5 Linear operators

In this chapter, X, Y, Z, etc. usually denote normed $\mathbb{K}$-vector spaces. We consider linear maps T from X to Y, where, following the notation for matrices, we usually write Tx instead of $T(x)$, and similarly ST instead of $S \circ T$ for linear maps $T : X \to Y$ and $S : Y \to Z$. In functional analysis, only the continuous linear maps are of importance (see E9.2), which are those linear maps for which $T(x)$ can be estimated by x:

5.1 Lemma. If $T : X \to Y$ is linear and $x_0 \in X$, then the following are equivalent:

(1) T is continuous.

(2) T is continuous at x_0.

(3) $\sup_{\|x\|_X \le 1} \|Tx\|_Y < \infty$.

(4) There exists a constant C with $\|Tx\|_Y \le C\|x\|_X$ for all $x \in X$.

Property 5.1(4) written with quantifiers reads

$$\exists\, C \ge 0 \,:\, \big(\, \forall\, x \in X \,:\, \|Tx\|_Y \le C\|x\|_X \,\big)$$

Proof (2)⇒(3). There exists an $\varepsilon > 0$ such that $T\Big(\overline{\mathrm{B}_\varepsilon(x_0)}\Big) \subset \mathrm{B}_1(T(x_0))$. Let $x \in \overline{\mathrm{B}_1(0)}$. Then $x_0 + \varepsilon x \in \overline{\mathrm{B}_\varepsilon(x_0)}$, and hence

$$T(x_0) + \varepsilon T(x) = T(x_0 + \varepsilon x) \in \mathrm{B}_1(T(x_0)) \,,$$

which implies that $T(x) \in \mathrm{B}_{\frac{1}{\varepsilon}}(0)$. ☐

Proof (3)⇒(4). Let C be the supremum in (3). Then for $x \neq 0$

$$\|T(x)\|_Y = \|x\|_X \cdot \left\| T\Big(\frac{x}{\|x\|_X}\Big) \right\|_Y \le \|x\|_X \cdot C \,.$$

☐

Proof (4)⇒(1). For $x, x_1 \in X$ we have that

$$\|T(x) - T(x_1)\|_Y = \|T(x - x_1)\|_Y \le C\|x - x_1\|_X \longrightarrow 0 \quad \text{as } x \to x_1 \,,$$

i.e. T is continuous at x_1. This is true for all x_1. ☐

H.W. Alt, *Linear Functional Analysis*, Universitext,
DOI 10.1007/978-1-4471-7280-2_5

5.2 Linear operators. We define

$$\mathscr{L}(X;Y) := \{ T : X \to Y \ ; \ T \text{ is linear and continuous} \} .$$

We call maps in $\mathscr{L}(X;Y)$ ***linear operators***. This is true in general for topological vector spaces X and Y (see 5.23). In the literature, if they are normed spaces, elements of $\mathscr{L}(X;Y)$ are often also called ***bounded operators***. If X and Y are normed spaces, on recalling 5.1(3), we define for every linear operator $T \in \mathscr{L}(X;Y)$ the ***operator norm*** of T by

$$\|T\|_{\mathscr{L}(X;Y)} := \sup_{\|x\|_X \le 1} \|Tx\|_Y < \infty . \tag{5-5}$$

In the following, we often use the abbreviation $\|T\|$ for $\|T\|_{\mathscr{L}(X;Y)}$. It follows from the proof of 5.1 that $\|T\|_{\mathscr{L}(X;Y)}$ is the smallest number satisfying

$$\|Tx\|_Y \le \|T\|_{\mathscr{L}(X;Y)} \|x\|_X \quad \text{for all } x \in X . \tag{5-6}$$

We set $\mathscr{L}(X) := \mathscr{L}(X;X)$ and denote the ***identity*** on X by Id (or by I). Clearly, $\mathrm{Id} \in \mathscr{L}(X)$.

5.3 Theorem. Let X, Y, and Z be normed spaces.

(1) $\mathscr{L}(X;Y)$ equipped with $\|\cdot\|_{\mathscr{L}(X;Y)}$ in (5-5) is a normed space.

(2) $\mathscr{L}(X;Y)$ is a Banach space if Y is a Banach space.

(3) If $T \in \mathscr{L}(X;Y)$ and $S \in \mathscr{L}(Y;Z)$, then $ST \in \mathscr{L}(X;Z)$ and

$$\|ST\|_{\mathscr{L}(X;Z)} \le \|S\|_{\mathscr{L}(Y;Z)} \cdot \|T\|_{\mathscr{L}(X;Y)} .$$

(4) $\mathscr{L}(X)$ is a Banach algebra if X is a Banach space. Here the product in $\mathscr{L}(X)$ is given by the composition of operators.

Proof (1). For $T_1, T_2 \in \mathscr{L}(X;Y)$ and $x \in X$

$$\|(T_1 + T_2)x\|_Y \le \|T_1 x\|_Y + \|T_2 x\|_Y \le \big(\|T_1\| + \|T_2\|\big)\|x\|_X .$$

Hence $T_1 + T_2 \in \mathscr{L}(X;Y)$ with $\|T_1 + T_2\| \le \|T_1\| + \|T_2\|$, i.e. the operator norm satisfies the triangle inequality. ☐

Proof (2). If $(T_k)_{k\in\mathbb{N}}$ is a Cauchy sequence in $\mathscr{L}(X;Y)$, then for $x \in X$, since $\|T_k x - T_l x\|_Y \le \|T_k - T_l\| \cdot \|x\|_X$, the sequence $(T_k x)_{k\in\mathbb{N}}$ is a Cauchy sequence in Y. As Y is complete, we have that

$$Tx := \lim_{k\to\infty} T_k x \quad \text{in } Y$$

exists pointwise, and it follows easily that $T : X \to Y$ is linear. It then follows that

$$\|(T - T_j)x\|_Y = \lim_{k\to\infty} \|(T_k - T_j)x\|_Y \le \liminf_{k\to\infty} \|T_k - T_j\| \cdot \|x\|_X ,$$

and so $T - T_j \in \mathscr{L}(X;Y)$, by 5.1(4), and

$$\|T - T_j\|_{\mathscr{L}(X;Y)} \le \liminf_{k\to\infty} \|T_k - T_j\|_{\mathscr{L}(X;Y)} \longrightarrow \quad \text{as } j \to \infty$$

(cf. the proof of completeness of $C^0(S;Y)$ in 3.2). □

Proof (3). On noting that

$$\|S(Tx)\|_Z \le \|S\| \cdot \|Tx\|_Y \le \|S\| \cdot \|T\| \cdot \|x\|_X ,$$

we have that $ST \in \mathscr{L}(X;Z)$ with $\|ST\| \le \|S\| \cdot \|T\|$. □

Proof (4). Follows from (3) and (2). □

5.4 Remarks.

(1) If X is finite-dimensional, then every linear map $T : X \to Y$ is continuous, i.e. in $\mathscr{L}(X;Y)$. For noncontinuous linear maps, see E9.2.

(2) Every $T \in \mathscr{L}(X;Y)$ is Lipschitz continuous, since

$$\|T(x) - T(y)\|_Y \le \|T\| \cdot \|x - y\|_X .$$

If follows that for $R > 0$ and $M > 0$

$$A := \left\{ T\big|_{\overline{\mathrm{B}_R(0)}} \; ; \; T \in \mathscr{L}(X;Y),\ \|T\|_{\mathscr{L}(X;Y)} \le M \right\}$$

is a bounded and equicontinuous subset of $C^0\big(\overline{\mathrm{B}_R(0)};Y\big)$. However, the Arzelà-Ascoli theorem is not valid in this context. Observe that A as a subset of $C^0\big(\overline{\mathrm{B}_R(0)};Y\big)$ is not (!) precompact, unless X and Y are finite-dimensional. Only then are the domain and the image set of these continuous functions precompact, which played an essential role in the proof of 4.12.

(3) Linear operators occur as Fréchet derivatives of nonlinear maps $F : X \to Y$. We define $T \in \mathscr{L}(X;Y)$ to be the ***Fréchet derivative*** of F at $x \in X$, if

$$\frac{F(y) - F(x) - T(y - x)}{\|y - x\|_X} \longrightarrow 0 \quad \text{in } Y \text{ as } y \to x \text{ in } X \text{ with } y \ne x.$$

This is the linear approximation property of the mapping $y \mapsto F(y)$ near x, given by the mapping $y \mapsto F(x) + T(y - x)$. Using quantifiers this definition reads

$$\begin{aligned} &\forall\, \varepsilon > 0 \; : \; \exists\, \delta > 0 \; : \; \forall\, y \in X \; : \\ &\|y - x\|_X \le \delta \implies \|F(y) - F(x) - T(y - x)\|_Y \le \varepsilon \cdot \|y - x\|_X . \end{aligned}$$

Proof (1). If n is the dimension of X and $\{e_1, \dots, e_n\}$ is a basis of X, then for $x = \sum_{i=1}^n x_i e_i \in X$

$$\|Tx\|_Y \le \sum_{i=1}^n |x_i| \|Te_i\|_Y \le \left(\sum_{i=1}^n \|Te_i\|_Y \right) \cdot \max_{i=1,\dots,n} |x_i| .$$

If we take, for instance,

$$\|x\| := \max_{i=1,\dots,n} |x_i|$$

as the norm in X (recall lemma 4.8), then, by 5.1, the inequality proves the continuity of T with

$$\|T\|_{\mathscr{L}(X;Y)} \le \sum_{i=1}^{n} \|Te_i\|_Y .$$

□

We now give a list of special linear operators and some notation. The detailed study of the properties of each class of linear operators will be the subject of the following chapters.

5.5 Definitions.

(1) The space $X' := \mathscr{L}(X;\mathbb{K})$ is the ***dual space*** to X. The elements of X' are also called ***linear functionals***. This is true for general topological vector spaces. If X a normed space, then the norm from (5-5) for $T \in X'$ is

$$\|T\|_{X'} := \sup_{\|x\|_X \le 1} |Tx| . \tag{5-7}$$

(2) The set of ***compact (linear) operators*** from X to Y is defined by

$$\mathscr{K}(X;Y) := \left\{ T \in \mathscr{L}(X;Y) \; ; \; \overline{T(\mathrm{B}_1(0))} \text{ is compact} \right\} .$$

If Y is complete, then we can replace "$\overline{T(\mathrm{B}_1(0))}$ is compact" in the definition by "$T(\mathrm{B}_1(0))$ is precompact" (see 4.7(5)).

(3) A linear map $P : X \to X$ is called a ***(linear) projection*** if $P^2 = P$. We denote the set of ***continuous (linear) projections*** by

$$\mathscr{P}(X) := \left\{ P \in \mathscr{L}(X) \; ; \; P^2 = P \right\} .$$

(4) For $T \in \mathscr{L}(X;Y)$ we denote by

$$\mathscr{N}(T) \quad \big(\text{or } \ker(T)\big) \quad := \{x \in X \; ; \; Tx = 0\}$$

the ***null space*** (or ***kernel***) of T. The continuity of T immediately yields that $\mathscr{N}(T)$ is a closed subspace. The ***range*** (or ***image***) of T is defined by

$$\mathscr{R}(T) \quad \big(\text{or } \mathrm{im}(T)\big) \quad := \{Tx \in Y \; ; \; x \in X\} .$$

The subspace $\mathscr{R}(T)$ in general is not closed (see the example 5.6(3)). We will often denote the image of a linear map also as $T(X) = \mathscr{R}(T)$.

(5) $T \in \mathscr{L}(X;Y)$ is called a ***(linear continuous) embedding*** of X into Y if T is injective, i.e. if $\mathscr{N}(T) = \{0\}$.

Observe: In general, the term embedding is used only for very special maps T, see for example the embedding theorems in Chapter 10.

(6) Let X and Y be complete spaces. If $T \in \mathscr{L}(X;Y)$ is bijective, then $T^{-1} \in \mathscr{L}(Y;X)$ (see the inverse mapping theorem 7.8, which plays an essential role in functional analysis). Then T is called an ***invertible (linear) operator***, or a ***(linear continuous) isomorphism***.

(7) $T \in \mathscr{L}(X;Y)$ is called an ***isometry*** (see the definition in 2.24) if

$$\|Tx\|_Y = \|x\|_X \quad \text{for all } x \in X.$$

(8) If $T \in \mathscr{L}(X;Y)$, then

$$(T'y')(x) := y'(Tx) \quad \text{for } y' \in Y',\ x \in X$$

defines a linear map $T' : Y' \to X'$, the ***adjoint map*** of T. We also call T' the ***adjoint operator*** of T, because $T' \in \mathscr{L}(Y', X')$.

Proof (8). For $x \in X$ and $y' \in Y'$,

$$|(T'y')(x)| = |y'(Tx)| \le \|y'\|_{Y'}\|Tx\|_Y \le \|y'\|_{Y'} \cdot \|T\| \cdot \|x\|_X\,,$$

so that, by (5-7),

$$\|T'y'\|_{X'} \le \|y'\|_{Y'} \cdot \|T\|\,,$$

hence, by (5-5), $T' \in \mathscr{L}(Y', X')$ with $\|T'\| \le \|T\|$ (see also 12.1, where we will show that $\|T'\| = \|T\|$). □

Dual spaces will be investigated in Chapter 6. In particular, we will characterize the dual spaces of $C^0(S)$ and $L^p(\mu)$, i.e. we will introduce measure and function spaces, respectively, that are isomorphic to these dual spaces. Continuous linear projections will be considered in Chapter 9. In Chapter 10, we will present the most important types of compact operators, and Chapter 11 will be devoted to the spectral theorem for compact operators. Results on adjoint maps can be found in Chapter 12.

We now give some examples of linear operators.

5.6 Examples.

(1) Let $S \subset \mathbb{R}^n$ be compact and let $(S, \mathcal{B}, \mu)$ be a measure space with $\mu(S) < \infty$, and such that $\mathcal{B}$ contains the Borel sets of S. Then $C^0(S) \subset L^1(\mu)$ and

$$T_\mu f := \int_S f \,\mathrm{d}\mu \quad \text{for } f \in C^0(S)$$

defines a functional $T_\mu \in C^0(S)'$ (see 6.22 and theorem 6.23). For example, if $\mu = \delta_x$ is the Dirac measure for $x \in S$, then $T_{\delta_x} f = f(x)$.

(2) Examples of operators in $\mathscr{L}\big(C^0(S)\big)$, $S \subset \mathbb{R}^n$ compact, are the multiplication operators

$$(T_g f)(x) := f(x)g(x) \quad \text{for } f \in C^0(S)\,,$$

for a fixed $g \in C^0(S)$.

(3) An example of an operator $T \in \mathscr{L}\big(C^0(S);C^1(S)\big)$ with $S = [0,1]$ is

$$(Tf)(x) := \int_0^x f(\xi)\,\mathrm{d}\xi \quad \text{for } f \in C^0(S).$$

One may also consider T as an operator in $\mathscr{L}\big(C^0(S)\big)$. Then $\mathscr{R}(T)$ is not closed in $C^0(S)$, since $\mathscr{R}(T) = \{g \in C^1(S)\,;\ g(0) = 0\}$ is a proper subset of the closure $\overline{\mathscr{R}(T)} = \{g \in C^0(S)\,;\ g(0) = 0\}$. Similarly, T can be defined as an operator in $\mathscr{L}\big(L^1(S)\big)$. Then $\mathscr{R}(T) = \{g \in W^{1,1}(]0,1[)\,;\ g(0) = 0\}$ (see E3.6), which is a proper dense subset of $\overline{\mathscr{R}(T)} = L^1(S)$.

(4) Let $1 \le p \le \infty$ and let $\frac{1}{p} + \frac{1}{p'} = 1$. Then for $g \in L^{p'}(\mu)$ the Hölder inequality yields that

$$T_g f := \int_S f\overline{g}\,\mathrm{d}\mu \quad \text{for } f \in L^p(\mu)$$

defines a functional $T_g \in L^p(\mu)'$ (see theorem 6.12).

(5) If p, p' are as in (4) and $g^s \in L^{p'}(\Omega)$ for $|s| \le m$ with $g = (g^s)_{|s|\le m}$, then

$$T_g f := \sum_{|s|\le m} \int_\Omega \partial^s f \cdot \overline{g^s}\,\mathrm{d}\mathrm{L}^n \quad \text{for } f \in W^{m,p}(\Omega)$$

defines a functional $T_g \in W^{m,p}(\Omega)'$.

(6) Let p be as in (4) and let $(\varphi_k)_{k\in\mathbb{N}}$ be a Dirac sequence. Then 4.13(2) yields that

$$T_k f(x) := \int_{\mathbb{R}^n} \varphi_k(x-y)f(y)\,\mathrm{d}y = (\varphi_k * f)(x)$$

defines an operator $T_k \in \mathscr{L}\big(L^p(\mathbb{R}^n)\big)$ with $\|T_k\| \le 1$. It follows from 4.15(2) that, if $p < \infty$,

$$(T_k - \mathrm{Id})f \to 0 \quad \text{in } L^p(\mathbb{R}^n) \text{ as } k \to \infty$$

for every $f \in L^p(\mathbb{R}^n)$. However, T_k does not converge in the operator norm (see E5.6).

We now prove some fundamental properties of linear operators.

5.7 Neumann series. Let X be a Banach space and let $T \in \mathscr{L}(X)$ with

$$\limsup_{m\to\infty} \|T^m\|^{\frac{1}{m}} < 1$$

(in particular, this is satisfied if $\|T\| < 1$). Then $\mathrm{Id} - T$ is bijective and $(\mathrm{Id} - T)^{-1} \in \mathscr{L}(X)$ with

$$(\mathrm{Id} - T)^{-1} = \sum_{n=0}^{\infty} T^n \quad \text{in } \mathscr{L}(X).$$

Proof. For $k \in \mathbb{N}$ let $S_k := \sum_{n=0}^{k} T^n$. Choose $m \in \mathbb{N}$ and $\theta < 1$ with $\|T^n\| \le \theta^n$ for $n \ge m$. Then for $m \le k < l$

$$\|S_l - S_k\| = \Big\| \sum_{n=k+1}^{l} T^n \Big\| \le \sum_{n=k+1}^{l} \|T^n\| \le \sum_{n=k+1}^{\infty} \theta^n \longrightarrow 0 \quad \text{as } k \to \infty.$$

Since $\mathscr{L}(X)$ is complete, there exists the limit

$$S := \lim_{k\to\infty} S_k \quad \text{in } \mathscr{L}(X)\,.$$

It follows that as $k \to \infty$

$$\begin{aligned}(\mathrm{Id} - T)S \;&\longleftarrow\; (\mathrm{Id} - T)S_k \\ &= \sum_{n=0}^{k} (T^n - T^{n+1}) = \mathrm{Id} - T^{k+1} \;\longrightarrow\; \mathrm{Id} \quad \text{in } \mathscr{L}(X)\,,\end{aligned}$$

because for $k \ge m$ we have that $\|T^{k+1}\| \le \theta^{k+1} \to 0$ as $k \to \infty$. Similarly, one can show that $S(\mathrm{Id} - T) = \mathrm{Id}$. Hence S is the inverse of $\mathrm{Id} - T$. ☐

As a consequence, we obtain that in the space of linear operators, perturbations of invertible operators are again invertible.

5.8 Theorem on invertible operators. Let X, Y be Banach spaces. Then the set of invertible operators in $\mathscr{L}(X;Y)$ is an open subset. More precisely: If $X \ne \{0\}$ and $Y \ne \{0\}$, then for $T, S \in \mathscr{L}(X;Y)$ we have that

$$\left.\begin{array}{l} T \text{ invertible,} \\ \|S - T\| < \|T^{-1}\|^{-1} \end{array}\right\} \quad\Longrightarrow\quad S \text{ invertible.}$$

Proof. Let $R := T - S$. Then $S = T(\mathrm{Id} - T^{-1}R) = (\mathrm{Id} - RT^{-1})T$, where $\|T^{-1}R\| \le \|T^{-1}\| \cdot \|R\| < 1$, and similarly $\|RT^{-1}\| < 1$. Applying 5.7 yields the desired result. ☐

5.9 Analytic functions of operators. Let

$$f(z) := \sum_{n=0}^{\infty} a_n z^n$$

be a power series in $\mathbb{K}$ with radius of convergence $\varrho > 0$. Let X be a Banach space over $\mathbb{K}$. If $T \in \mathscr{L}(X)$, then

$$\limsup_{m\to\infty} \|T^m\|^{\frac{1}{m}} < \varrho \quad\Longrightarrow\quad f(T) := \sum_{n=0}^{\infty} a_n T^n \quad \text{exists in } \mathscr{L}(X)\,.$$

Proof. There exists an r with $0 < r < \varrho$ and an $n \in \mathbb{N}$ with $\|T^m\| \le r^m$ for $m \ge n$. For $n \le m \le k$ it then holds that

$$\Big\| \sum_{i=m}^{k} a_i T^i \Big\| \le \sum_{i=m}^{k} |a_i| \big\|T^i\big\| \le \sum_{i=m}^{\infty} |a_i| r^i \longrightarrow 0 \quad \text{as } m \to \infty$$

thanks to the assumption on the power series. □

5.10 Examples. Let X be a Banach space.

(1) ***Exponential function.*** For all $T \in \mathscr{L}(X)$ we define

$$\exp(T) \ \big(\text{ or } \ \mathrm{e}^T\big) \ := \sum_{n=0}^{\infty} \frac{1}{n!} T^n \quad \in \mathscr{L}(X)\,.$$

For $T, S \in \mathscr{L}(X)$

$$S\,T = T\,S \quad \Longrightarrow \quad \mathrm{e}^{T+S} = \mathrm{e}^T\,\mathrm{e}^S\,.$$

(2) ***Evolution equation.*** For $T \in \mathscr{L}(X)$ the function $A(s) := \mathrm{e}^{sT}$ for $s \in \mathbb{R}$ defines an $A \in C^\infty\big(\mathbb{R}\,; \mathscr{L}(X)\big)$ with

$$\frac{\mathrm{d}}{\mathrm{d}s} A(s) = T\,A(s) = A(s)\,T\,.$$

(3) ***Logarithm.*** For $T \in \mathscr{L}(X)$ with $\|\mathrm{Id} - T\| < 1$ we define

$$\log(T) := -\sum_{n=1}^{\infty} \frac{1}{n} (\mathrm{Id} - T)^n \quad \in \mathscr{L}(X)\,.$$

(4) For $T \in \mathscr{L}(X)$ with $\|T\| < 1$ the function $A(s) := \log(\mathrm{Id} - sT)$ for $|s| < 1$ defines an $A \in C^\infty(]-1,1[\,; \mathscr{L}(X))$ with

$$\frac{\mathrm{d}}{\mathrm{d}s} A(s) = -T\,(\mathrm{Id} - sT)^{-1} = -(\mathrm{Id} - sT)^{-1}\,T$$

and $\exp(A(s)) = \mathrm{Id} - sT$.

The following theorem shows that linear operators commute with the integral (and hence it is a linear version of Jensen's inequality in E4.9).

5.11 Theorem. Let $(S, \mathcal{B}, \mu)$ be a measure space and let Y and Z be Banach spaces. If $f \in L^1(\mu\,; Y)$ and $T \in \mathscr{L}(Y; Z)$, then $T \circ f \in L^1(\mu\,; Z)$ and

$$T\left(\int_S f\,\mathrm{d}\mu\right) = \int_S T \circ f\,\mathrm{d}\mu\,.$$

Explanation: Setting $I_Y f := \int_S f\,\mathrm{d}\mu$ defines $I_Y \in \mathscr{L}\big(L^1(\mu; Y); Y\big)$, and similarly I_Z. In addition, let $\widetilde{T}$ be the operator corresponding to T lifted to functions, i.e. $(\widetilde{T} f)(x) := T(f(x))$ defines $\widetilde{T} \in \mathscr{L}(L^1(\mu; Y); L^1(\mu; Z))$. The theorem then says that

$$T\,I_Y = I_Z\,\widetilde{T}\,,$$

i.e. in this sense, the integral commutes with linear operators.

Proof. Approximate f in $L^1(\mu;Y)$ with step functions

$$f_k = \sum_{i=1}^{n_k} \mathcal{X}_{E_{ki}} \alpha_{ki} \quad \text{with } \alpha_{ki} \in Y \text{ and } \mu(E_{ki}) < \infty,$$

with E_{ki}, $i = 1, \dots, n_k$, being pairwise disjoint. Then as $k \to \infty$

$$\begin{aligned} T\Big(\int_S f \,\mathrm{d}\mu\Big) \longleftarrow T\Big(\int_S f_k \,\mathrm{d}\mu\Big) &= T\Big(\sum_i \mu(E_{ki})\alpha_{ki}\Big) \\ &= \sum_i \mu(E_{ki}) T\alpha_{ki} = \int_S T\circ f_k \,\mathrm{d}\mu\,. \end{aligned}$$

Since

$$\int_S \|T\circ f_k - T\circ f_l\|_Z \,\mathrm{d}\mu \le \|T\| \int_S \|f_k - f_l\|_Y \,\mathrm{d}\mu \longrightarrow 0 \quad \text{as } k,l \to \infty,$$

we have that $(T\circ f_k)_{k\in\mathbb{N}}$ is a Cauchy sequence in $L^1(\mu;Z)$. It follows that there exists a $g \in L^1(\mu;Z)$ such that

$$T\circ f_k \longrightarrow g \text{ in } L^1(\mu;Z)$$

as $k \to \infty$, and hence also

$$\int_S T\circ f_k \,\mathrm{d}\mu \longrightarrow \int_S g \,\mathrm{d}\mu\,.$$

For a subsequence $k \to \infty$ it holds that $T\circ f_k \to g$ almost everywhere in S, and for a further subsequence $k \to \infty$ we have that $f_k \to f$ and hence also $T\circ f_k \to T\circ f$ almost everywhere in S. Consequently, $g = T\circ f$ almost everywhere. □

The linear operators between function spaces that are most important in applications are differential and integral operators.

5.12 Hilbert-Schmidt integral operators. Let $\Omega_1 \subset \mathbb{R}^{n_1}$, $\Omega_2 \subset \mathbb{R}^{n_2}$ be Lebesgue measurable, $1 < p < \infty$ and $1 < q < \infty$, and let $K : \Omega_1 \times \Omega_2 \to \mathbb{K}$ be Lebesgue measurable with

$$\|K\| := \left(\int_{\Omega_1} \Big(\int_{\Omega_2} |K(x,y)|^{p'} \,\mathrm{d}y\Big)^{\frac{q}{p'}} \,\mathrm{d}x\right)^{\frac{1}{q}} < \infty\,, \tag{5-8}$$

where $\frac{1}{p} + \frac{1}{p'} = 1$. Then

$$(Tf)(x) := \int_{\Omega_2} K(x,y) f(y)\,\mathrm{d}y$$

defines an operator $T \in \mathscr{L}\big(L^p(\Omega_2;\mathbb{K})\,;L^q(\Omega_1;\mathbb{K})\big)$ with $\|T\| \le \|K\|$. We call K the ***integral kernel*** of the operator T.

Remark: In 10.15 we will show that T is a compact operator.

Proof. We first assume that all of the following integrals exist. Then using the Hölder inequality we have that

$$\int_{\Omega_1} |Tf(x)|^q\,\mathrm{d}x = \int_{\Omega_1} \left| \int_{\Omega_2} K(x,y)f(y)\,\mathrm{d}y \right|^q \mathrm{d}x$$
$$\le \int_{\Omega_1} \left| \int_{\Omega_2} |K(x,y)|^{p'}\,\mathrm{d}y \right|^{\frac{q}{p'}} \cdot \Big(\int_{\Omega_2} |f(y)|^p\,\mathrm{d}y \Big)^{\frac{q}{p}} \mathrm{d}x = \|K\|^q \cdot \|f\|^q_{L^p(\Omega_2)}\,,$$

which yields the desired result. The existence of the integrals can now be justified retrospectively, similarly to the proof of 4.13(1), and it follows in particular that $Tf \in L^q(\Omega_1)$. Here we note that the assumption (5-8) states that $K(x,\bullet) \in L^{p'}(\Omega_2)$ for almost all $x \in \Omega_1$, and that the function $x \mapsto \|K(x,\bullet)\|_{L^{p'}(\Omega_2)}$ lies in $L^q(\Omega_1)$. □

Now we introduce the set of locally integrable functions.

5.13 Definition. Let $\Omega \subset \mathbb{R}^n$ be open.

(1) We let $D \subset\subset \Omega$ be a shorthand notation for a set $D \subset \mathbb{R}^n$ which is precompact with $\overline{D} \subset \Omega$.

Remark: One also says that D is a ***relatively precompact*** subset of Ω, which means that the closure of D is compact in the relative topology of Ω.

(2) For $1 \le p \le \infty$, let

$$L^p_{\mathrm{loc}}(\Omega) := \big\{ f : \Omega \to \mathbb{K}\ ;\ f\big|_D \in L^p(D) \text{ for all } D \subset\subset \Omega \big\}\,,$$

the vector space of ***locally in Ω p-integrable*** functions.

(3) Equipped with the Fréchet metric

$$\varrho(f) := \sum_{i\in\mathbb{N}} 2^{-i} \frac{\|f\|_{L^p(K_i)}}{1 + \|f\|_{L^p(K_i)}} \qquad \text{for } f \in L^p_{\mathrm{loc}}(\Omega)$$

this is a complete metric space. Here $(K_i)_{i\in\mathbb{N}}$ is a sequence of compact sets, which is an exhaustion of Ω (see (3-2)).

(4) Analogously we define $W^{m,p}_{\mathrm{loc}}(\Omega)$, i.e.

$$W^{m,p}_{\mathrm{loc}}(\Omega) := \{ f : \Omega \to \mathbb{K}\,;\ f\big|_D \in W^{m,p}(D) \text{ for all open sets } D \subset\subset \Omega \}\,.$$

With this we state the following.

5.14 Linear differential operators. Let $\Omega \subset \mathbb{R}^n$ be open and assume $a_s : \Omega \to \mathbb{K}$ for multi-indices s with $|s| \leq m$. Then

$$(Tf)(x) := \sum_{|s|\leq m} a_s(x)\partial^s f(x)$$

defines an operator

(1) $T \in \mathscr{L}\big(C^m(\Omega); C^0(\Omega)\big)$, if $a_s \in C^0(\Omega)$ for $|s| \leq m$.
Remark: $T \in \mathscr{L}\big(C^m(\overline{\Omega}); C^0(\overline{\Omega})\big)$, if all $a_s \in C^0(\overline{\Omega})$ and Ω is bounded.
(2) $T \in \mathscr{L}\big(C^{m,\alpha}(\Omega)\,; C^{0,\alpha}(\Omega)\big)$ with $0 < \alpha \leq 1$ provided $a_s \in C^{0,\alpha}(\Omega)$ for $|s| \leq m$.
Remark: $T \in \mathscr{L}\big(C^{m,\alpha}(\overline{\Omega})\,; C^{0,\alpha}(\overline{\Omega})\big)$, if $a_s \in C^{0,\alpha}(\overline{\Omega})$ and Ω is bounded.
(3) $T \in \mathscr{L}\big(W^{m,p}_{\text{loc}}(\Omega);\, L^p_{\text{loc}}(\Omega)\big)$ with $1 \leq p \leq \infty$, provided $a_s \in L^\infty_{\text{loc}}(\Omega)$ for $|s| \leq m$.
Remark: $T \in \mathscr{L}\big(W^{m,p}(\Omega);\, L^p(\Omega)\big)$, if $a_s \in L^\infty(\Omega)$.

In each case we call T a ***linear differential operator*** of ***order*** m, and we call a_s for $|s| \leq m$ the ***coefficients*** of the differential operator.

Distributions

We now want to consider the functionals in 5.6 in a more general setting. To this end, we restrict the functionals to the common vector space $C_0^\infty(\Omega)$ (here set $S := \overline{\Omega}$ in 5.6). Hence we consider functions and measures only in Ω, i.e. as in 5.14 without boundary conditions. This leads to the following

5.15 Notation. Let $\Omega \subset \mathbb{R}^n$ be open.

(1) Let $(\Omega, \mathcal{B}, \mu)$ be a measure space such that $\mathcal{B}$ contains the Borel sets of Ω and such that μ is finite on compact subsets. Then

$$[\mu](\zeta) \text{ (or } \mathrm{T}_\Omega(\mu)(\zeta)) \;:= \int_\Omega \zeta \,\mathrm{d}\mu \quad \text{for } \zeta \in C_0^\infty(\Omega)$$

defines a linear map $[\mu]$ $\big(\text{or } \mathrm{T}_\Omega(\mu)\big) : C_0^\infty(\Omega) \to \mathbb{K}$.
Remark: With the notation in 5.6(1) we have that $[\mu] = \mathrm{T}_\Omega(\mu) = T_\mu\big|_{C_0^\infty(\Omega)}$.

Note: The integral in this definition is the Riemann integral (see 6.22). Hence for the measures considered here one has $C_0^0(\Omega) \subset L^1(\mu)$.
(2) Let $f \in L^1_{\text{loc}}(\Omega)$. Then

$$[f](\zeta) \text{ (or } \mathrm{T}_\Omega(f)(\zeta)) \;:= \int_\Omega \zeta \cdot f \,\mathrm{dL}^n \quad \text{for } \zeta \in C_0^\infty(\Omega)$$

defines a linear map $[f]$ $\big($or $\mathrm{T}_\Omega(f)\big) : C_0^\infty(\Omega) \to \mathbb{K}$.
Observe: This is a special case of (1), on setting $\mu(E) := \int_E f \, \mathrm{dL}^n$ for Lebesgue measurable sets $E \subset\subset \Omega$ (see the definition 5.13(1)).
Remark: With the notation in 5.6(4) one has $[f] = \mathrm{T}_\Omega(f) = T_{\overline{f}}\big|_{C_0^\infty(\Omega)}$.

5.16 Lemma. Let $\Omega \subset \mathbb{R}^n$ be open and consider the map in 5.15(2)

$$f \mapsto [f] = \mathrm{T}_\Omega(f) \text{ from } L^1_{\mathrm{loc}}(\Omega) \text{ to } \{T : C_0^\infty(\Omega) \to \mathbb{K}\,;\ T \text{ linear}\}.$$

(1) This map is linear and injective.

(2) The function f can be reconstructed from $[f] = \mathrm{T}_\Omega(f)$.

(3) The definition of the weak derivatives $\partial^s f$ of a function $f \in W^{m,1}_{\mathrm{loc}}(\Omega)$ in (3-17) can now be written as

$$(-1)^{|s|}[f](\partial^s\zeta) = [\partial^s f](\zeta) \quad \text{for } \zeta \in C_0^\infty(\Omega),\ |s| \le m. \tag{5-9}$$

Proof (1). This follows from 4.22 (applied to sets $D \subset\subset \Omega$, or note that the fundamental lemma holds in $L^1_{\mathrm{loc}}(\Omega)$). □

Proof (2). To see this, choose $\zeta_\varepsilon = \varphi_\varepsilon * \mathcal{X}_E$ with $E \subset\subset \Omega$ as in the proof of 4.22. Then $[f](\zeta_\varepsilon) \to \int_E f \, \mathrm{dL}^n$ as $\varepsilon \to 0$. Now choose $E = \mathrm{B}_\varepsilon(x)$ with $x \in \Omega$ and obtain for (a subsequence) $\varepsilon \to 0$ that

$$(\varphi_\varepsilon * f)(x) = \mathrm{L}^n\big(\mathrm{B}_\varepsilon(x)\big)^{-1} \int_{\mathrm{B}_\varepsilon(x)} f \, \mathrm{dL}^n \to f(x)$$

for L^n-almost all x. Here we have used 4.15(2). □

This means that knowledge of all the values $[f](\zeta)$ with $\zeta \in C_0^\infty(\Omega)$ provides full information on the function f almost everywhere in Ω. Hence we also call $C_0^\infty(\Omega)$ the space of ***test functions***. We transfer this to linear maps $T : C_0^\infty(\Omega) \to \mathbb{K}$, where the main property is motivated by the structure of the identity (5-9).

5.17 Distributions. Let $\Omega \subset \mathbb{R}^n$ be open and let $T : C_0^\infty(\Omega) \to \mathbb{K}$ be linear.

(1) For all multi-indices s, the ***distributional derivative*** $\partial^s T$ is the linear map $\partial^s T : C_0^\infty(\Omega) \to \mathbb{K}$ defined by

$$(\partial^s T)(\zeta) := (-1)^{|s|} T(\partial^s \zeta) \quad \text{for } \zeta \in C_0^\infty(\Omega)\,. \tag{5-10}$$

(2) We call the linear map T a ***distribution*** on Ω, and use the notation

$$T \in \mathscr{D}'(\Omega)\,,$$

if for all open sets $D \subset\subset \Omega$ there exist a constant C_D and a $k_D \in \mathbb{N} \cup \{0\}$ such that

$$|T(\zeta)| \le C_D \|\zeta\|_{C^{k_D}(\overline{D})} \quad \text{for all } \zeta \in C_0^\infty(\Omega) \text{ with } \operatorname{supp}(\zeta) \subset D. \tag{5-11}$$

If $k = k_D$ can be chosen independently of D, then k (if chosen minimally) is called the ***order*** of T.

(3) If T is a distribution, then so is $\partial^s T$ for all multi-indices s. If T is a distribution of order k, then $\partial^s T$ is a distribution of order $k + |s|$.

Proof (3). We have $|(\partial^s T)(\zeta)| \le C_D \|\partial^s \zeta\|_{C^{k_D}(\overline{D})} \le C_D \|\zeta\|_{C^{k_D+|s|}(\overline{D})}$. ☐

5.18 Examples.

(1) For $f \in W^{m,p}(\Omega)$ and $|s| \le m$

$$\partial^s[f] = [\partial^s f] \quad \text{in } \mathscr{D}'(\Omega). \tag{5-12}$$

Hence the definition of $W^{m,p}(\Omega)$ can also be formulated as follows: A function $f \in L^p(\Omega)$ is in $W^{m,p}(\Omega)$ if all its distributional derivatives up to order m can be identified with functions in $L^p(\Omega)$.

(2) For $f \in L^1_{\text{loc}}(\Omega)$ and $\zeta \in C_0^\infty(D)$ with $D \subset\subset \Omega$

$$[f](\zeta) = \int_\Omega \zeta \cdot f \, \mathrm{dL}^n \quad \text{with} \quad |[f](\zeta)| \le \|f\|_{L^1(D)} \cdot \|\zeta\|_{C^0(\overline{D})} .$$

It follows that $[f] \in \mathscr{D}'(\Omega)$ and is of order 0.

(3) For μ is as in 5.15(1) and for $\zeta \in C_0^\infty(D)$ with $D \subset\subset \Omega$

$$[\mu](\zeta) = \int_\Omega \zeta \, \mathrm{d}\mu \quad \text{with} \quad |[\mu](\zeta)| \le \mu(D) \|\zeta\|_{C^0(\overline{D})} .$$

It follows that $[\mu] \in \mathscr{D}'(\Omega)$ and is of order 0.

(4) As an example, let $\Omega = \mathbb{R}$ and, given $c_-, c_+ \in \mathbb{R}$, let

$$f(x) := \begin{cases} c_+ & \text{for } x > 0, \\ c_- & \text{for } x < 0. \end{cases}$$

By (2), $[f]$ is a distribution of order 0. With the definitions in 5.17(1) and 5.15 it follows that

$$[f]'(\zeta) = -[f](\zeta') = \big(c_+ - c_-\big)\zeta(0) = \big(c_+ - c_-\big)[\delta_0](\zeta) ,$$

where δ_0 is the Dirac measure at the point 0. Hence $[f]'$ is also a distribution of order 0. In addition,

$$[f]''(\zeta) = -[f]'(\zeta') = -\big(c_+ - c_-\big)\zeta'(0) .$$

Hence $[f]''$ is a distribution of order 1, if $c_- \ne c_+$.

(5) Let $(\varphi_k)_{k\in\mathbb{N}}$ be a general Dirac sequence and let δ_0 be the Dirac measure at $0 \in \mathbb{R}^n$. Then it holds as $k \to \infty$ that

$$[\varphi_k](\zeta) \longrightarrow [\delta_0](\zeta) \quad \text{for all } \zeta \in C_0^\infty(\mathbb{R}^n)\,,$$

i.e. $[\varphi_k]$ converges to $[\delta_0]$ as $k \to \infty$ pointwise as a linear map. The name Dirac sequence originates from this property.

(6) As a further example, let $f(x) := \log|x|$ for $x \in \mathbb{R}^n \setminus \{0\}$. Then $f \in L^1_{\text{loc}}(\mathbb{R}^n)$, and so, by (2), $[f]$ is a distribution of order 0 on $\mathbb{R}^n$. For $1 \le i \le n$

$$(\partial_i[f])(\zeta) = \begin{cases} \displaystyle\int_{\mathbb{R}^n} \zeta(x)\frac{x_i}{|x|^2}\,\mathrm{d}x & \text{for } n \ge 2, \\ \displaystyle\lim_{\varepsilon \searrow 0} \int_{\mathbb{R}\setminus[-\varepsilon,\varepsilon]} \zeta(x)\frac{1}{x}\,\mathrm{d}x & \text{for } n = 1. \end{cases}$$

In order to prove this, verify with the help of Gauß's theorem that as $\varepsilon \searrow 0$

$$\begin{aligned}(\partial_i[f])(\zeta) = -[f](\partial_i\zeta) \quad \longleftarrow \quad & -\int_{\mathbb{R}^n\setminus \mathrm{B}_\varepsilon(0)} \partial_i\zeta \cdot f\,\mathrm{dL}^n \\ &= \int_{\mathbb{R}^n\setminus \mathrm{B}_\varepsilon(0)} \zeta\partial_i f\,\mathrm{dL}^n + \int_{\partial \mathrm{B}_\varepsilon(0)} \nu_i\zeta f\,\mathrm{dH}^{n-1}\,,\end{aligned}$$

where $\nu_i(x) = \frac{x_i}{|x|}$ is the i-th component of the outer normal to the set $\mathrm{B}_\varepsilon(0)$ (see A8.5(3) for the general situation). It can be seen that the second integral converges to zero as $\varepsilon \searrow 0$. In the case $n \ge 2$ the function $x \mapsto x_i|x|^{-2}$ is in $L^1_{\text{loc}}(\mathbb{R}^n)$, but not for $n = 1$. Hence for $n \ge 2$ it holds that $\partial_i[f]$ is a distribution of order 0, while for $n = 1$ it can be shown that it is a distribution of order 1.

The essential estimate (5-11) is used in order to approximate distributions with C^∞-functions by means of convolutions.

5.19 Approximation of distributions. Let $\Omega \subset \mathbb{R}^n$ and let $T \in \mathscr{D}'(\Omega)$. For $\varphi \in C_0^\infty(\mathrm{B}_r(0))$ and $x \in \Omega$ with $\mathrm{B}_r(x) \subset \Omega$,

$$(\varphi * T)(x) := T(\varphi(x - \bullet)) \tag{5-13}$$

is well defined, since $\varphi(x - \bullet) \in C_0^\infty(\Omega)$. Moreover, it holds that:

(1) For $T = [f]$ with $f \in L^1_{\text{loc}}(\Omega)$ it follows that

$$(\varphi * [f])(x) = (\varphi * f)(x) \quad \text{if } \mathrm{B}_r(x) \subset \Omega\,.$$

(2) If $D \subset\subset \Omega$ with $\mathrm{B}_r(D) \subset \Omega$, then $\varphi * T \in C^\infty(D)$, with derivatives $\partial^s(\varphi * T) = (\partial^s\varphi) * T$.

(3) Let $D \subset\subset \Omega$ and let $(\varphi_\varepsilon)_{\varepsilon>0}$ be a standard Dirac sequence. For small ε we have that $\varphi_\varepsilon * T \in C^\infty(D)$ and for all $\zeta \in C_0^\infty(D)$

$$[\varphi_\varepsilon * T](\zeta) \longrightarrow T(\zeta) \quad \text{as } \varepsilon \to 0.$$

Proof (1). It holds that

$$(\varphi * [f])(x) = [f](\varphi(x - \bullet)) = \int_\Omega \varphi(x-y)f(y)\,\mathrm{d}y = (\varphi * f)(x)\,,$$

since $\operatorname{supp}(\varphi(x - \bullet)) \subset \Omega$ (formally set $f = 0$ in the exterior of Ω). □

Proof (2). Let k_D be chosen for T and D as in (5-11). On introducing the difference quotients $\partial_i^h \psi(x) := \frac{1}{h}(\psi(x + h\mathbf{e}_i) - \psi(x))$, the linearity of T yields that

$$\partial_i^h(\varphi * T)(x) = T\left(\partial_i^h \varphi(x - \bullet)\right)\,.$$

We have that $\partial_i^h \varphi(x - \bullet) \to \partial_i \varphi(x - \bullet)$ in $C^{k_D}(\overline{D})$ as $h \to 0$, and hence it follows from (5-11) that

$$T(\partial_i^h \varphi(x - \bullet)) \longrightarrow T(\partial_i \varphi(x - \bullet)) = \big((\partial_i \varphi) * T\big)(x)\,.$$

This shows that the partial derivative $\partial_i(\varphi * T)(x) = \big((\partial_i\varphi) * T\big)(x)$ exists. The desired result for higher derivatives now follows by induction on the order of the derivative. □

Proof (3). We have that

$$[\varphi_\varepsilon * T](\zeta) = \int_\Omega \zeta(x) \underbrace{(\varphi_\varepsilon * T)(x)}_{= T\left(\varphi_\varepsilon(x - \bullet)\right)} \,\mathrm{d}x\,.$$

Now it holds that (the proof is given below)

$$\int_\Omega \zeta(x) T\big(\varphi_\varepsilon(x - \bullet)\big)\,\mathrm{d}x = T\left(\int_\Omega \zeta(x)\varphi_\varepsilon(x - \bullet)\,\mathrm{d}x\right)\,. \tag{5-14}$$

The argument of T on the right-hand side is $\zeta_\varepsilon(\bullet)$, if $\zeta_\varepsilon := \varphi_\varepsilon^- * \zeta$ with $\varphi_\varepsilon^-(y) := \varphi_\varepsilon(-y)$. Since $\zeta_\varepsilon \to \zeta$ in $C^{k_D}(\overline{D})$ as $\varepsilon \to 0$, it follows that $T(\zeta_\varepsilon) \to T(\zeta)$, if k_D for T and D is chosen as in (5-11), and so we have shown that

$$[\varphi_\varepsilon * T](\zeta) = T(\zeta_\varepsilon) \longrightarrow T(\zeta) \quad \text{as } \varepsilon \to 0.$$

The identity (5-14) is closely related to theorem 5.11 and the proof is analogous: Approximate ζ uniformly by step functions ζ_j with a common compact support in D. Then (5-14) holds for ζ_j because of the linearity of T. The left-hand side converges as $j \to \infty$, since $T(\varphi_\varepsilon(x - \bullet))$ is continuous, recall (2). The right-hand side converges using the same argument as above, since $\varphi_\varepsilon^- * \zeta_j \to \varphi_\varepsilon^- * \zeta$ in $C^{k_D}(\overline{D})$. □

For functional analysis purposes, the following result is of importance: The vector space $C_0^\infty(\Omega)$ can be equipped with a topology $\mathcal{T}$ in such a way that T is a distribution if and only if T lies in the corresponding dual space, i.e. if $T : C_0^\infty(\Omega) \to \mathbb{K}$ is linear and continuous with respect to the topology $\mathcal{T}$. We denote $C_0^\infty(\Omega)$, equipped with the topology $\mathcal{T}$, by $\mathscr{D}(\Omega)$ (see 5.21). The dual space $\mathscr{D}(\Omega)'$ is then the same as $\mathscr{D}'(\Omega)$ (see 5.23).

5.20 Topology on $C_0^\infty(\Omega)$. Let $\Omega \subset \mathbb{R}^n$ be open. Define

$$p(\zeta) := \sum_{k=0}^{\infty} 2^{-k} \frac{\|\zeta\|_{C^k(\overline{D})}}{1 + \|\zeta\|_{C^k(\overline{D})}} \quad \text{for } \zeta \in C_0^\infty(\Omega) \text{ with } \operatorname{supp}(\zeta) \subset D \subset\subset \Omega\,,$$

where the right-hand side is independent of the choice of D. Choose an open cover $(D_j)_{j\in\mathbb{N}}$ of Ω with sets $D_j \subset\subset D_{j+1} \subset \Omega$ for all $j \in \mathbb{N}$. For every sequence $\varepsilon = (\varepsilon_j)_{j\in\mathbb{N}}$ with $\varepsilon_j > 0$ for $j \in \mathbb{N}$ define

$$U_\varepsilon := \operatorname{conv}\Big(\bigcup_{j\in\mathbb{N}} \{\, \zeta \in C_0^\infty(\Omega)\ ;\ \operatorname{supp}(\zeta) \subset D_j \text{ and } p(\zeta) < \varepsilon_j \,\} \Big)\,.$$

Finally, define

$$\mathcal{T} := \big\{\, U \subset C_0^\infty(\Omega)\ ;\ \text{for } \zeta \in U \text{ there exists an } \varepsilon \text{ with } \zeta + U_\varepsilon \subset U \,\big\}\,.$$

Then:

(1) p is a Fréchet metric with $p(r\zeta) \le rp(\zeta)$ for $r \ge 1$.

(2) For all ε it holds that $U_\varepsilon \in \mathcal{T}$.

(3) $\mathcal{T}$ is a topology. Hence the sets U_ε form a neighbourhood basis (see the definition (4-17)) of 0 with respect to $\mathcal{T}$.

(4) $\mathcal{T}$ is independent of the choice of cover $(D_j)_{j\in\mathbb{N}}$.

We remark that $\mathcal{T}$ is stronger than the topology induced by p. This follows from the fact that the p-ball $\mathrm{B}_\varrho(0) \subset C_0^\infty(\Omega)$ is a neighbourhood in the $\mathcal{T}$-topology, namely, $\mathrm{B}_\varrho(0) = U_\varepsilon$ with $\varepsilon = (\varepsilon_j)_{j\in\mathbb{N}}$ and $\varepsilon_j = \varrho$.

Proof (2). Let $\zeta \in U_\varepsilon$. Consider a finite convex combination

$$\zeta = \textstyle\sum_{k=1}^{k_0} \alpha_k \zeta_k \in U_\varepsilon \quad \text{with } k_0 \in \mathbb{N},\ \alpha_k > 0,\ \sum_{k=1}^{k_0} \alpha_k = 1\,, \tag{5-15}$$

where $\zeta_k \in C_0^\infty(D_{j_k})$ with $p(\zeta_k) < \varepsilon_{j_k}$. Choose $0 < \theta < 1$ such that $p(\zeta_k) < \theta\varepsilon_{j_k}$ for all $k = 1, \dots, k_0$, and set $\delta = (\delta_j)_{j\in\mathbb{N}}$ with $\delta_j := (1-\theta)\varepsilon_j$. We claim that $\zeta + U_\delta \subset U_\varepsilon$. To see this, let

$$\eta = \textstyle\sum_{l=1}^{l_0} \beta_l \eta_l \in U_\delta \quad \text{with } l_0 \in \mathbb{N},\ \beta_l > 0,\ \sum_{l=1}^{l_0} \beta_l = 1\,,$$

where $\eta_l \in C_0^\infty(D_{m_l})$ with $p(\eta_l) < \delta_{m_l}$. Then, on noting (1),

$$p\big(\tfrac{1}{\theta}\zeta_k\big) \le \tfrac{1}{\theta} p(\zeta_k) < \varepsilon_{j_k} \quad \text{and} \quad p\big(\tfrac{1}{1-\theta}\eta_l\big) \le \tfrac{1}{1-\theta} p(\eta_l) < \varepsilon_{m_l}\,,$$

i.e. $\frac{1}{\theta}\zeta_k$ and $\frac{1}{1-\theta}\eta_l$ are elements of U_ε. Hence the convexity of U_ε yields that

$$\zeta + \eta = \theta \sum_{k=1}^{k_0} \alpha_k \cdot \tfrac{1}{\theta}\zeta_k + (1-\theta)\sum_{l=1}^{l_0} \beta_l \cdot \tfrac{1}{1-\theta}\eta_l \in U_\varepsilon \,.$$

This shows that $U_\varepsilon \in \mathcal{T}$. □

Proof (3). We need to show that $U^1 \cap U^2 \in \mathcal{T}$, if $U^1, U^2 \in \mathcal{T}$. But this follows from $U_\varepsilon \subset U_{\varepsilon^1} \cap U_{\varepsilon^2}$, where $\varepsilon_j := \min(\varepsilon_j^1, \varepsilon_j^2)$ for $j \in \mathbb{N}$. □

Proof (4). Let $\left(\widetilde{D}_j\right)_{j\in\mathbb{N}}$ be another cover and let $\widetilde{U}_{\widetilde{\varepsilon}}$ with $\widetilde{\varepsilon} = (\widetilde{\varepsilon}_j)_{j\in\mathbb{N}}$ be a set defined as above, now with respect to this cover. Since $\overline{D}_j$ is compact with $\overline{D}_j \subset \Omega$, for each $j \in \mathbb{N}$ there exists an $m_j \in \mathbb{N}$ with $D_j \subset \widetilde{D}_{m_j}$. Setting $\varepsilon_j := \widetilde{\varepsilon}_{m_j}$ for $j \in \mathbb{N}$ and $\varepsilon = (\varepsilon_j)_{j\in\mathbb{N}}$ then yields that $U_\varepsilon \subset \widetilde{U}_{\widetilde{\varepsilon}}$. □

5.21 The space $\mathscr{D}(\Omega)$. We denote the vector space $C_0^\infty(\Omega)$, equipped with the topology $\mathcal{T}$ from 5.20, by $\mathscr{D}(\Omega)$. Then $\mathscr{D}(\Omega)$ is a ***locally convex topological vector space***, i.e. it holds that:

(1) $\mathscr{D}(\Omega)$ with $\mathcal{T}$ is a Hausdorff space.

(2) $\mathscr{D}(\Omega)$ is a vector space and addition and scalar multiplication are continuous (as maps from $\mathscr{D}(\Omega) \times \mathscr{D}(\Omega)$ to $\mathscr{D}(\Omega)$ and from $\mathbb{K} \times \mathscr{D}(\Omega)$ to $\mathscr{D}(\Omega)$, respectively).

(3) For $\zeta \in U$ with $U \in \mathcal{T}$ there exists a convex set $V \in \mathcal{T}$ with $\zeta \in V \subset U$.

Proof (3). By their definition, the sets U_ε in 5.20 are convex. □

Proof (2). We claim for every U_ε that $U_\delta + U_\delta \subset U_\varepsilon$, where $\delta = (\delta_j)_{j\in\mathbb{N}}$ with $\delta_j := \frac{1}{2}\varepsilon_j$, which implies the continuity of the addition. For the proof let

$$\zeta_l \in C_0^\infty\big(D_{j_l}\big) \quad \text{ with } p\big(\zeta_l\big) < \delta_{j_l} \text{ for } l = 1, 2.$$

We have that $\zeta_1 + \zeta_2 = \frac{1}{2}(2\zeta_1 + 2\zeta_2)$ with $p(2\zeta_l) \le 2p(\zeta_l) \le 2\delta_{j_l} = \varepsilon_{j_l}$, and so $\zeta_1 + \zeta_2 \in U_\varepsilon$, as U_ε is convex. Then the same also holds for arbitrary elements $\zeta_1, \zeta_2 \in U_\delta$.

In order to show the continuity of the scalar multiplication at the point $(\alpha_0, \zeta_0) \in \mathbb{K} \times \mathscr{D}(\Omega)$, let U_ε be given. Let $\zeta_0 \in C_0^\infty(D_{j_0})$ and write

$$\alpha\zeta - \alpha_0\zeta_0 = \frac{1}{2}\big(2(\alpha - \alpha_0)\zeta_0 + 2\alpha(\zeta - \zeta_0)\big)\,.$$

Let $|\alpha - \alpha_0| < \gamma \le \frac{1}{2}$ and let $\zeta - \zeta_0 \in C_0^\infty(D_j)$ with $p(\zeta - \zeta_0) < \delta_j$, where γ, δ_j need to be chosen. Now it holds that $\|2\gamma\zeta_0\|_{C^k(\overline{D}_{j_0})} \to 0$ as $\gamma \to 0$ for all $k \in \mathbb{N}$, and so it follows (as in 2.23(2)) that

$$p\big(2(\alpha - \alpha_0)\zeta_0\big) \le p(2\gamma\zeta_0) \to 0 \quad \text{ as } \gamma \to 0.$$

If we now choose $\gamma \le \frac{1}{2}$ with $p(2\gamma\zeta_0) < \varepsilon_{j_0}$, then $2(\alpha - \alpha_0)\zeta_0 \in U_\varepsilon$. In addition, since $|2\alpha| \le 2(|\alpha_0| + \gamma) \le 2|\alpha_0| + 1$,

$$p\big(2\alpha(\zeta-\zeta_0)\big) \le (1+2|\alpha_0|)p(\zeta-\zeta_0) < \varepsilon_j\,,$$

if we set $\delta_j := (1+2|\alpha_0|)^{-1}\varepsilon_j$. This implies that also $2\alpha(\zeta-\zeta_0) \in U_\varepsilon$, and hence $\alpha\zeta \in \alpha_0\zeta_0 + U_\varepsilon$. Then the same also follows for all $\zeta \in \zeta_0 + U_\delta$, where $\delta := (\delta_j)_{j\in\mathbb{N}}$. □

Proof (1). Let $\zeta^1, \zeta^2 \in \mathscr{D}(\Omega)$ with $\zeta^1 \neq \zeta^2$ and $\zeta := \zeta^1 - \zeta^2$. We claim that

$$\big(\zeta^1 + U_\varepsilon\big) \cap \big(\zeta^2 + U_\varepsilon\big) = \emptyset\,,$$

if $\varepsilon = (\varrho)_{j\in\mathbb{N}}$ and $\varrho > 0$ is sufficiently small. Indeed, if $\eta^1, \eta^2 \in U_\varepsilon$ with $\zeta^1 + \eta^1 = \zeta^2 + \eta^2$, then also $-\eta^1 \in U_\varepsilon$, and so

$$\zeta = \zeta^1 - \zeta^2 = (-\eta^1) + \eta^2 \in U_\varepsilon + U_\varepsilon \subset U_{2\varepsilon}\,,$$

on recalling the proof of (2). Now write ζ as a convex combination as in (5-15), so that

$$\frac{\|\zeta_k\|_{C^0}}{1+\|\zeta_k\|_{C^0}} \le p(\zeta_k) < 2\varrho\,.$$

This implies, if $\varrho < \frac{1}{2}$, that

$$0 \neq \|\zeta\|_{C^0} \le \textstyle\sum_{k=1}^{k_0} \alpha_k \|\zeta_k\|_{C^0} \le \max_{k=1,\dots,k_0} \|\zeta_k\|_{C^0} < \frac{2\varrho}{1-2\varrho}\,,$$

which is not possible, if ϱ depending on ζ was chosen sufficiently small. □

5.22 Lemma. For every sequence $(\zeta_m)_{m\in\mathbb{N}}$ in $\mathscr{D}(\Omega)$ it holds that:

$$\zeta_m \to 0 \quad \text{as } m \to \infty \text{ in } \mathscr{D}(\Omega)$$

if and only if

(1) There exists an open $D \subset\subset \Omega$ such that $\zeta_m \in C_0^\infty(D)$ for all m.

(2) For all $D \subset\subset \Omega$ and all $k \in \mathbb{N}$ it holds that $\|\zeta_m\|_{C^k(\overline{D})} \to 0$ as $m \to \infty$.

Proof ⇐. On noting that $\overline{D}$ is compact and $\overline{D} \subset \Omega$, the cover in 5.20 contains a D_j such that $D \subset D_j$. Then for a given ε it follows from (2) (as in 2.23(2)) that $p(\zeta_m) < \varepsilon_j$ for large m, and so $\zeta_m \in U_\varepsilon$. □

Proof ⇒. If we assume that (1) is not satisfied, then there exist an open cover $(D_j)_{j\in\mathbb{N}}$ of Ω with $D_j \subset\subset \Omega$ and $D_{j-1} \subset D_j$, as well as $x_j \in D_j \setminus \overline{D_{j-1}}$ and a subsequence $m_j \to \infty$, such that $\zeta_{m_j}(x_j) \neq 0$. Then

$$U := \Big\{\, \zeta \in \mathscr{D}(\Omega)\ ;\ \sum_{j\in\mathbb{N}} \frac{2}{|\zeta_{m_j}(x_j)|} \|\zeta\|_{C^0(\overline{D_j}\setminus D_{j-1})} \le 1 \,\Big\}$$

is a convex subset of $\mathscr{D}(\Omega)$. On noting that for all j

$$\{\, \zeta \in C_0^\infty(D_j)\ ;\ p(\zeta) < \varepsilon_j \,\} \subset U\,, \quad \text{where} \quad \varepsilon_j := \Big(1 + \sum_{i\le j} \frac{2}{|\zeta_{m_i}(x_i)|}\Big)^{-1},$$

we have that $U_\varepsilon \subset U$, if $\varepsilon = (\varepsilon_j)_{j\in\mathbb{N}}$ and U_ε is defined with respect to the cover $(D_j)_{j\in\mathbb{N}}$. The definition of the topology and the fact that $\zeta_m \to 0$ in $\mathscr{D}(\Omega)$ as $m \to \infty$ yield that $\zeta_m \in U_\varepsilon$ for large m. But it follows from the construction of U that the ζ_{m_j} do not lie in U, a contradiction. This shows (1).

Now for $k \in \mathbb{N}$ and $\delta > 0$ choose $\varepsilon = (\varepsilon_j)_{j\in\mathbb{N}}$ with $2^k\varepsilon_j = \left(1+\frac{1}{\delta}\right)^{-1} > 0$ for all j, which yields that

$$U_\varepsilon \subset \left\{ \zeta \in C_0^\infty(\Omega)\ ;\ \|\zeta\|_{C^k} \le \delta \right\}.$$

For large m we have that $\zeta_m \in U_\varepsilon$, and so $\|\zeta_m\|_{C^k} \le \delta$. This shows (2). □

5.23 The dual space of $\mathscr{D}(\Omega)$. Consider (see 5.5(1)) the dual space

$$\mathscr{D}(\Omega)' = \{T : \mathscr{D}(\Omega) \to \mathbb{K}\,;\ T \text{ is linear and continuous}\}$$

of $\mathscr{D}(\Omega)$. Then (with the notation in 5.17(2))

$$\mathscr{D}(\Omega)' = \mathscr{D}'(\Omega).$$

Proof $\subset$. Let $T \in \mathscr{D}(\Omega)'$. If $T \notin \mathscr{D}'(\Omega)$, then there exist a $D \subset\subset \Omega$ and $\zeta_m \in C_0^\infty(D)$ with

$$1 = |T\zeta_m| > m\|\zeta_m\|_{C^m(\overline{D})} \quad \text{for } m \in \mathbb{N}.$$

For all $k \in \mathbb{N}$ it then follows that $\|\zeta_m\|_{C^k(\overline{D})} \to 0$ as $m \to \infty$, and so 5.22 yields $\zeta_m \to 0$ as $m \to \infty$ in $\mathscr{D}(\Omega)$. Now the continuity of T implies that $T\zeta_m \to 0$ as $m \to \infty$, which is a contradiction. □

Proof $\supset$. Let $T \in \mathscr{D}'(\Omega)$, let $(D_j)_{j\in\mathbb{N}}$ be the exhaustion from 5.20 and let

$$|T\zeta| \le C_j\|\zeta\|_{C^{k_j}(\overline{D_j})} \quad \text{for } \zeta \in C_0^\infty(D_j).$$

For $\delta > 0$ let $\varepsilon = (\varepsilon_j)_{j\in\mathbb{N}}$ be defined by $\varepsilon_j := 2^{-k_j}\frac{\delta}{C_j+\delta}$. Then

$$\zeta \in C_0^\infty(D_j) \text{ with } p(\zeta) < \varepsilon_j \quad \Longrightarrow \quad |T\zeta| \le C_j\|\zeta\|_{C^{k_j}(\overline{D_j})} \le \delta.$$

As T is linear, it follows that $|T\zeta| \le \delta$ for all $\zeta \in U_\varepsilon$ (with U_ε as in 5.20). This proves the continuity of T. □

E5 Exercises

E5.1 Commutator. Let X be a nontrivial normed vector space and let $P, Q : X \to X$ be linear maps with $PQ - QP = \mathrm{Id}$. Then P and Q cannot both be continuous. (This relation, which appears in quantum mechanics, is called the Heisenberg relation.)

Solution. It follows inductively for $n \in \mathbb{N}$ that

$$PQ^n - Q^n P = nQ^{n-1}\,, \tag{E5-1}$$

on noting that for such n we have that

$$\begin{aligned} PQ^{n+1} - Q^{n+1}P &= \underbrace{(PQ^n - Q^n P)}_{=nQ^{n-1}} Q + Q^n \underbrace{(PQ - QP)}_{=1Q^0=\mathrm{Id}} \\ &= nQ^{n-1}Q + Q^n = (n+1)Q^n\,. \end{aligned}$$

Assuming that $P, Q \in \mathscr{L}(X)$, it follows from (E5-1) that

$$n\big\|Q^{n-1}\big\| \le 2\|P\| \cdot \|Q^n\| \le 2\|P\| \cdot \|Q\| \cdot \big\|Q^{n-1}\big\|\,,$$

and hence $Q^{n-1} = 0$ for large n, that is, for $n > 2\|P\| \cdot \|Q\|$. It follows inductively from (E5-1) that $Q^{n-m} = 0$ for $m = 1, \ldots, n$, i.e. $\mathrm{Id} = Q^0 = 0$, a contradiction if $X \neq \{0\}$. □

E5.2 Nonexistence of the inverse. For noncomplete normed spaces, the inverse in 5.7 in general does not exist.

Solution. We give a counterexample. Let $Y := \ell^2(\mathbb{R})$ and let

$$X := \left\{ x = (x_i)_{i\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}\ ;\ \text{only finitely many } x_i \neq 0 \right\} \subset \ell^2(\mathbb{R}) = Y\,,$$

i.e. X is equipped with the Y-norm. Let $\varepsilon > 0$. For the ***shift operator***

$$(Tx)_i := \begin{cases} 0 & \text{for } i = 1, \\ \varepsilon x_{i-1} & \text{for } i > 1, \end{cases}$$

it holds that $T \in \mathscr{L}(Y)$ and $\|T\| = \varepsilon$. Hence for $\varepsilon < 1$ we can apply 5.7 for Y and T, and obtain, for instance, that

$$(\mathrm{Id} - T)^{-1} e_1 = \sum_{n=0}^{\infty} T^n e_1 = \left(\varepsilon^{i-1}\right)_{i\in\mathbb{N}} \notin X\,.$$

On the other hand, $Tx \in X$ for $x \in X$. Hence 5.7 is not valid for X and $T\big|_X$ (X is not complete and $\overline{X} = Y$). □

E5.3 Unique extension of linear maps. Let $Z \subset X$ be a dense subspace and let $T \in \mathscr{L}(Z; Y)$. Then there exists a unique continuous extension $\widetilde{T}$ of T to X. Moreover, $\widetilde{T} \in \mathscr{L}(X; Y)$.

Solution. T is uniformly continuous on Z (in fact Lipschitz continuous with Lipschitz constant $\|T\|$). Hence, on recalling E4.18,

$$\widetilde{T}x := \lim_{z\in Z\,:\,z\to x} Tz \quad \text{for } x \in X$$

defines a unique continuous extension of T to X. In addition, the linearity of T carries over to $\widetilde{T}$. □

E5.4 Limit of linear maps. Let $(T_k)_{k\in\mathbb{N}}$ be a bounded sequence in $\mathscr{L}(X;Y)$ and let $D \subset X$ be dense. If there exists

$$\lim\nolimits_{k\to\infty} T_k x \quad \text{for } x \in D, \tag{E5-2}$$

then there exists

$$Tx := \lim_{k\to\infty} T_k x \quad \text{for all } x \in X$$

and $T \in \mathscr{L}(X;Y)$.

Solution. Let $\|T_k\| \le C < \infty$ for all k and let $Z := \operatorname{span}(D)$. Then it follows from (E5-2) that

$$Tz := \lim_{k\to\infty} T_k z$$

exists for all $z \in Z$, and that T is linear on Z. Since

$$\|Tz\| = \lim_{k\to\infty} \|T_k z\| \le C\|z\|\,,$$

it holds that $T \in \mathscr{L}(Z;Y)$. Let $\widetilde{T} \in \mathscr{L}(X;Y)$ be the unique extension of T to X from E5.3. Then it holds for all $x \in X$ and $z \in Z$ that

$$\begin{aligned} \left\|\widetilde{T}x - T_k x\right\| &\le \left\|\widetilde{T}z - T_k z\right\| + \left(\left\|\widetilde{T}\right\| + C\right)\|x - z\| \\ &\longrightarrow \left(\left\|\widetilde{T}\right\| + C\right)\|x - z\| \quad \text{as } k \to \infty. \end{aligned}$$

As $\overline{Z} = X$, we can choose $\|x - z\|$ arbitrarily small. This shows that

$$\widetilde{T}x = \lim_{k\to\infty} T_k x \quad \text{for all } x \in X.$$

□

E5.5 Pointwise convergence of operators. Let $T, T_k \in \mathscr{L}(X;Y)$, $k \in \mathbb{N}$, with $\|T_k\| \le C < \infty$ and let $D \subset X$ be dense. If for all $x \in D$

$$T_k x \longrightarrow Tx \quad \text{as } k \to \infty,$$

then this also holds for all $x \in X$.

Solution. See the second part of the solution of E5.4. □

E5.6 Convergence of operators. Let T_k be defined as in 5.6(6) with $1 \le p < \infty$. Does it hold that $T_k \longrightarrow \mathrm{Id}$ as $k \to \infty$ in the space $\mathscr{L}\big(L^p(\mathbb{R}^n)\big)$?

Solution. No! As an example, let $n = 1$ and $\varphi_k = \psi_{\varepsilon_k}$ with $\varepsilon_k \to 0$ as $k \to \infty$, where $\psi_\varepsilon(x) := \frac{1}{2\varepsilon}$ for $|x| < \varepsilon$ and $\psi_\varepsilon(x) := 0$ for $|x| > \varepsilon$. Then consider $T_k\varphi_k = \psi_{\varepsilon_k} * \psi_{\varepsilon_k}$. Direct calculations yield that

$$\begin{aligned}
&\psi_\varepsilon * \psi_\varepsilon(x) = \max\Big(0, \frac{1}{2}\varepsilon\big(1 - \frac{|x|}{2\varepsilon}\big)\Big)\,,\\
&\|\psi_\varepsilon\|_{L^p} = (2\varepsilon)^{\frac{1}{p}-1}\,,\\
&\|\psi_\varepsilon * \psi_\varepsilon - \psi_\varepsilon\|_{L^p} = (1+p)^{-\frac{1}{p}} \cdot (4\varepsilon)^{\frac{1}{p}-1}\,.
\end{aligned}$$

Consequently,

$$\|T_k - \mathrm{Id}\| \ge \frac{\|T_k\psi_{\varepsilon_k} - \psi_{\varepsilon_k}\|_{L^p}}{\|\psi_{\varepsilon_k}\|_{L^p}} = \frac{1}{2}\Big(\frac{1+p}{2}\Big)^{-\frac{1}{p}} > 0\,.$$

□