{j=1}^m \varphi_j^q|\,dx\Big)^{\frac{1}{q}} \le \prod_{j=1}^m \|\varphi_j^q\|_{L^{\frac{p_i}{q}}}^{\frac{1}{q}} = \prod_{j=1}^m \|\varphi_j\|_{L^{p_i}}.$$

Therefore, we get $\prod_{j=1}^{m+1} \varphi_j \in L^1(\Omega)$, and the formula (4.2) is proved.

□

Next we introduce Sobolev spaces of integer. These spaces are defined over an arbitrary domain $\Omega \subset \mathbb{R}^n$ and are vector subspaces of various spaces $L^p(\Omega)$. We define a function $\|\cdot\|_{W^{m,p}(\Omega)}$, where m is a nonnegative integer and $1 \le p \le \infty$, as follows:

$$\|\varphi\|_{W^{m,p}(\Omega)} = \Big(\sum_{0\le|\mathbf{a}|\le m} \|D^{\mathbf{a}}\varphi\|_{L^p(\Omega)}^p\Big)^{\frac{1}{p}} \qquad \text{if } 1 \le p < \infty\,, \tag{4.3}$$

$$\|\varphi\|_{W^{m,\infty}(\Omega)} = \max_{0\le|\mathbf{a}|\le m} \|D^{\mathbf{a}}\varphi\|_{L^\infty(\Omega)} \tag{4.4}$$

for any function φ for which the right side makes sense. In situations where no confusion of domains may occur we shall write $\|\cdot\|_{W^{m,p}}$ in place of $\|\cdot\|_{W^{m,p}(\Omega)}$. (4.3) and (4.4) are norms on any linear space on which the right side takes finite values provided functions are identified in the space if they are equal a.e. in Ω. We define to any given values of m and p the Sobolev spaces:

$$\begin{aligned} W^{m,p}(\Omega) &= \text{the completation of } \{\varphi \in C^m(\Omega) : \|\varphi\|_{W^{m,p}} < \infty\} \text{ with} \\ &\quad\ \text{respect to the norm } \|\cdot\|_{W^{m,p}(\Omega)}\,, \end{aligned}$$

$$W_0^{m,p}(\Omega) = \text{the closure of } C_0^\infty(\Omega) \text{ in the space } W^{m,p}(\Omega). \tag{4.5}$$

Equipped with the appropriate norms (4.3) and (4.4), these are called *Sobolev spaces* over Ω. Clearly, $W^{0,p}(\Omega) = L^p(\Omega)$, and if $1 \le p < \infty$, $W_0^{0,p}(\Omega) = L^p(\Omega)$ by Theorem 2.19 on page 31 in [1].

Lemma 4.6. *$W^{m,p}(\Omega)$ is a Banach space.*

Proof. Let us refer the reader to Theorem 3.2 on page 45 in [1]. □

A function φ defined a.e. on Ω is said to be *locally integrable* on Ω provided $\varphi \in L^1(M)$ for every measurable $M \subset\subset \Omega$. In this case we write $\varphi \in L^1_{\rm loc}(\Omega)$. We now define the concept of weak derivative of a locally integrable function $\varphi \in L^1_{\rm loc}(\Omega)$. If there exists a $\phi \in L^1_{\rm loc}(\Omega)$, such that

$$\int_\Omega \varphi D^{\mathbf{a}}\psi\, dx = (-1)^{|\mathbf{a}|} \int_\Omega \phi\psi\, dx \qquad \text{for all } \psi \in C_0^\infty(\Omega)\,,$$

it is unique up to sets of measure zero and it is called the *weak* or *distributional partial derivative* of φ and is denoted by $D^{\mathbf{a}}\varphi$. If φ is sufficiently smooth to have continuous partial derivative $D^{\mathbf{a}}\varphi$ in the usual (classical) sense, then $D^{\mathbf{a}}\varphi$ is also a distributional derivative of φ. For example a function φ, continuous on $\mathbb{R}$, which has a bounded derivative φ' except at finitely many points, has a derivative in the distributional sense.

We can introduce $W^{m,p}(\Omega)$ for any $m \ge 0$ and $1 \le p < \infty$ in a different way:

$$\begin{aligned} W^{m,p}(\Omega) = \{\varphi \in L^p(\Omega) : D^{\mathbf{a}}\varphi \in L^p(\Omega) \text{ for } 0 \le |\mathbf{a}| \le m, \\ D^{\mathbf{a}}\varphi \text{ is the weak partial derivative}\}, \end{aligned}$$

(Theorem 3.16 on page 52 in [1]).

Lemma 4.7. *$W^{m,p}(\Omega)$ is separable if $1 \le p < \infty$. In particular, $W^{m,2}(\Omega)$ is a separable Hilbert space with inner product*

$$\langle \varphi, \psi \rangle_{W^{m,2}} = \sum_{0 \le |\mathbf{a}| \le m} \langle D^{\mathbf{a}}\varphi, D^{\mathbf{a}}\psi \rangle_{L^2} = \sum_{0 \le |\mathbf{a}| \le m} \int_\Omega D^{\mathbf{a}}\varphi\, D^{\mathbf{a}}\psi\, dx.$$

Proof. We refer to Theorem 3.5 on page 47 in [1]. □

We write $H^m(\Omega)$ in place of $W^{m,2}(\Omega)$ and $H_0^m(\Omega)$ instead of $W_0^{m,2}(\Omega)$.

Lemma 4.8. *Let $\Omega \subset \mathbb{R}^n$ be open and bounded with Lipschitz-continuous boundary Γ. For $m \ge 1$ and real p with $1 \le p < \infty$ there exists a continuous linear extension operator $\mathcal{F} : W^{m,p}(\Omega) \to W^{m,p}(\mathbb{R}^n)$ such that*

$$\mathcal{F}(\varphi)|_\Omega = \varphi \quad \textit{for all } \varphi \in W^{m,p}(\Omega).$$

Proof. Let us refer the reader to Theorem 1.2 on page 5 in [8]. □

If the boundary Γ is Lipschitz continuous, one can show that there exists an operator $\tau_\Gamma : H^1(\Omega) \to L^2(\Gamma)$, linear and continuous, such that

$$\tau_\Gamma \varphi = \text{trace of } \varphi \text{ on } \Gamma \text{ for every } \varphi \in C^1(\overline{\Omega}).$$

It then seems natural to call $\tau_\Gamma \varphi$ *the trace of* φ *on* Γ, and denote it by $\varphi|_\Gamma$ even if φ is a general function in $H^1(\Omega)$. A deeper analysis shows that by taking all the traces of all functions of $H^1(\Omega)$ one does not obtain the whole space $L^2(\Gamma)$ but only a subset of it. Further, such a subspace contains $H^1(\Gamma)$. Hence we have,

$$H^1(\Gamma) \subsetneq \tau_\Gamma(H^1(\Omega)) \subsetneq L^2(\Gamma) \equiv H^0(\Gamma).$$

Therefore we introduce the space

$$H^{1/2}(\Gamma) = \tau_\Gamma(H^1(\Omega))$$

with

$$\|g\|_{H^{1/2}} = \inf_{\substack{\varphi \in H^1(\Omega) \\ \tau_\Gamma \varphi = g}} \|\varphi\|_{H^1}.$$

In a similar way one can see that the traces of functions in $H^2(\Omega)$ belong to a space $H^{3/2}(\Gamma)$. We define

$$H^{3/2}(\Gamma) = \tau_\Gamma(H^2(\Omega))$$

and

$$\|g\|_{H^{3/2}} = \inf_{\substack{\varphi \in H^2(\Omega) \\ \tau_\Gamma \varphi = g}} \|\varphi\|_{H^2}.$$

We shall need a special form of the *Sobolev embedding theorem*. The normed space V_1 is said to be *continuous embedded* in the normed space V_2, and write $V_1 \hookrightarrow V_2$ to designate this embedding, provided

1) V_1 is a vector subspace of V_2, and
2) the identity operator $\mathcal{I}_{V_1,V_2}$ defined on V_1 into V_2 by $\mathcal{I}_{V_1,V_2}(v) = v$ for all $v \in V_1$ is continuous.

We say, V_1 is *compact embedded* in V_2, $V_1 \hookrightarrow\hookrightarrow V_2$, if the embedding operator $\mathcal{I}_{V_1,V_2}$ is compact.

Definition 4.9. *The bounded domain $\Omega \subset \mathbb{R}^n$ has a* locally Lipschitz boundary, *if each point x on the boundary $\partial\Omega$ should have a neighborhood $U(x)$ such that $\partial\Omega \cap U(x)$ is the graph of a Lipschitz-continuous function.*

Lemma 4.10. *Let $j \geq 0$ and Ω be a bounded domain in $\mathbb{R}^n$, such that $\partial\Omega$ is locally Lipschitz continuous. Suppose $2m > n > 2(m-1)$. Then*

$$H^{j+m}(\Omega) \hookrightarrow C^{j,s}(\overline{\Omega})$$

for $s \in (0, m - n/2]$.

Proof. We choose $p = 2$ and apply Theorem 5.4, Part II, on page 98 in [1]. ☐

Remark 4.11. *If the assumptions of Lemma* 4.10 *are satisfied, there exists a constant $K > 0$ such that*

$$\|\varphi\|_{C^{j,s}} \leq K \|\varphi\|_{H^{j+m}} \qquad \textit{for all } \varphi \in H^{j+m}(\Omega).$$

Now we mention a special case of the *Rellich-Kondrachov theorem*. For a proof we refer the reader to Theorem 6.2 on page 144 in [1].

Lemma 4.12. *Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with locally Lipschitz boundary, j, m be integers, $j \geq 0$, $m \geq 1$, and let $1 \leq p < \infty$. Then:*

$$\begin{array}{lll} W^{j+m,p}(\Omega) \hookrightarrow\hookrightarrow W^{j,q}(\Omega) & \text{if} \quad n > mp \quad \text{and} & 1 \leq q < \frac{np}{n-mp}\,, \\ W^{j+m,p}(\Omega) \hookrightarrow\hookrightarrow W^{j,q}(\Omega) & \text{if} \quad n = mp \quad \text{and} & 1 \leq q < \infty\,, \\ W^{j+m,p}(\Omega) \hookrightarrow\hookrightarrow W^{j,q}(\Omega) & \text{if} \quad n < mp \quad \text{and} & 1 \leq q \leq \infty. \end{array}$$

Remark 4.13. *If we choose $n \leq 3$, $j = 0$, $m = 1$, $p = 2$, $q = 4$ we get $H^1(\Omega)$ is compact embedded in $L^4(\Omega)$ from Lemma 4.12. By Lemma 1.28 the embedding operator maps weakly convergent sequences in $H_0^1(\Omega)$ into norm convergent sequences in $L^4(\Omega)$. In particular, if $n = 1$ holds, we have $H^1(\Omega) \hookrightarrow\hookrightarrow L^q(\Omega)$ for $1 \leq q \leq \infty$.*

The following lemma characterize weak convergence in $H^1(0,1)$.

Lemma 4.14. *For every $f \in H^1(0,1)'$ there exists $\varphi_1, \varphi_2 \in L^2(0,1)$ such that*

$$\langle f, y\rangle_{(H^1)',H^1} = \int_0^1 y'\varphi_1 + y\varphi_2\,dx$$

for all $y \in H^1(0,1)$.

Proof. Let us refer the reader to Theorem 3.8 on page 48 in [1]. □

The next lemma gives a useful application of Green's formula. For a proof we refer to Lemma 1.4 on page 10 in [8].

Lemma 4.15. *Let Ω be a bounded open subset of $\mathbb{R}^n$ with Lipschitz–continuous boundary Γ and $\mathbf{n} = (n_1, \ldots, n_n)$ the outward unit normal.*

1) *For $u, v \in H^1(\Omega)$ and $1 \leq i \leq n$ we have*

$$\int_\Omega uv_{x_i}\,dx = -\int_\Omega u_{x_i} v\,dx + \int_\Gamma \tau_\Gamma(uv) n_i\,ds.$$

2) *If in addition $u \in H^2(\Omega)$ we derive*

$$\sum_{i=1}^n \int_\Omega u_{x_i}\,v_{x_i}\,dx = -\sum_{i=1}^n \int_\Omega u_{x_i x_i} v\,dx + \sum_{i=1}^n \int_\Gamma \tau_\Gamma\left(v u_{x_i}\right) n_i\,ds. \tag{4.6}$$

Remark 4.16. *By using*

$$\Delta u = \sum_{i=1}^n u_{x_i x_i}\,, \quad \nabla u = \left(u_{x_1}, \ldots, u_{x_n}\right)\,, \quad \mathbf{n} = (n_1, \ldots, n_n)$$

we conclude from (4.6)

$$\int_\Omega \nabla u \cdot \nabla v\,dx = -\int_\Omega \Delta u\,v\,dx + \int_\Gamma \tau_\Gamma\left(v\nabla u\right)\cdot \mathbf{n}\,ds.$$

In the remaining of this section we recall the Fourier transform. By $\iota = \sqrt{-1}$ we denote the *imaginary unit* in $\mathbb{C}$.

Definition 4.17. *For $u \in L^1(\mathbb{R})$ we define the* Fourier transform *of u, denoted by $\hat{u}$:*

$$\hat{u}(y) = \int_{\mathbb{R}} \exp(-\iota xy)\,u(x)\,dx \quad \text{for all } y \in \mathbb{R}.$$

Remark 4.18. *The mapping $u \mapsto \hat{u}$ defined by Definition* 4.17 *is obviously linear. From the inequality*

$$|\hat{u}(y)| \le \|u\|_{L^1(\mathbb{R})} \qquad \text{for all } y \in \mathbb{R}$$

we deduce:

$$\text{(4.7)} \qquad \begin{cases} \text{if } u \in L^1(\mathbb{R}),\ \hat{u} \text{ is a bounded continuous function on } \mathbb{R} \text{ with} \\ \qquad\qquad \|\hat{u}\|_{L^\infty(\mathbb{R})} \le \|u\|_{L^1(\mathbb{R})}. \end{cases}$$

In addition we have the Riemann-Lebesgue theorem (see also Theorem IX.7 *on page 327 in* [13]*):*

$$\text{(4.8)} \qquad \hat{u}(y) \to 0 \quad \text{in} \quad \mathbb{C} \quad \text{when} \quad |y| \to \infty\,.$$

Let $v \in C_0^\infty(\mathbb{R})$. Then we derive, as a result of an integration by parts:

$$\hat{v}(y) = \frac{1}{\iota y} \int_{\mathbb{R}} \exp(-\iota xy)\, v'(x)\, dx,$$

from which we have

$$\text{(4.9)} \qquad |\hat{v}(y)| \le \frac{1}{|y|}\ \|v'\|_{L^1(\mathbb{R})} \to 0, \quad \text{when} \quad |y| \to \infty.$$

Now, as $C_0^\infty(\mathbb{R})$ is dense in $L^1(\mathbb{R})$, if $u \in L^1(\mathbb{R})$, for all $\varepsilon > 0$, we find $v \in C_0^\infty(\mathbb{R})$, such that $\|v - u\|_{L^1(\mathbb{R})} \le \varepsilon/2$. Thus from

$$\hat{u}(y) = \widehat{(u - v)}(y) + \hat{v}(y),$$

we derive using (4.9)

$$|\hat{v}(y)| \le \|u - v\|_{L^1(\mathbb{R})} + \frac{1}{|y|}\ \|v'\|_{L^1(\mathbb{R})} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

for $|y|$ sufficiently large (see (4.9)*), from which* (4.8) *follows.*

We put

$$\mathcal{S}(\mathbb{R}) = \mathcal{S} = \{u \in C^\infty(\mathbb{R}) :\ \text{for all } \alpha, l \in \mathbb{N},\ x^\alpha u^{(l)}(x) \to 0 \text{ as } |x| \to \infty\}.$$

$\mathcal{S}$ is the space of functions of class C^∞ of *rapid decay at infinity*, which is not a normed space, but of which the topology can be defined by the (denumerable) sequence of semi-norms

$$u \mapsto \sup_{x \in \mathbb{R}} |x^k u^{(l)}(x)| = d_{kl}(u)$$

which yields a complete metrisable space:

$$d(u, v) = \sum_{k,l \in \mathbb{N}} a_{kl} \frac{d_{kl}(u - v)}{1 + d_{kl}(u - v)} \quad \text{for all } u, v \in \mathcal{S},$$

where the coefficients a_{kl} are chosen to be such that $\sum_{k,l \in \mathbb{N}} a_{kl} = 1$, is a distance on $\mathcal{S}$. We should notice that if $u \in \mathcal{S}$, then $xu^{(l)}(x) \in L^p(\mathbb{R})$ for all $p \ge 1$ and for all $k, l \in \mathbb{N}$. Further, $\mathcal{S}$ is dense in $L^p(\mathbb{R})$ for all p with $1 \le p < \infty$ (on the contrary $\mathcal{S}$ is not dense in $L^\infty(\mathbb{R})$). For $u \in \mathcal{S}$ we can thus define its Fourier transform by Definition 4.17, as well as the Fourier transform of $x^k u^{(l)}(x)$ for all $k, l \in \mathbb{N}$. Hence, we also have $\hat{u} \in \mathcal{S}$. Further we have the inversion formula

$$u(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \exp(\iota xy) \hat{u}(y)\, dy \quad \text{for all } x \in \mathbb{R}.$$

Let $u, v \in \mathcal{S}$. Then the following properties are valid:

1) $\widehat{u^{(k)}} = (\iota y)^k \hat{u}$.
2) *Parseval's formula*:
$$\int_{\mathbb{R}} u(x)\overline{v(x)}\,dx = \int_{\mathbb{R}} \hat{u}(y)\overline{\hat{v}(y)}\,dy.$$
3) *Plancharel's formula*:
$$\int_{\mathbb{R}} |u(x)|^2\,dx = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} |\hat{u}(y)|^2\,dy.$$

5. Evolution problems: variational methods

In the optimal control of parabolic systems we shall need to make frequent use of the notion of integral of a Banach space-valued function φ defined on an interval of $\mathbb{R}$. We begin therefore with a brief discussion of the Bochner integral, referring the reader to [16], for instance, for further details and proofs of our assertions.

Let B be a Banach space with norm denoted by $\|\cdot\|_B$ and $\{M_1, \dots, M_m\}$ be a finite collection of mutually disjoint, measurable subsets of $\mathbb{R}$, each having finite measure, and let $\{b_1, \dots, b_m\}$ be a corresponding collection of points of B. The function φ on $\mathbb{R}$ defined by

$$\varphi(t) = \sum_{j=1}^{m} \chi_{M_j}(t)\, b_j,$$ [iii]

is called a *simple function*. For simple functions we define

$$\int_{\mathbb{R}} \varphi(t)\,dt = \sum_{j=1}^{m} \mu(M_j)\, b_j$$

where $\mu(M)$ denotes the (Lebesgue) measure of M. Let M be a measurable set in $\mathbb{R}$ and φ an arbitrary function defined a.e. on M into B. The function φ is called (*strongly*) *measurable* on M if there exists a sequence $\{\varphi_n\}_{n\in\mathbb{N}}$ of simple functions with supports in M such that

$$\lim_{n\to\infty} \|\varphi_n(t) - \varphi(t)\|_B = 0 \qquad \text{a.e. in } M. \tag{5.1}$$

It can be shown that any function φ whose range is separable is measurable provided the scalar–valued function $\langle f, \varphi(\cdot)\rangle_{B',B}$ is measurable on M for each $f \in B'$. We suppose that a sequence of simple functions φ_n satisfying (5.1) can be chosen in such a way that

$$\lim_{n\to\infty}\int_M \|\varphi_n(t) - \varphi(t)\|_B\,dt = 0.$$

Then φ is called (*Bochner*) *integrable* on M and we define

$$\int_M \varphi(t)\,dt = \lim_{n\to\infty}\int_{\mathbb{R}} \varphi_n(t)\,dt. \tag{5.2}$$

The integrals on the right side of (5.2) do converge in (the norm topology of) B to a limit which is independent of the choice of approximating sequence $\{\varphi_n\}_{n\in\mathbb{N}}$.

[iii] χ_M denotes the characteristic function of M: $\chi_M(t) = 1$ if $t \in M$ and $\chi_M(t) = 0$ if $t \notin M$.

A measurable function φ is integrable on M if and only if $\|\varphi(\cdot)\|_B$ is (Lebesgue) integrable on M:

$$\Big\| \int_M \varphi(t)\,dt \Big\|_B \le \int_M \|\varphi\|_B\,dt.$$

Definition 5.1. *Let $-\infty \le a < b \le \infty$. We denote by $L^p(a,b;B)$* [iv] *the linear space of (equivalence classes of) functions φ measurable on (a,b) into B such that*

1) *φ is measurable for dt,*
2)

$$\begin{cases} \|\varphi\|_{L^p(a,b;B)} = \Big(\int_a^b \|\varphi(t)\|_B\,dt \Big)^{\frac{1}{p}} < \infty & \text{if } 1 \le p < \infty\,, \\ \|\varphi\|_{L^\infty(a,b;B)} = \operatorname{ess\,sup}_{t\in(a,b)} \|\varphi(t)\|_B < \infty & \text{if } p = \infty. \end{cases} \tag{5.3}$$

If $\varphi \in L^p(c,d;B)$ for every c,d with $a<c<d<b$, then we write $\varphi \in L^p_{\mathrm{loc}}(a,b;B)$, and, if $p=1$, call φ locally integrable.

Proposition 5.2. *For $1 \le p \le \infty$, $L^p(a,b;B)$ is a Banach space.*

Proof. We refer to Proposition 1 on page 469 in [7]. ☐

Proposition 5.3. *If B is a Banach space, a and b are finite, $f \in B'$ and $\varphi \in L^p(a,b;B)$ for $p \ge 1$ we have*

$$\Big\langle f, \int_a^b \varphi(t)\,dt \Big\rangle_{B',B} = \int_a^b \langle f, \varphi(t)\rangle_{B',B}\,dt.$$

Proof. Let us refer the reader to Corollary 2 on page 470 in [7]. ☐

Definition 5.4. *By $\mathcal{D}(a,b)$ we denote the linear space $C_0^\infty(a,b)$. We say, the sequence $\{\varphi_n\}_{n\in\mathbb{N}}$ tends to zero in $\mathcal{D}(a,b)$ if there is a closed subspace $M \subset (a,b)$ such that $\varphi_n^{(i)}(t) = 0$ for all $i \in \mathbb{N}$, for all $n \in \mathbb{N}$ and for all $t \in (a,b)\setminus M$ and it follows $\|\varphi_n^{(i)}\|_{L^\infty} \to 0$ as $n \to \infty$ for all $i \in \mathbb{N}$. We call every continuous linear mapping of $\mathcal{D}(a,b)$ into a Banach space B a* vectorial distribution *over (a,b) with values into B, and we write $\mathcal{D}'(a,b;B) = L(\mathcal{D}(a,b),B)$.*

Remark 5.5. *Let (B_1,B_2) be a pair of Banach spaces with $B_1 \hookrightarrow B_2$. Then we derive*

$$L^p(a,b;B_1) \hookrightarrow L^p(a,b;B_2) \quad \text{for } 1 \le p \le \infty.$$

Proposition 5.6. *Let B be a Banach space and $u \in L^1_{\mathrm{loc}}(a,b;B)$. Then the mapping*

$$\varphi \mapsto \int_a^b \varphi(t)u(t)\,dt$$

is a distribution over (a,b) with values in B.

Proof. We refer the reader to Proposition 4 on page 470 in [7]. ☐

Remark 5.7. *We identify the function u with the distribution with which it is associated.*

[iv] When there is no risk of confusion, we shall write the simplified notation $L^p(B)$.

Proposition 5.8. *Let B be a Banach space. The functions $u, v \in L^1_{\mathrm{loc}}(a,b;B)$ define the same distributions if and only if u and v are equal (in a scalar sense) a.e.*

Proof. Let us refer to Proposition 5 on page 471 in [7]. □

Remark 5.9. *The Proposition 5.8 means that for all $f \in B'$ the functions: $t \to \langle f, u(t)\rangle_{B',B}$ and $t \to \langle f, v(t)\rangle_{B',B}$ are equal almost everywhere. If B is separable, this implies $u = v$ a.e.*

Definition 5.10. *Let B be a Banach space, $f \in \mathcal{D}'(a,b;B)$ and m a nonnegative integer. Then the mapping $\varphi \mapsto (-1)^m f(\frac{d^m\varphi}{dt^m})$, $\varphi \in \mathcal{D}(a,b)$, is a distribution — the* distributional derivative *— that we denote by $\frac{d^m f}{dt^m}$. We have:*

$$\frac{d^m f}{dt^m}\,(\varphi) = (-1)^m f\left(\frac{d^m\varphi}{dt^m}\right) \quad \text{for all } \varphi \in \mathcal{D}(a,b).$$

Remark 5.11. *Let B_1 and B_2 be two separable Banach spaces. If $u \in L^1_{\mathrm{loc}}(a,b;B)$ and if B is a space of functions of the variable x, for instance $B = L^p(\Omega)$, then u is identified with a function $u(t,x)$. $u(t)$ denotes the mapping $x \mapsto u(t,x)$ for almost all t. The distributional derivative $\frac{du}{dt}$ is identified with the derivative $\frac{\partial u}{\partial t}$ in $\mathcal{D}'(a,b;B)$. We use the following notation for the derivative of u with respect to t:*

$$\frac{du}{dt} \quad \text{or} \quad u' \quad \text{or} \quad u_t.$$

Definition 5.12. *Let B be a Banach space and $u \in L^2(a,b;B)$. Then for all $\varphi \in \mathcal{D}(a,b)$:*

$$\frac{du}{dt}\,(\varphi) = -\int_a^b u(t)\varphi'(t)\,dt.$$

We say that $u' = \frac{du}{dt} \in L^2(a,b;B)$ if there exists $v \in L^2(a,b;B)$ such that:

$$\begin{cases} \text{for all } \varphi \in \mathcal{D}(a,b)\,, \quad v(\varphi) = -u(\varphi'), \\ \text{i.e.: } \displaystyle\int_a^b v(t)\varphi(t)\,dt = \int_a^b u(t)\varphi'(t)\,dt. \end{cases}$$

The space we shall introduce next is of fundamental importance. We consider two real, separable Hilbert spaces V, H. It is supposed that V is dense in H so that, by identifying H and its dual H', we have

$$V \hookrightarrow H \equiv H' \hookrightarrow V', \tag{5.4}$$

each space being dense in the following.

Definition 5.13. *Let $a, b \in \mathbb{R}\cup\{-\infty,+\infty\}$, $a < b$. Moreover, V and H are Hilbert spaces satisfying* (5.4)*. The space $W(a,b;V)$ is given by*

$$W(a,b;V) = \Big\{\varphi : \varphi \in L^2(a,b;V),\ \frac{d\varphi}{dt} \in L^2(a,b;V')\Big\}.^{\mathrm{v}}$$

Proposition 5.14. *The space $W(a,b;V)$ endowed with the norm*

$$\|\varphi\|_{W(V)} = \Big(\|\varphi\|^2_{L^2(V)} + \|\tfrac{d\varphi}{dt}\|^2_{L^2(V')}\Big)^{\frac{1}{2}} = \Big(\int_a^b \|\varphi(t)\|^2_V + \|\tfrac{d\varphi(t)}{dt}\|^2_{V'}\,dt\Big)^{\frac{1}{2}}$$

is a Hilbert space.

[v] When there is no risk of confusion, we shall write the simplified notation $W(V)$.

Proof. We refer the reader to Proposition 6 on page 473 in [7]. ☐

We are now interested in regularity properties of elements belonging to $W(V)$. For a proof of the following lemma we refer to Theorem 1 on page 472 in [7].

Lemma 5.15. *For $a, b \in \mathbb{R}$, every $\varphi \in W(V)$ is almost everywhere equal to a continuous function of (a, b) in H. Further, we have:*

$$W(V) \hookrightarrow C([a,b]; H),$$

the space $C([a,b]; H)$ being equipped with the norm of uniform convergence.

Remark 5.16. *By Lemma 5.15 it makes sense to speak of the traces $\varphi(a), \varphi(b) \in H$ for $\varphi \in W(V)$ with $[a,b] \subset \mathbb{R}$. Moreover, we can show, that the mapping $\varphi \mapsto \varphi(a)$ from $W(V)$ is surjective (see Remark 5 on page 477 in [7]).*

Lemma 5.17. *Let $[a,b] \subset \mathbb{R}$ and $\varphi, \phi \in W(V)$. Then*

$$\int_a^b \langle \tfrac{d\varphi(t)}{dt}, \phi(t) \rangle_{V',V}\, dt + \int_a^b \langle \tfrac{d\phi(t)}{dt}, \varphi(t) \rangle_{V',V}\, dt = \langle \varphi(b), \phi(b) \rangle_H - \langle \varphi(a), \phi(a) \rangle_H$$

Proof. We refer to Theorem 2 on page 477 in [7]. ☐

A very useful property is

Proposition 5.18. *For $\varphi \in W(V)$ and $\psi \in V$ we obtain:*

$$\Big\langle \frac{d\varphi(\cdot)}{dt}, \psi \Big\rangle_{V',V} = \frac{d}{dt}\, \langle \varphi(\cdot), \psi \rangle_H$$

in the distributional sense.

Proof. Let us refer to Proposition 7 on page 477 in [7]. ☐

Let X and Y be two separable Hilbert spaces with $X \hookrightarrow Y$ and X being dense in Y. We now define the space $W(a, b; X, Y)$ by

$$W(a,b;X,Y) = \Big\{ \varphi : \varphi \in L^2(a,b;X),\ \tfrac{d\varphi}{dt} \in L^2(a,b;Y) \Big\}$$

equipped with the norm

$$\|\varphi\|_{W(X,Y)} = \Big(\|\varphi\|^2_{L^2(X)} + \Big\| \frac{d\varphi}{dt} \Big\|^2_{L^2(Y)} \Big)^{\frac{1}{2}} = \Big(\int_a^b \|\varphi(t)\|^2_X + \Big\| \frac{d\varphi(t)}{dt} \Big\|^2_Y \, dt \Big)^{\frac{1}{2}}$$

It can be shown that $W(a, b; X, Y)$ is a Hilbert space and that

$$\left\{ \begin{array}{l} \text{i) } X \text{ is dense in } [X,Y]_\theta\,, \theta \in [0,1]^{\text{vi}}, \\ \text{ii) } W(a,b;X,Y) \hookrightarrow C([a,b]; [X,Y]_{1/2}) \end{array} \right. \tag{5.5}$$

(see (1.61) on page 480 in [7]).

We are given two real, separable Hilbert spaces V and H. V is supposed to be dense in H and we identify H with its dual H'. Moreover: $V \hookrightarrow H \hookrightarrow V'$. We denote by $W(V)$ the space $W(0, T; V)$ with $0 < T < \infty$. By Lemma 5.15 we derive

---

[vi]With $[X,Y]_1 = Y$ and $[X,Y]_0 = X$, the space $[X,Y]_\theta$ is called the *holomorphic interpolant* of the spaces X and Y (see Chapter VIII, §3 in [6]).

$W(V) \hookrightarrow C([0,T];H)$. For each $t \in [0,T]$ we are given a continuous bilinear form over $V \times V$ and we make the hypothesis:

$$(5.6) \qquad \begin{cases} \text{for every } \varphi, \psi \in V, \text{ the function } t \mapsto a(t;\varphi,\psi) \text{ is measurable and} \\ \text{there exists a constant } K = K(T) > 0 \text{ (independent of } t \in (0,T), \\ \varphi, \psi \in V) \text{ such that} \\ \qquad |a(t;\varphi,\psi)| \le K \, \|\varphi\|_V \, \|\psi\|_V \quad \text{for all } \varphi, \psi \in V. \end{cases}$$

Therefore, for each $t \in [0,T]$ the bilinear form $a(t;\varphi,\psi)$ defines a continuous linear operator $\mathcal{A}(t)$ from V into V' with

$$\sup_{t \in (0,T)} \|\mathcal{A}(t)\|_{L(V,V')} \le K.$$

We make the following assumption (of coercivity over V with respect to H):

$$(5.7) \qquad \begin{cases} \text{there exists } \lambda, \alpha \text{ constants, } \alpha > 0 \text{ such that} \\ a(t;\varphi,\varphi) + \lambda \|\varphi\|_H^2 \ge \alpha \, \|\varphi\|_V \quad \text{for all } t \in [0,T] \text{ and } \varphi \in V. \end{cases}$$

We give some examples of bilinear forms $a(t;\varphi,\psi)$.

Example 5.19. 1) *We take $V = H_0^1(\Omega)$, $H = L^2(\Omega)$.*

$$a(t;\varphi,\psi) = a(\varphi,\psi) = \langle \psi, \varphi \rangle_{H_0^1}.$$

Then (5.6) *and* (5.7) *holds with $\alpha = 1$ and $\lambda = 0$.*

2) *We take $V = H^1(\Omega)$, $H = L^2(\Omega)$.*

$$a(t;\varphi,\psi) = \langle \varphi, \psi \rangle_{H^1}.$$

Then (5.6) *and* (5.7) *holds with $\alpha = 1$ and $\lambda = 0$.*

3) *Now let V be a closed subspace of $H^1(\Omega)$ with*

$$H_0^1(\Omega) \hookrightarrow V \hookrightarrow H^1(\Omega) \text{ and } H = L^2(\Omega).$$

We set $Q = \Omega \times (0,T)$ and

$$a(t;\varphi,\psi) = \sum_{i,j=1}^{n} \int_\Omega a_{ij}(x,t) \frac{\partial \varphi}{\partial x_i} \frac{\partial \psi}{\partial x_j} \, dx + \int_\Omega a_0(x,t) \varphi \psi \, dx$$

where $a_{ij}, a_0 \in L^\infty(Q)$, $1 \le i,j \le n$ and

$$\sum_{i,j=1}^{n} a_{ij}(x,t) \xi_i \xi_j \ge \alpha \sum_{i=1}^{n} \xi_i^2$$

for a constant $\alpha > 0$ and for $\xi_i \in \mathbb{R}$ a.e. in Q. Then, for λ large enough we derive for all $\varphi \in H^1(\Omega)$:

$$a(t;\varphi,\varphi) + \lambda \|\varphi\|_{L^2}^2 \ge \alpha \, \|\varphi\|_V .$$

Let

$$(5.8) \qquad u_0 \in H, \quad f \in L^2(V')$$

be given. We are looking for

$$(5.9) \qquad u \in W(V),$$

such that

$$\left\{ \begin{array}{l} \dfrac{d}{dt}\, \langle u(\cdot), \varphi \rangle_H + a(\cdot; u(\cdot), \varphi) = \langle f, \varphi \rangle_{V',V} \\ \text{in the sense of } \mathcal{D}'(a,b) \text{ for all } \varphi \in V, \end{array} \right. \tag{5.10}$$

and

$$u(0) = u_0. \tag{5.11}$$

Remark 5.20. 1) *From Lemma* 5.15 *the initial condition* (5.11) *is senseful.*

2) *Due to Proposition* 5.18 *we have*

$$\frac{d}{dt}\, \langle u(\cdot), \varphi \rangle_H = \left\langle \frac{du(\cdot)}{dt}, \varphi \right\rangle_{V',V} \quad \textit{for all } \varphi \in V.$$

Remark 5.21. *If we set $u = we^{kt}$, $k \in \mathbb{R}$, w satisfies*

$$\left\langle \frac{dw(\cdot)}{dt}, \varphi \right\rangle_{V',V} + a(\cdot; w(\cdot), \varphi) + k\, \langle w(\cdot), \varphi \rangle_H = \langle e^{-kt} f(\cdot), \varphi \rangle_{V',V}$$

and

$$w(0) = u_0$$

by changing u to ue^{kt} and choosing k, we can assume that (5.7) *holds with $\lambda = 0$ (that has no consequences since T is finite). In the following we, we shall therefore make the hypothesis:*

$$a(\cdot\,; \varphi, \varphi) \geq \alpha\, \|\varphi\|_V^2 \quad \textit{for all } t \in [0,T]^{\text{vii}} \textit{ and } \varphi \in V. \tag{5.12}$$

Theorem 5.22. *We suppose V, H are given and satisfy $V \hookrightarrow H \hookrightarrow V'$ and $a(\cdot\,; u, \varphi)$ satisfies* (5.6), (5.12). *u_0 and f are given and satisfy* (5.8). *Then there exists a unique solution of* (5.9)-(5.11).

Proof. Let us refer the reader to Theorem 1 on page 512 and Theorem 2 on page 513 in [7]. ☐

If u is the solution of (5.10)-(5.11), we derive

$$\tfrac{1}{2}\, \|u(t)\|_H^2 + \int_0^t a(s; u(s), u(s))\, ds = \tfrac{1}{2}\, \|u_0\|_H^2 + \int_0^t \langle f(s), u(s) \rangle_{V',V}\, ds\,,$$

the so called *energy equality*, as the quantity

$$E(t) = \frac{1}{2}\, \|u_0\|_H^2 + \int_0^t \langle f(s), u(s) \rangle_{V',V}\, ds$$

represents the energy of the system.

Theorem 5.23. *Let $a(t; \varphi, \psi)$ satisfy* (5.6) *and* (5.12), (u_0, f), $(u_0^*, f^*) \in H \times L^2(V')$ *and let u and u^* be the corresponding solutions of* (5.9)-(5.11). *Then*

$$\begin{aligned} \|u - u^*\|_{L^1(H)} &\leq \left(\|u_0 - u_0^*\|_H^2 + \frac{1}{\alpha}\, \|f - f^*\|_{L^2(V')}^2 \right)^{\frac{1}{2}} \\ \|u - u^*\|_{L^2(V)} &\leq \frac{1}{\sqrt{\alpha}} \left(\|u_0 - u_0^*\|_H^2 + \frac{1}{\alpha}\, \|f - f^*\|_{L^2(V')}^2 \right)^{\frac{1}{2}} \end{aligned}$$

Proof. We refer to Theorem 3 on page 520 in [7]. ☐

[vii] Or likewise $t \in [0,T]$ a.e.

Remark 5.24. *We assume $f \in L^1(H)$. Then it can be proved that* (5.9)-(5.11) *has a unique solution in the space*

$$W^*(V) = \{\varphi : \varphi \in L^2(V),\, \varphi' \in L^2(V') + L^1(H)\}.$$

The same estimates as in Theorem 5.23 *hold (see Remark* 6 *on page 521 and Theorem* 4 *on page 522 in* [7]*). Thus problem* (5.9)-(5.11) *can be considered with*

$$f = f_1 + f_2\,, \quad f_1 \in L^1(H)\,, \quad f_2 \in L^2(V')$$

and there exists a unique solution to this problem, the assumptions being those of the beginning of this section.

6. Principal Notations

In the following list of symbols we give the symbol and a descriptive name or phrase for an explanation. The number at the indicate the pagenumber on which the symbols are introduced.

$M \setminus M_1$	complement of M_1 in M	1
$\overline{M}$	closure of a set M	1
M°	interior of M	1
∂M	boundary of a set M	1
$M_1 \times M_2$	Cartesian product of the sets M_1 and M_2	1
$\mathrm{ran}\,(f)$	range of the function f	1
$\ker\,(f)$	kernel of the function f	1
$g \circ f$	compositive mapping given by $x \mapsto g(f(x))$	1
$\|\cdot\|_V$	norm on a (real) normed linear space V	2
$v_n \stackrel{n\to\infty}{\longrightarrow} v$	(strong) convergence of the sequence $\{v_n\}_{n\in\mathbb{N}}$ to v	2
$B(v;\rho)$	open ball of radius ρ about the point v	2
$U(v)$	neighborhood of v	2
$L(V_1, V_2)$	set of bounded linear operators $\mathcal{A} : V_1 \to V_2$	3
$\|\mathcal{A}\|_{L(V_1,V_2)}$	norm of $\mathcal{A} \in L(V_1, V_2)$	3
$\mathcal{A}^{-1}$	inverse of a bounded linear operator $\mathcal{A}$	3
$L(V)$	set of bounded linear operators $\mathcal{A} : V \to V$	3
$\mathcal{I}_V$	identity on a normed linear space V	3
$\langle\cdot,\cdot\rangle_X$	(real) inner product on a Hilbert space X	3
$a(\cdot,\cdot)$	continuous bilinear form	4
$X \oplus Y$	direct sum of the Hilbert spaces X and Y	4
$M^\perp$	orthogonal complement of a closed space $M \subset X$	4
$\mathcal{P}_M$	linear projection onto the subset M	5
B'	dual space of a Banach space B	5
$\langle\cdot,\cdot\rangle_{B',B}$	duality pairing of B' with its Banach space B	5
$\mathcal{J}_X$	Riesz isomorphism of which maps a Hilbert space X onto its dual X'	6
$x^n \stackrel{n\to\infty}{\rightharpoonup} x$	weak convergence of a sequence $\{x_n\}_{n\in\mathbb{N}}$ to x	7
$\mathcal{K}$	compact operator	8
$\rho(\mathcal{A})$	resolvent set of a bounded linear operator $\mathcal{A}$	8
$\sigma(\mathcal{A})$	spectrum of a bounded linear operator $\mathcal{A}$	8
$\mathcal{A}^\star$	adjoint of a bounded linear operator $\mathcal{A}$	9
$\delta f(x;y)$	Fréchet differential of f at x in the direction y	11
$f'(x)$	Fréchet derivative of f at the point x	11

J	cost functional	11
e	constraint function	11
$\mathcal{L}$	Lagrange functional	11
x^*	(local) optimal solution of a constrained minimizing problem	12
λ^*	Lagrange multiplier of a constrained minimizing problem	12
$\lvert \cdot \rvert_2$	Eucledian norm in $\mathbb{R}^n$	16
Ω	open set of $\mathbb{R}^n$	16
$\mathbf{a}$	multi index	17
$\operatorname{supp}(f)$	support of f	17
$C^m(\Omega)$	linear space consisting of m-times continuously differentiable functions	17
$C^m(\overline{\Omega})$	Banach space which is subspace of $C^m(\Omega)$	17
$\lVert \cdot \rVert_{C^m}$	norm on $C^m(\overline{\Omega})$	17
$C_0(\Omega)$, $C_0^\infty(\Omega)$	subspaces of $C(\Omega)$ respectively $C^m(\Omega)$	17
$C^{m,s}(\overline{\Omega})$	Hölder spaces	17
$\lVert \cdot \rVert_{C^{m,s}}$	norm on $C^{m,s}(\overline{\Omega})$	17
$L^p(\Omega)$	L^p-spaces	17
$\lVert \cdot \rVert_{L^p}$	norm on $L^p(\Omega)$	17
$L^1_{\mathrm{loc}}(\Omega)$	linear space consisting locally integrable functions	19
$W^{m,p}(\Omega)$	Sobolev spaces	19
$\lVert \cdot \rVert_{W^{m,p}}$	norm on $W^{m,p}(\Omega)$	18
$W_0^{m,p}(\Omega)$	Sobolev spaces	19
$\lVert \cdot \rVert_{W_0^{m,p}}$	norm on $W_0^{m,p}(\Omega)$	18
$H^m(\Omega)$, $H_0^m(\Omega)$	Hilbert spaces consisting of weak differentiable functions	19
$\lVert \cdot \rVert_{H^m}$, $\lVert \cdot \rVert_{H_0^m}$	norm on $H^m(\Omega)$ respectively $H_0^m(\Omega)$	19
$V_1 \hookrightarrow V_2$	continuous embedding of a normed linear space V_1 into a normed linear space V_2	20
$V_1 \hookrightarrow\hookrightarrow V_2$	compact embedding of a normed linear space V_1 into a normed linear space V_2	20
χ_M	characteristic function of a set M	23
$L^p(0,T;B)$	Banach space consisting of functions with vector values in a Banach space B	24
$\lVert \cdot \rVert_{L^p(a,b;B)}$	norm on the Banach space $L^p(0,T;B)$	24
$W(V)$	Hilbert space consisting of vector-valued functions	25
$\lVert \cdot \rVert_{W(V)}$	norm on $W(V)$	25

References

[1] R. A. Adams. *Sobolev Spaces*. Academic Press, New York, 1975.

[2] H. W. Alt. *Lineare Funktionalanalysis. Eine anwendungsorientierte Einführung*. Springer-Verlag, Berlin, 1992.

[3] H. Brezis. *Analyse fonctionnelle. Théorie et applications*. Masson, 1987.

[4] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods*, volume 15 of *Springer Series in Computational Mathematics*. Springer–Verlag, 1991.

[5] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology. Volume 2: Functional and Variational Methods*. Springer-Verlag, Berlin, 1988.

[6] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology. Volume 3: Spectral Theory and Applications*. Springer-Verlag, Berlin, 1990.

[7] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology. Volume 5: Evolution Problems I*. Springer-Verlag, Berlin, 1992.

[8] V. Girault and P.-A. Raviart. *Finite Element Methods for Navier–Stokes Equations. Theory and Algorithms*. Springer-Verlag, 1980.

[9] D. Gilbarg and N. S. Trudinger. *Elliptic Differential Equations of Second Order*. Springer-Verlag, Berlin, 1977.

[10] F. Hirzebruch and W. Scharlau. *Einführung in die Funktionalanalysis*. B.I.-Wissenschafts-verlag, 1971.

[11] D. G. Luenberger. *Optimization by Vector Space Methods*. John Wiley & Sons, New York, 1969.

[12] H. Maurer and J. Zowe. First and second order necessary and sufficient optimality conditions for infinite-dimensional programming problems. *Math. Programming*, 16:98–110, 1979.

[13] M. Reed and B. Simon. *Methods of Modern Mathematical Physics I: Functional Analysis*. Academic Press, New York, 1980.

[14] A. E. Taylor and D. C. Lay. *Introduction to Functional Analysis*. John Wiley & Sons, New York, 1958.

[15] A. Wouk. *A Course of Applied Functional Analysis*. Wiley-Interscience, New York, 1979.

[16] K. Yosida. *Functional Analysis*. Springer-Verlag, 1965.

S. Volkwein, Institut für Mathematik, Karl-Franzens-Universität Graz, Heinrich-strasse 36, A-8010 Graz, Austria

E-mail address: `stefan.volkwein@uni-graz.at`