Functional Analysis 2

Lecture Script

Friedrich Philipp

TU Berlin, WS 2014/15

Live-LaTeXed by Maximilian König. A big THANKS goes to him!

Proofread by Friedrich Philipp and Maximilian König. However, there might still be hidden mistakes or typos in the text. Corrections are welcome at fmphilipp@gmail.com or maxkoe@math.tu-berlin.de.

Contents

1. Spectral Theory for Bounded Selfadjoint Operators

Throughout this chapter, let H be a Hilbert space over $\mathbb{C}$ with scalar product $\langle \cdot, \cdot \rangle$. The set of all bounded linear operators from H to H is denoted by $L(H)$. The range (i.e., image) and the kernel (i.e., null space) of $T \in L(H)$ are denoted by $\operatorname{ran} T$ and $\ker T$, respectively.

1.1. Basic Spectral Properties of Selfadjoint Operators

The following definition covers the basic notions of spectral theory for bounded linear operators in Hilbert spaces.

Definition 1.1. Let $T \in L(H)$.

(i) The **spectrum** of T is defined by

$$\sigma(T) := \{\lambda \in \mathbb{C} : T - \lambda I \text{ is not bijective}\}^{1}.$$

(ii) $\lambda \in \mathbb{C}$ is called an **eigenvalue** of T if $\ker(T - \lambda) \neq \{0\}$.

(iii) $\varrho(T) := \mathbb{C} \setminus \sigma(T)$ is called the **resolvent set** of T. For $\lambda \in \varrho(T)$, the operator $(T - \lambda)^{-1} \in L(H)$[2] is called the **resolvent** of T in λ.

(iv) $r(T) := \sup\{|\lambda| : \lambda \in \sigma(T)\}$ is called the **spectral radius** of T.

Let us collect some facts proved in the course Functional Analysis I (see [5]).

Theorem 1.2. *Let $T \in L(H)$. Then the following statements hold.*

(i) $r(T) \leq \|T\|$.

(ii) $\sigma(T)$ *is compact. In particular,* $\varrho(T)$ *is open.*

(iii) $\sigma(T) \neq \varnothing$ (*since in this chapter we only consider Hilbert spaces over* $\mathbb{C}$) *and*

$$r(T) = \lim_{n\to\infty} \|T^n\|^{\frac{1}{n}}.$$

(iv) *If* $T = T^*$ *then* $\|T\| = \sup\{|\langle Tx, x\rangle| : x \in H,\ \|x\| = 1\}$.

Definition 1.3. Let $T \in L(H)$. The set

$$W(T) := \{\langle Tx, x\rangle : x \in H, \|x\| = 1\}$$

is called the **numerical range** of T.

Lemma 1.4. *For $T \in L(H)$ we have $\sigma(T) \subset \overline{W(T)}$.*

[1] As it is usual in spectral theory, we will write for short $T - \lambda$ instead of $T - \lambda I$.
[2] The inverse operator is bounded by the open mapping theorem.

Proof. Let $\lambda \notin \overline{W(T)}$. Then $d := \operatorname{dist}\left(\lambda, \overline{W(T)}\right) > 0$, and for $x \in H$, $\|x\| = 1$, we have

$$d \le |\langle Tx, x\rangle - \lambda| = |\langle Tx, x\rangle - \langle \lambda x, x\rangle| = |\langle (T-\lambda)x, x\rangle| \le \|(T-\lambda)x\|.$$

This implies $\|(T-\lambda)x\| \ge d\|x\|$ for all $x \in H$. Hence $T-\lambda$ is injective and $\operatorname{ran}(T-\lambda)$ is closed. To see that $\operatorname{ran}(T-\lambda) = H$, let $x_0 \in \operatorname{ran}(T-\lambda)^\perp$, $\|x_0\| = 1$. Then

$$\langle Tx_0, x_0\rangle - \lambda = \langle (T-\lambda)x_0, x_0\rangle = 0,$$

which implies $\lambda \in W(T)$, contradicting our initial assumption. □

Recall that T is called **selfadjoint** if $T = T^*$. Equivalently, $\langle Tx, y\rangle = \langle x, Ty\rangle$ for all $x, y \in H$. In this case, $\langle Tx, x\rangle = \langle x, Tx\rangle = \overline{\langle Tx, x\rangle}$ and hence $W(T) \subset \mathbb{R}$. This implies that $\sigma(T) \subset \mathbb{R}$. We define

$$m(T) := \inf W(T) \qquad \text{and} \qquad M(T) := \sup W(T).$$

From Lemma 1.4 and Theorem 1.2(iv), we conclude

$$\sigma(T) \subset [m(T), M(T)] \qquad \text{and} \qquad \|T\| = \max\{|m(T)|, |M(T)|\}. \tag{1.1}$$

Proposition 1.5. *Let $T = T^* \in L(H)$. Then the following holds.*

(i) $r(T) = \|T\|$.

(ii) $m(T) = \min \sigma(T)$.

(iii) $M(T) = \max \sigma(T)$.

Proof. For the proof of (i), it suffices to show that $\|T^{2^k}\| = \|T\|^{2^k}$ for all $k \in \mathbb{N}_0$. For $k = 0$, this is trivial. For $k > 0$, we define $S := T^{2^{k-1}}$. We assume that the above identity holds for $k-1$. Since S is selfadjoint, so is S^2, which gives

$$\left\|T^{2^k}\right\| = \|S^2\| = \sup\left\{\left|\langle S^2x, x\rangle\right| : \|x\| = 1\right\} = \|S\|^2 = \left\|T^{2^{k-1}}\right\|^2 = \|T\|^{2^k}.$$

Let us prove (ii) and (iii). For this, we only need to show that $m := m(T)$ and $M := M(T)$ are contained in $\sigma(T)$. First, assume that $T \ge 0$, i.e., $\langle Tx, x\rangle \ge 0$ for all $x \in H$. Then (1.1) and (i) imply $M = \|T\| = r(T)$. Hence, $M \in \sigma(T)$.

Now, let T be arbitrary, and define the operators $S_1 := T - m$ and $S_2 := -T + M$. It is easy to see that $m(S_1) = m(S_2) = 0$ and $M(S_1) = M(S_2) = M - m$. In particular, $S_1 \ge 0$ and $S_2 \ge 0$. Therefore, $M - m \in \sigma(S_1) \cap \sigma(S_2)$. Hence,

$$\sigma(S_1) = \{\lambda - m : \lambda \in \sigma(T)\} \quad \text{and} \quad \sigma(S_2) = \{-\lambda + M : \lambda \in \sigma(T)\}$$

implies that $m, M \in \sigma(T)$. □

1.2. Continuous Functional Calculus

A functional calculus is a means to define $f(T)$ for functions f and operators T.

Example 1.6. (a) Let $p(z) = \sum_{k=0}^{n} a_k z^k$ be a polynomial. Then we define $p(T) := \sum_{k=0}^{n} a_k T^k$ with $T^0 := I$.

(b) If $R(z) = \frac{p(z)}{q(z)}$ with two polynomials p and q, and the zeros of q are contained in $\varrho(T)$, then define

$$R(T) := p(T)q(T)^{-1}.$$

To justify this definition, we write $q(z) = q_0(z - z_0) \cdot \ldots \cdot (z - z_n)$ with $q_0 \in \mathbb{C}$ and $z_0, \ldots, z_n \in \varrho(T)$. Hence, $q(T) = q_0(T - z_0) \cdot \ldots \cdot (T - z_n)$ is invertible.

(c) Let $f(z) = \sum_{k=0}^{\infty} a_k z^k$ be a power series with radius of convergence r_f. Then for $T \in L(H)$ with $r(T) < r_f$, the series

$$f(T) := \sum_{k=0}^{\infty} a_k T^k$$

converges in $L(H)$. Indeed, if $\varepsilon > 0$ and $\left\|T^k\right\|^{\frac{1}{k}} < r_f - \varepsilon$ for $k \geq n_0$, then

$$\left\| \sum_{k=n}^{n+m} a_k T^k \right\| \leq \sum_{k=n}^{n+m} |a_k| \cdot \left(\left\|T^k\right\|^{\frac{1}{k}} \right)^k < \sum_{k=n}^{n+m} |a_n| (r_f - \varepsilon)^k \overset{n,m\to\infty}{\longrightarrow} 0.$$

(d) $e^T = \sum_0^{\infty} \frac{T^k}{k!}$ is defined for each $T \in L(H)$.

Examples (b) and (c) indicate that in order to define $f(T)$, the function f should at least be defined on $\sigma(T)$.

Lemma 1.7. *For $T \in L(H)$ and a polynomial p, we have*

$$\sigma(p(T)) = p(\sigma(T)) = \{p(\lambda) : \lambda \in \sigma(T)\}.$$

Proof. Exercise. □

Lemma 1.8 (Tietze-Urysohn-Lemma for Metric Spaces). *Let X be a metric space and $A \subset X$ a closed subset. Then any continuous function $f : A \to [a, b]$ admits a continuous extension $F : X \to [a, b]$ with $\|F\|_\infty = \|f\|_\infty$.*

Proof. It suffices to prove the claim for $a = 1$, $b = 2$. Then we define $F(x) = f(x)$ for $x \in A$ and

$$F(x) = \frac{\inf \{f(y) \operatorname{dist}(x, y) : y \in A\}}{\inf \{\operatorname{dist}(x, y) : y \in A\}}, \quad x \notin A.$$

Proving that F indeed maps into $[1, 2]$ and is coninuous will be an exercise. □

Theorem 1.9 (Continuous Functional Calculus). *Let $T \in L(H)$ be selfadjoint. Then there is a unique mapping $\Phi : C(\sigma(T)) \to L(H)$ with the following properties:*

(i) $\Phi(\mathbb{1}) = I$ *and* $\Phi(id) = T$.

(ii) Φ *is an involutive homomorphism between algebras, i.e.,*

  (a) Φ *is linear.*

  (b) Φ *is multiplicative, i.e.* $\Phi(fg) = \Phi(f)\Phi(g)$.

  (c) Φ *is involutive, i.e.* $\Phi(\overline{f}) = \Phi(f)^*$.

(iii) Φ *is continuous.*

The mapping Φ in Theorem 1.9 is called the **continuous functional calculus** of T, and we write $f(T) := \Phi(f)$.

Proof of Theorem 1.9. Let $\mathbb{P}$ denote the set of polynomials $f : \sigma(T) \to \mathbb{C}$, $f(t) = \sum_{k=0}^n a_k t^k$, $t \in \sigma(T)$. Note that $f \in \mathbb{P}$ implies $\overline{f} \in \mathbb{P}$ since f is defined on $\sigma(T) \subset \mathbb{R}$. For $f \in \mathbb{P}$, define $\Phi_0(f) := f(T)$ (cf. Example 1.6 (a)). It is easily seen that Φ_0 satisfies (i) and (ii). But also (iii), since for a polynomial f as above,

$$\begin{aligned}\|\Phi_0(f)\|^2 &= \|\Phi_0(f)^*\Phi_0(f)\| = \|\Phi_0(\overline{f})\Phi_0(f)\| = \|\Phi_0(\overline{f}f)\| = r(\Phi_0(\overline{f}f)) \\ &= \sup\left\{|\lambda| : \lambda \in \sigma(\Phi_0(\overline{f}f))\right\} = \sup\left\{|\lambda| : \lambda \in (\overline{f}f)(\sigma(T))\right\} \\ &= \sup\left\{|f(\lambda)|^2 : \lambda \in \sigma(T)\right\} = \|f\|_\infty^2,\end{aligned}$$

where $\|f\|_\infty := \sup\{|f(t)| : t \in \sigma(T)\}$. Hence, $\Phi_0 : \mathbb{P} \to L(H)$ is an isometry. It follows from a well known theorem of Weierstraß and Lemma 1.8 that $\mathbb{P}$ is dense in $C(\sigma(T))$. Thus, Φ_0 extends to an isometry Φ on $C(\sigma(T))$.

To show (ii)(c) for Φ, let $f \in C(\sigma(T))$ and $(f_n) \subset \mathbb{P}$ such that $\|f_n - f\|_\infty \to 0$. Then

$$\Phi(\overline{f}) = \Phi\left(\lim_{n\to\infty} \overline{f_n}\right) = \lim_{n\to\infty} \Phi_0(\overline{f_n}) = \lim_{n\to\infty} \Phi_0(f_n)^* = \left(\lim_{n\to\infty} \Phi_0(f_n)\right)^* = \Phi(f)^*.$$

(ii)(b) is proved similarly.

If $\Psi : C(\sigma(T)) \to L(H)$ is another mapping with (i)–(iii), then for $f \in \mathbb{P}$, $f(t) = \sum_{k=0}^n a_k t^k$, we have

$$\Psi(f) = \Psi\left(\sum_{k=0}^n a_k id^k\right) = \sum_{k=0}^n a_k \Psi\left(id^k\right) = \sum_{k=0}^n a_k \Psi(id)^k = \sum_{k=0}^n a_k T^k = \Phi(f).$$

Thus, $\Psi|_{\mathbb{P}} = \Phi|_{\mathbb{P}}$ and $\mathbb{P}$ is dense in $C(\sigma(T))$. This implies $\Psi = \Phi$. $\square$

Theorem 1.10. *Let $T \in L(H)$ be selfadjoint and $f \in C(\sigma(T))$. Then the following statements hold.*

(i) $\|f(T)\| = \|f\|_\infty$.

(ii) *If f is real-valued and non-negative then $f(T)$ is non-negative, i.e. $\langle Tx, x\rangle \geq 0$ for all $x \in H$.*

(iii) *If $Tx = \lambda x$ then $f(T)x = f(\lambda)x$.*

(iv) $\sigma(f(T)) = f(\sigma(T))$.

(v) *$f(T)$ is normal, i.e. $f(T)^*f(T) = f(T)f(T)^*$.*

(vi) *$f(T)$ is selfadjoint if and only if f is real-valued.*

(vii) *If $W \in L(H)$ commutes with T, then it also commutes with $f(T)$.*

Proof. (v) is clear, and (i) was a byproduct in the proof of Theorem 1.9. (vi) follows from

$$\|f - \overline{f}\|_\infty = \|\Phi(f) - \Phi(\overline{f})\| = \|\Phi(f) - \Phi(f)^*\|, \quad f \in C(\sigma(T)).$$

For $f \in C(\sigma(T))$, $f \geq 0$, we also have $\sqrt{f} \in C(\sigma(T))$, and hence

$$\langle \Phi(f)x, x\rangle = \left\langle \Phi\left(\sqrt{f}^2\right) x, x\right\rangle = \left\langle \Phi\left(\sqrt{f}\right) \Phi\left(\sqrt{f}\right) x, x\right\rangle = \|\Phi(\sqrt{f})x\|^2 \geq 0.$$

This implies (ii). Statements (iii) and (vii) can easily be seen by approximating f uniformly by polynomials on $\sigma(T)$.

For the proof of the inclusion "$\subset$" in (iv), let $\mu \notin f(\sigma(T))$. Then $g := (f - \mu)^{-1} \in C(\sigma(T))$. We have $g(f-\mu) = (f-\mu)g = \mathbb{1}$. Hence $g(T)(f(T)-\mu) = (f(T)-\mu)g(T) = I$. So, $\mu \in \varrho(f(T))$. For the converse inclusion, assume $\mu = f(\lambda) \in f(\sigma(T))$ for some $\lambda \in \sigma(T)$. Choose a sequence of polynomials (f_n) such that $\|f - f_n\|_\infty \leq \frac{1}{n}$ for each $n \in \mathbb{N}$. We have $f_n(\lambda) \in f_n(\sigma(T)) = \sigma(f_n(T))$. In an exercise it will be shown, that there exists a sequence $(x_n) \subset H$ with $\|x_n\| = 1$ and $\|(f_n(T) - f_n(\lambda))\, x_n\| \leq \frac{1}{n}$. Hence

$$\begin{aligned}\|(f(T) - \mu)\, x_n\| &\leq \|f(T)x_n - f_n(T)x_n\| + \|f_n(T)x_n - f_n(\lambda)x_n\| + |f_n(\lambda) - f(\lambda)| \\ &\leq 2\|f_n - f\|_\infty + \frac{1}{n} \to 0.\end{aligned}$$

This shows, that $\mu \in \sigma(f(T))$. □

1.3. Measurable Functional Calculus

Let $T = T^*$ and $H = \mathbb{C}^N$. So, $T = \sum_{i=1}^n \lambda_i E_i$, where the λ_i are the (mutually distinct) eigenvalues of T, and E_i is the orthogonal projection onto $\ker(T - \lambda_i)$. Let $f : \sigma(T) \to \mathbb{C}$ be an arbitrary function. Note that $\sigma(T)$ is a finite set. Hence, every such f is an element of $C(\sigma(T))$. It can now easily be seen that

$$f(T) = \sum_{i=1}^n f(\lambda_i)E_i.$$

Given T (e.g., as a matrix representation), an interesting question is how to extract the projections E_i from T. This can be done by using $f = \chi_{\{\lambda_j\}}$:

$$\chi_{\{\lambda_j\}}(T) = \sum_{i=1}^{n} \chi_{\{\lambda_j\}}(\lambda_i) E_i = E_j.$$

If now $T = T^* \in L(H)$ with $\dim H = \infty$, and λ is not an isolated point of $\sigma(T)$, then $\chi_{\{\lambda\}} : \sigma(T) \to \mathbb{C}$ is not continous. In general, χ_A is not an element of $C(\sigma(T))$ for a Borel set $A \subset \sigma(T)$. In this section, we are going to extend the continuous functional calculus to bounded measurable functions so that, in particular, $\chi_A(T)$ is defined.

Definition 1.11. A sesquilinear form $g : H^2 \to \mathbb{C}$ is called **bounded** if there exists a $C > 0$ such that $|g(x, y)| < C\|x\|\|y\|$ for all $x, y \in H$. As for linear operators, one sees that g is bounded if and only if it is continuous. The **quadratic form** of g is defined by $\varphi(x) := g(x, x)$ for all $x \in H$.

If $g : H^2 \to \mathbb{C}$ is sesquilinear and φ is its quadratic form, it is easily seen that the following holds:

$$g(x, y) := \frac{1}{4}\big(\varphi(x + y) - \varphi(x - y) + i\varphi(x + iy) - i\varphi(x - iy)\big), \quad x, y \in H. \tag{1.2}$$

This formula is called the **polarization formula**. It even holds for functions $g : H^2 \to \mathbb{C}$ satisfying

$$g(-x, y) = g(x, -y) = -g(x, y) \quad \text{and} \quad g(ix, y) = g(x, -iy) = ig(x, y) \tag{1.3}$$

for all $x, y \in H$. Hence, a sesquilinear form is uniquely defined by its quadratic form. In particular, two sesquilinear forms with the same quadratic form must be equal.

Lemma 1.12. *A function $\varphi : H \to \mathbb{C}$ is the quadratic form of a sesquilinear form if and only if the following two conditions are satisfied:*

(a) *$\varphi(\lambda x) = |\lambda|^2 \varphi(x)$ for all $\lambda \in \mathbb{C}$ and $x \in H$.*

(b) *$\varphi(x + y) + \varphi(x - y) = 2\varphi(x) + 2\varphi(y)$ for all $x, y \in H$.*

In this case, the associated sesquilinear form is bounded if and only if the following holds:

(c) *There exists $C > 0$ such that $|\varphi(x)| \le C\|x\|^2$ for all $x \in H$.*

Proof. Clearly, the quadratic form φ of a bounded sesquilinear form satisfies (a) and (b) and, in addition, (c) if the sesquilinear form is bounded. Conversely, let φ be a function satisfying (a) and (b). We define $g : H^2 \to \mathbb{C}$ by the polarization formula (1.2). Then, indeed, $g(x, x) = \varphi(x)$, $x \in H$. Let us prove that g is sesquilinear. For this, we observe that for $u, v, y \in H$ we have

$$\begin{aligned}
\varphi(u + v + y) &= -\varphi(u - (v + y)) + 2\varphi(u) + 2\varphi(v + y), \\
\varphi(u + v + y) &= -\varphi((u + v) - y) + 2\varphi(u + v) + 2\varphi(y).
\end{aligned}$$

From this, we conclude

$$\begin{aligned}2\varphi(u+v+y) &= -\varphi((u-y)-v)) - \varphi((u-y)+v) \\ &\qquad\qquad + 2\varphi(u) + 2\varphi(v+y) + 2\varphi(u+v) + 2\varphi(y) \\ &= -2\varphi(u-y) - 2\varphi(v) + 2\varphi(u) + 2\varphi(v+y) + 2\varphi(u+v) + 2\varphi(y),\end{aligned}$$

that is,

$$\varphi(u+v+y) = -\varphi(u-y) + \varphi(u+v) + \varphi(v+y) - \varphi(v) + \varphi(u) + \varphi(y). \tag{1.4}$$

Replacing y by $-y$ gives

$$\varphi(u+v-y) = -\varphi(u+y) + \varphi(u+v) + \varphi(v-y) - \varphi(v) + \varphi(u) + \varphi(y). \tag{1.5}$$

Replacing v by $-v$ in (1.5) yields

$$\varphi(u-v-y) = -\varphi(u+y) + \varphi(u-v) + \varphi(v+y) - \varphi(v) + \varphi(u) + \varphi(y). \tag{1.6}$$

Let us now define $\phi(x,y) := \varphi(x+y) - \varphi(x-y)$, $x, y \in H$. Then (1.4) and (1.5) as well as (1.4) and (1.6) give

$$\begin{aligned}\phi(u+v,y) &= \big(\varphi(u+y) - \varphi(u-y)\big) + \big(\varphi(v+y) - \varphi(v-y)\big) = \phi(u,y) + \phi(v,y), \\ \phi(u,v+y) &= \big(\varphi(u+y) - \varphi(u-y)\big) + \big(\varphi(u+v) - \varphi(u-v)\big) = \phi(u,y) + \phi(u,v).\end{aligned}$$

Thus, ϕ is additive in both components. The same can now be proved for g:

$$\begin{aligned}g(u+v,y) &= \frac{1}{4}\big(\phi(u+v,y) + i\phi(u+v,iy)\big) \\ &= \frac{1}{4}\big(\phi(u,y) + \phi(v,y) + i\phi(u,iy) + i\phi(v,iy)\big) \\ &= g(u,y) + g(v,y).\end{aligned}$$

and

$$\begin{aligned}g(u,v+y) &= \frac{1}{4}\big(\phi(u,v+y) + i\phi(u,iv+iy)\big) \\ &= \frac{1}{4}\big(\phi(u,v) + \phi(u,y) + i\phi(u,iv) + i\phi(u,iy)\big) \\ &= g(u,v) + g(u,y).\end{aligned}$$

Let $\lambda \in \mathbb{C}$. Then we have

$$\begin{aligned}g(\lambda x,x) &= \frac{1}{4}\big(\varphi((\lambda+1)x) - \varphi((\lambda-1)x) + i\varphi((\lambda+i)x) - i\varphi((\lambda-i)x)\big) \\ &= \frac{1}{4}\big(|\lambda+1|^2 - |\lambda-1|^2 + i|\lambda+i|^2 - i|\lambda-i|^2\big)\varphi(x) \\ &= \frac{1}{4}(4\operatorname{Re}\lambda + 4i\operatorname{Im}\lambda)g(x,x) = \lambda g(x,x).\end{aligned}$$

Now, define $s(x,y) := g(\lambda x, y)$, $x, y \in H$. It can easily be seen that g satisfies (1.3). Hence, also s does. Therefore, the polarization formula holds for s and its quadratic form φ_s. But we have $\varphi_s(x) = \lambda\varphi(x)$. It follows that

$$g(\lambda x, y) = s(x,y) = \frac{\lambda}{4}\big(\varphi(x+y) - \varphi(x-y) + i\varphi(x+iy) - i\varphi(x-iy)\big) = \lambda g(x,y).$$

The anti-homogenicity of g in the second component is proved similarly. Hence, g is indeed a sesquilinear form.

Assume now that φ also satisfies (c). Then,

$$\begin{aligned}|g(x,y)| &\le \frac{1}{4}\big(|\varphi(x+y)| + |\varphi(x-y)| + |\varphi(x+iy)| + |\varphi(x-iy)|\big)\\ &\le \frac{C}{4}\left(\|x+y\|^2 + \|x-y\|^2 + \|x+iy\|^2 + \|x-iy\|^2\right)\\ &= \frac{C}{4}\left(4\|x\|^2 + 4\|y\|^2\right) = C\left(\|x\|^2 + \|y\|^2\right).\end{aligned}$$

Thus, we have $\sup_{\|x\|=\|y\|=1} |g(x,y)| \le 2C$ and hence $|g(x,y)| \le 2C\|x\|\|y\|$. □

The next lemma is an easy application of the Riesz Representation Theorem for continuous linear functionals on a Hilbert space. Its proof is left to the reader.

Lemma 1.13. *If $g : H^2 \to \mathbb{C}$ is a bounded sesquilinear form then there exists a unique $T \in L(H)$ such that $g(x,y) = \langle Tx, y\rangle$ for all $x, y \in H$.*

For a compact set $\sigma \subset \mathbb{R}$, we denote by $B(\sigma)$ the space of all bounded measurable functions on σ. We equip $B(\sigma)$ with $\|\cdot\|_\infty$. Note that we do not consider equivalence classes in terms of equality a.e. on $B(\sigma)$.

Lemma 1.14. *The space $B(\sigma)$ has the following properties.*

(a) *$(B(\sigma), \|\cdot\|_\infty)$ is a Banach space.*

(b) *If $U \subset B(\sigma)$ is such that $C(\sigma) \subset U$, and for every sequence $(f_n) \subset U$ with $\sup \|f_n\|_\infty < \infty$ and $f_n \to f$ (pointwise), it holds that $f \in U$, then $U = B(\sigma)$.*

Proof. See [10, Lemma VII.1.5]. □

The following theorem is the main result of this section and extends the continuous functional calculus for selfadjoint operators.

Theorem 1.15 (Measurable Functional Calculus). *Let $T = T^* \in L(H)$. Then there exists a unique mapping $\Psi : B(\sigma(T)) \to L(H)$ with the following properties:*

(a) *$\Psi(\mathbb{1}) = I$ and $\Psi(\mathrm{id}) = T$.*

(b) *Ψ is an involutive homomorphism between algebras.*

(c) *Ψ is continuous.*

(d) *If* $(f_n) \subset B(\sigma(T))$ *is uniformly bounded, i.e.,* $\sup_{n\in\mathbb{N}} \|f_n\|_\infty < \infty$, *and* $f_n \to f$ *(pointwise), then* $\Psi(f_n)x \to \Psi(f)x$ *for all* $x \in H$.

The mapping Ψ in Theorem 1.15 is called the **measurable functional calculus** of T, and we write $f(T) := \Psi(f)$.

Proof of Theorem 1.15. Uniqueness: (a)–(c) imply that $\Psi|_{C(\sigma(T))} = \Phi$. Now, (d) and Lemma 1.14 yield uniqueness.

Existence: For $x \in H$, consider the functional $\ell_x : C(\sigma(T), \mathbb{R}) \to \mathbb{R}$, defined by $\ell_x f := \langle \Phi(f)x, x\rangle$. Obviously ℓ_x is linear in f, non-negative (i.e., $f \geq 0$ implies $\ell_x f \geq 0$), and we have

$$|\ell_x f| \leq \|\Phi(f)\| \cdot \|x\|^2 = \|x\|^2 \cdot \|f\|_\infty,$$

which implies $\ell_x \in C(\sigma(T), \mathbb{R})'$. By the Riesz Representation Theorem (Theorem A.1), there exists a finite positive Borel measure μ_x on $\sigma(T)$ with $\mu_x(\sigma(T)) = \|\ell_x\| \leq \|x\|^2$ and

$$\langle \Phi(f)x, x\rangle = \int_{\sigma(T)} f \,\mathrm{d}\mu_x, \qquad \forall\, f \in C(\sigma(T), \mathbb{R}).$$

Hence, if $f \in C(\sigma(T))$, $f = g + ih$, then

$$\langle \Phi(f)x, x\rangle = \langle \Phi(g)x, x\rangle + i\,\langle \Phi(h)x, x\rangle = \int_{\sigma(T)} g \,\mathrm{d}\mu_x + i \int_{\sigma(T)} h \,\mathrm{d}\mu_x = \int_{\sigma(T)} f \,\mathrm{d}\mu_x.$$

For $f \in C(\sigma(T))$ and $\lambda \in \mathbb{C}$, we have

$$\int f \,\mathrm{d}\mu_{\lambda x} = \langle \Phi(f)\lambda x, \lambda x\rangle = |\lambda|^2 \,\langle \Phi(f)x, x\rangle = |\lambda|^2 \int_{\sigma(T)} f \,\mathrm{d}\mu_x.$$

Now set

$$U = \left\{ f \in B(\sigma(T)) : \int f \,\mathrm{d}\mu_{\lambda x} = |\lambda|^2 \int f \,\mathrm{d}\mu_x \right\}.$$

We have $C(\sigma(T)) \subset U$. If $(f_n) \subset U$ with $\sup \|f_n\|_\infty < \infty$ and $f_n \to f$ pointwise, then by Lebesgue's Theorem we get

$$\int f \,\mathrm{d}\mu_{\lambda x} = \lim_{n\to\infty} \int f_n \,\mathrm{d}\mu_{\lambda x} = \lim_{n\to\infty} |\lambda|^2 \int f_n \,\mathrm{d}\mu_x = |\lambda|^2 \int f \,\mathrm{d}\mu_x.$$

Thus, $f \in U$. By Lemma 1.14, $U = B(\sigma(T))$ follows. Similarly, one shows that

$$\int f \,\mathrm{d}\mu_{x+y} + \int f \,\mathrm{d}\mu_{x-y} = 2\int f \,\mathrm{d}\mu_x + 2\int f \,\mathrm{d}\mu_y, \qquad f \in B(\sigma(T)),\ x, y \in H.$$

Moreover, we have

$$\left|\int f \,\mathrm{d}\mu_x\right| \;\leq\; \int |f| \,\mathrm{d}\mu_x \;\leq\; \|f\|_\infty \mu_x(\sigma(T)) \;\leq\; \|f\|_\infty \|x\|^2, \qquad f \in B(\sigma(T)),\ x \in H.$$

For fixed $f \in B(\sigma(T))$, define $\varphi_f(x) := \int f \,\mathrm{d}\mu_x$, $x \in H$. From Lemma 1.12, it follows that there exists a bounded sesquilinear form g_f on H such that $\varphi_f(x) = g_f(x, x)$ for all $x \in H$.

By Lemma 1.13, there exists a unique $\Psi(f) \in L(H)$ such that $g_f(x,y) = \langle \Psi(f)x, y\rangle$ for all $x, y \in H$. Alltogether, we have

$$\langle \Psi(f)x, x\rangle = \int f \, \mathrm{d}\mu_x, \qquad f \in B(\sigma(T)),\ x \in H. \tag{1.7}$$

Thus, for $f \in C(\sigma(T))$ and $x \in H$, $\langle \Phi(f)x, x\rangle = \int f \, \mathrm{d}\mu_x = \langle \Psi(f)x, x\rangle$, which implies

$$\langle (\Phi(f) - \Psi(f))x, x\rangle = 0.$$

By polarization, we get $\Phi(f) = \Psi(f)$, and thus (a) holds.

Linearity: In order to see that $\Psi : B(\sigma(T)) \to L(H)$ is linear, let $f, g \in B(\sigma(T))$ and $\lambda \in \mathbb{C}$. Then

$$\langle \Psi(\lambda f + g)x, x\rangle = \int (\lambda f + g) \, \mathrm{d}\mu_x = \lambda \langle \Psi(f)x, x\rangle + \langle \Psi(g)x, x\rangle = \langle (\lambda\Psi(f) + \Psi(g))x, x\rangle .$$

Polarization gives $\Psi(\lambda f + g) = \lambda\Psi(f) + \Psi(f)$.

Boundedness: Let $f \in B(\sigma(T))$ and $x \in H$. From the proof of Lemma 1.12, we obtain

$$\|\Psi(f)x\|^2 = g_f(x, \Psi(f)x) \leq 2\|f\|_\infty \|x\| \|\Psi(f)x\|.$$

This implies $\|\Psi(f)x\| \leq 2\|f\|_\infty \|x\|$.

Multiplicativity: Let (f_n) be a sequence as in (d). Then $\langle \Psi(f_n)x, y\rangle \to \langle \Psi(f)x, y\rangle$ for all $x, y \in H$. Indeed, for $x = y$ this follows from Lebesgue's Theorem, and by polarization for $x \neq y$. Let $g \in C(\sigma(T))$, and define

$$U = \{f \in B(\sigma(T)) : \Psi(fg) = \Psi(f)\Psi(g)\} .$$

Obviously, we have $C(\sigma(T)) \subset U$. Let $f_n \in U$ such that $\sup \|f_n\|_\infty < \infty$ and $f_n \to f$ pointwise. Then

$$\langle \Psi(f)\Psi(g)x, y\rangle = \lim_{n\to\infty} \langle \Psi(f_n)\Psi(g)x, y\rangle = \lim_{n\to\infty} \langle \Psi(f_n g)x, y\rangle = \langle \Psi(fg)x, y\rangle .$$

This shows $f \in U$. By Lemma 1.14, we have $\Psi(fg) = \Psi(f)\Psi(g)$ for all $f \in B(\sigma(T))$ and all $g \in C(\sigma(T))$. For $f \in B(\sigma(T))$, define

$$V = \{g \in B(\sigma(T)) : \Psi(fg) = \Psi(f)\Psi(g)\} .$$

By the above, $C(\sigma(T)) \subset V$. If $(g_n) \subset V$, $\sup \|g_n\| < \infty$, and $g_n \to g$ pointwise, it follows similarly as above that $g \in V$. Hence, Ψ is multiplicative. The involutivity can be proved in a similar manner.

For the proof of (iv), recall that $z_n \to z$ if $z_n \to z$ weakly and $\|z_n\| \to \|z\|$. It is therefore sufficient to prove that

$$\|\Psi(f_n)x\| \to \|\Psi(f)x\|.$$

But this follows from

$$\begin{aligned}\|\Psi(f_n)x\|^2 = \langle \Psi(f_n)x, \Psi(f_n)x\rangle = \langle \Psi(f_n)^*\Psi(f_n)x, x\rangle = \left\langle \Psi\left(|f_n|^2\right)x, x\right\rangle \\ \to \left\langle \Psi\left(|f|^2\right)x, x\right\rangle = \|\Psi(f)x\|^2.\end{aligned}$$

The theorem is proved. □

The rest of this section is devoted to the properties of the measurable functional calculus.

Proposition 1.16. *Let $T \in L(H)$ be selfadjoint and $f \in B(\sigma(T))$. Then the following statements hold.*

(i) $\|f(T)\| \leq \|f\|_\infty$.

(ii) *If $f \geq 0$, then $f(T) \geq 0$.*

(iii) *If $Tx = \lambda x$, then $f(T)x = f(\lambda)x$.*

(iv) $\sigma(f(T)) \subset \overline{f(\sigma(T))}$.

Proof. (i). Let f be real-valued. Then $\Psi(f) = \Psi(f)^*$. Hence

$$\begin{aligned}\|\Psi(f)\| &= \sup\{|\langle\Psi(f)x, x\rangle| : x \in H, \|x\| = 1\} = \sup\left\{\left|\int f \, \mathrm{d}\mu_x\right| : x \in H, \|x\| = 1\right\}\\ &\leq \sup\{\|f\|_\infty \mu_x(\sigma(T)) : x \in H, \|x\| = 1\} \leq \|f\|_\infty.\end{aligned}$$

For general f, we have $\|\Psi(f)\|^2 = \|\Psi(f)\Psi(f)^*\| = \|\Psi(|f|^2)\| \leq \||f|^2\|_\infty = \|f\|_\infty^2$.

(ii). We have

$$\langle\Psi(f)x, x\rangle = \int f \, \mathrm{d}\mu_x \geq 0.$$

(iii). Let $Tx = \lambda x$ and set

$$U = \{f \in B(\sigma(T)) : \Psi(f)x = f(\lambda)x\}.$$

Then $C(\sigma(T)) \subset U$. Let $f_n \in U$ such that $\sup_n \|f_n\|_\infty < \infty$ and $f_n \to f$ pointwise. Then $\Psi(f_n)x \to \Psi(f)x$ implies $f \in U$. The application of Lemma 1.13 yields the claim.

(iv). Let $\mu \notin \overline{f(\sigma(T))}$. Then for each $\lambda \in \sigma(T)$,

$$|\mu - f(\lambda)| \geq \operatorname{dist}\left(\mu, \overline{f(\sigma(T))}\right) =: d > 0.$$

Set $g(\lambda) = \frac{1}{\mu - f(\lambda)}$, $\lambda \in \sigma(T)$. Then $\|g\|_\infty \leq \frac{1}{d}$, and thus $g \in B(\sigma(T))$. Moreover,

$$(\mu - f(\lambda))g(\lambda) = g(\lambda)(\mu - f(\lambda)) = 1, \quad \lambda \in \sigma(T).$$

Finally $(\mu - f(T))g(T) = g(T)(\mu - f(T)) = \mathrm{id}$. Thus $\mu \in \varrho(f(T))$. □

Theorem 1.17. *Let $T \in L(H)$ be selfadjoint, $g \in B(\sigma(T))$ a real-valued function, and $f \in B(\overline{g(\sigma(T))})$. Then*

$$(f \circ g)(T) = f(g(T)).$$

Proof. Let Ψ_T be the measurable functional calculus of T and $\Psi_{g(T)}$ the measurable fuctional calculus of $g(T)$. We have to show, that

$$\Psi_T(f \circ g) = \Psi_{g(T)}(f).$$

If f is a polynomial, say

$$f(t) = \sum_{k=1}^{n} a_k t^k.$$

then $f \circ g = \sum_{k=1}^{n} a_k g^k$. Therefore

$$\Psi_{g(T)}(f) = \sum_{k=1}^{n} a_k g(T)^k = \sum_{k=1}^{n} a_k \Psi_T(g)^k = \Psi_T\left(\sum_{k=1}^{n} a_k g^k\right) = \Psi_T(f \circ g).$$

Now let $f \in C(\overline{g(\sigma(T))})$ and (p_n) a sequence of polynomials uniformly convergent to f on $\overline{g(\sigma(T))}$. Then

$$\|\Psi_T(f \circ g) - \Psi_T(p_n \circ g)\| = \|\Psi_T((f - p_n) \circ g)\| \leq \|(f - p_n) \circ g\|_\infty \leq \|f - p_n\|_\infty \to 0,$$

which shows

$$\Psi_{g(T)}(f) = \lim_{n\to\infty} \Psi_{g(T)}(p_n) = \lim_{n\to\infty} \Psi_T(p_n \circ g) = \Psi_T(f \circ g).$$

Finally, for $x \in H$ define

$$U = \left\{ f \in B(\overline{g(\sigma(T))}) : \Psi_{g(T)}(f)x = \Psi_T(f \circ g)x \right\}.$$

We just proved $C(\overline{g(\sigma(T))}) \subset U$. Let $f_n \in U$, $\sup \|f_n\|_\infty < \infty$, and $f_n \to f$ pointwise. Note, that $\sup \|f_n \circ g\|_\infty < \infty$ and $f_n \circ g \to f \circ g$. Thus

$$\Psi_{g(T)}(f)x = \lim_{n\to\infty} \Psi_{g(T)}(f_n)x = \lim_{n\to\infty} \Psi_T(f_n \circ g)x = \Psi_T(f \circ g)x.$$

This shows $f \in U$. Using again Lemma 1.14, the theorem is proved. □

2. Spectral Theory for Unbounded Selfadjoint Operators

Many linear operators arising from applications in, e.g., quantum mechanics, are unbounded and not defined on the entire Hilbert space. In this chapter, we shall see that one can still (very well) deal with these kind of operators. We will define the notion of selfadjointness for unbounded operators and introduce the reader into the well-developed spectral theory for selfadjoint operators.

2.1. Unbounded Operators

Example 2.1. (a) Let $T = T^* \in L(H)$. Then $\overline{\operatorname{ran} T} = (\ker T)^\perp$. Assume $\ker T = \{0\}$. Then $\operatorname{ran} T$ is dense in H. However, it is possible that $\operatorname{ran} T \neq H$. So T^{-1} can only be defined on $\operatorname{ran} T$ and is not bounded if $\operatorname{ran} T \neq H$.

(b) *Differential Operators*: Let for example $T = \frac{\mathrm{d}}{\mathrm{d}t} : C^1(I) \subset L^2(I) \to L^2(I)$. $C^1(I)$ is a dense subset of $L^2(I)$.

In this section, we will be dealing with linear operators

$$T : H \supset \operatorname{dom} T \to H$$

where the **domain** $\operatorname{dom} T$ of T is a (in general, non-closed) subspace of T. In fact, in most cases, $\operatorname{dom} T$ is dense.

Definition 2.2. Let $T : H \supset \operatorname{dom} T \to H$ be a linear operator.

(i) T is called **densely defined** if $\overline{\operatorname{dom} T} = H$.

(ii) $G(T) := \{(x, Tx) : x \in \operatorname{dom} T\}$ is called **graph** of T.

(iii) T is called **closed** if $G(T) \subset \operatorname{dom} T \times H$ is closed in $H \times H$.

(iv) T is called **closable** if $\overline{G(T)}$ is the graph of some linear operator $S : H \to H$. In this case, the operator S is called the **closure** of T and is denoted by $\overline{T}$.

(v) We write $T \subset S$ if $G(T) \subset G(S)$ or, equivalently, $\operatorname{dom} T \subset \operatorname{dom} S$ and $Sx = Tx$ for all $x \in \operatorname{dom} T$.

In the next two lemmas, we characterize closedness and closability of linear operators. The first one simply follows from the definition of the norm on $H \times H$.

Lemma 2.3. *Let $T : H \supset \operatorname{dom} T \to H$ be a linear operator. Then the following statements are equivalent.*

(i) *T is closed.*

(ii) *$(x_n) \subset \operatorname{dom} T$, $x_n \to x \in H$, $Tx_n \to y \in H$, implies $x \in \operatorname{dom} T$ and $Tx = y$.*

Lemma 2.4. *Let $T : H \supset \operatorname{dom} T \to H$ be a linear operator. Then the following statements are equivalent.*

(i) T *is closable.*

(ii) T *has a closed linear extension.*

(iii) $(\{0\} \times H) \cap \overline{G(T)} = \{(0,0)\}$.

(iv) $(x_n) \subset \operatorname{dom} T$, $x_n \to 0$, $Tx_n \to y \in H$, *implies* $y = 0$.

Proof. (i)⇒(ii). If T is closable then $G(T) \subset \overline{G(T)} = G(\overline{T})$, and hence $T \subset \overline{T}$ so that $\overline{T}$ is a closed linear extension of T.

(ii)⇒(iii). Let S be a closed linear extension of T. Then it follows from $G(T) \subset G(S)$ that $\overline{G(T)} \subset \overline{G(S)} = G(S)$. Hence, if $y \in H$ such that $(0, y) \in \overline{G(T)}$, then $y = S(0) = 0$, and (iii) is proved.

(iii)⇔(iv). This equivalence is straightforward and follows from the characterization of the closure by sequences.

(iii)⇒(i). Define the set $\mathcal{D} := \{x \in H \,|\, \exists y \in H : (x, y) \in \overline{G(T)}\}$. Let $x \in \mathcal{D}$, $y, z \in H$, and assume that $(x, y), (x, z) \in \overline{G(T)}$. Then also $(0, y - z) \in \overline{G(T)}$, and (iii) implies $y = z$. Hence, the $y \in H$ in the definition of $\mathcal{D}$ is unique, and we define $Sx := y$. Hence, $\operatorname{dom} S := \mathcal{D}$. The linearity of S follows from that of $\overline{G(T)}$, and it is obvious that $G(S) = \overline{G(T)}$. □

Example 2.5. The operator $T : \ell^2 \supset \operatorname{dom} T \to \ell^2$ with $\operatorname{dom} T := \{(x_n) \in \ell^2 : (nx_n) \in \ell^2\}$, defined by $T(x_n) := (nx_n)$, is a closed and densely defined operator.

The inverse mapping theorem generalizes to closed operators.

Theorem 2.6. *Let the linear operator* $T : H \supset \operatorname{dom} T \to H$ *be closed and bijective. Then* $T^{-1} \in L(H)$.

Proof. On $\operatorname{dom} T$, we define the so-called **graph norm** $\|\cdot\|_T$ of T by $\|x\|_T := \|x\| + \|Tx\|$, $x \in \operatorname{dom} T$. It is straightforward to see that the closedness of T implies that $H_T := (\operatorname{dom} T, \|\cdot\|_T)$ is a Banach space. Hence, by the inverse mapping theorem, T^{-1} can be seen as an operator in $L(H, H_T)$. Thus, for $y \in H$, we have $\|T^{-1}y\| \le \|T^{-1}y\| + \|y\| = \|T^{-1}y\|_T \le C\|y\|$ with some $C > 0$. □

Definition 2.7. Let T be densely defined. Then define

$$\operatorname{dom} T^* = \{y \in H \,|\, \exists\, z \in H \,\forall\, x \in H : \langle Tx, y\rangle = \langle x, z\rangle\}\,.$$

The $z \in H$ as in the definition is unique, and we define $T^*y = z$ for all $y \in \operatorname{dom} T^*$.

Note that the adjoint T^* is only defined for densely defined operators. The next lemma is obvious.

Lemma 2.8. *If* T *is densely defined, we have for all* $x \in \operatorname{dom} T$ *and* $y \in \operatorname{dom} T^*$*:*

$$\langle Tx, y\rangle = \langle x, T^*y\rangle\,.$$

Theorem 2.9. *Let* T *be densely defined. Then the following statements hold.*

(i) T^* *is closed.*

(ii) *If* T^* *is densely defined then* T *is closable and* $T^{**} = \overline{T}$.

Proof. (i). Let $(y_n) \subset \operatorname{dom} T^*$ with $y_n \to y$ and $T^* y_n \to z$. For each $x \in \operatorname{dom} T$, we have

$$\langle Tx, y\rangle = \lim_{n\to\infty} \langle Tx, y_n\rangle = \lim_{n\to\infty} \langle x, T^* y_n\rangle = \langle x, z\rangle .$$

This shows $y \in \operatorname{dom} T^*$ and $T^* y = z$, and (i) follows.

(ii). Let $x \in \operatorname{dom} T$. Then for all $y \in \operatorname{dom} T^*$ we have $\langle T^* y, x\rangle = \langle y, Tx\rangle$. This proves $x \in \operatorname{dom} T^{**}$ and $T^{**} x = Tx$. Equivalently, $T \subset T^{**}$. In particular, T is closable. Next, we shall show that $\overline{G(T)} = G(T^{**})$. This then implies $\overline{T} = T^{**}$. Since $T \subset T^{**}$, we have $G(T) \subset G(T^{**})$ and, by (i), $\overline{G(T)} \subset G(T^{**})$. For the converse inclusion, we prove

$$\left(\overline{G(T)}\right)^{\perp} \subset G(T^{**})^{\perp}.$$

Let $(u, v) \in G(T)^{\perp}$. Then $\langle u, x\rangle + \langle v, Tx\rangle = 0$ for all $x \in \operatorname{dom} T$, and therefore,

$$\langle Tx, v\rangle = \langle x, -u\rangle \quad \forall x \in \operatorname{dom} T.$$

Hence, $v \in \operatorname{dom} T^*$. and $T^* v = -u$. For any $(w, T^{**} w) \in G(T^{**})$, we have

$$\langle (w, T^{**} w), (u, v)\rangle_{H\times H} = \langle w, u\rangle + \langle T^{**} w, v\rangle = \langle w, u\rangle + \langle w, T^* v\rangle = 0.$$

This proves the claim. □

Lemma 2.10. *Let* $T : H \supset \operatorname{dom} T \to H$ *be a linear operator. The the following statements hold.*

(i) *If* T *is densely defined, then* $(\operatorname{ran} T)^{\perp} = \ker T^*$.

(ii) *If* T *is closed and there exists* $c > 0$ *such that*

$$\|Tx\| \geq c\|x\| \quad \forall x \in \operatorname{dom} T,$$

then T *is injective and the range of* T *is closed.*

Proof. Exercise. □

Definition 2.11. A linear operator $T : H \supset \operatorname{dom} T \to H$ is called **symmetric** if for all $x, y \in \operatorname{dom} T$ we have $\langle Tx, y\rangle = \langle x, Ty\rangle$.

Note that a symmetric operator T satisfies $\langle Tx, x\rangle \in \mathbb{R}$ for all $x \in \operatorname{dom} T$. It is seen by polarization that this condition is in fact equivalent to the symmetry of T. Moreover, a densely defined operator T is symmetric if and only if $T \subset T^*$.

Lemma 2.12. *Let* T *be a closed and densely defined symmetric operator. Then* $\operatorname{ran}(T-\lambda)$ *is closed and* $\ker(T - \lambda) = \{0\}$ *for all* $\lambda \in \mathbb{C} \setminus \mathbb{R}$.

Proof. Let $\lambda = a + ib$ with $b \neq 0$ and $x \in \operatorname{dom} T$. Then

$$\|(T-\lambda)x\|^2 = \|(T-a)x - ibx\|^2 = \|(T-a)x\|^2 + b^2\|x\|^2 \geq b^2\|x\|^2,$$

since $\langle (T-a)x, ibx\rangle = -ib\langle (T-a)x, x\rangle \in i\mathbb{R}$. □

Definition 2.13. A linear operator $T : H \supset \operatorname{dom} T \to H$ is called **selfadjoint** if it is densely defined and $T = T^*$.

Note that a selfadjoint operator is always densely defined (by definition) and closed, since T^* is always closed by Theorem 2.9.

Clearly, a bounded operator $T \in L(H)$ is symmetric if and only if it is selfadjoint. For unbounded linear operators this is in general not the case.

Theorem 2.14. *Let T be densely defined and symmetric and $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Then the following statements are equivalent:*

(i) *T is selfadjoint.*

(ii) *T is closed and* $\ker(T^* - \lambda) = \ker(T^* - \overline{\lambda}) = \{0\}$.

(iii) $\operatorname{ran}(T - \lambda) = \operatorname{ran}(T - \overline{\lambda}) = H$.

Proof. We prove the theorem for $\lambda = i$. For general $\lambda \in \mathbb{C}\setminus\mathbb{R}$, the reasoning is analogous. (i)⇒(ii) directly follows from Lemma 2.12, and (ii)⇒(iii) is also clear, since $\operatorname{ran}(T \pm i)$ are closed by Lemma 2.12 and $(\operatorname{ran}(T \pm i))^\perp = \ker(T^* \mp i) = \{0\}$. Hence, $\operatorname{ran}(T \pm i) = H$.

(iii)⇒(i). Since T is symmetric, we obtain $T \subset T^*$. Therefore, it is sufficient to prove $\operatorname{dom} T^* \subset \operatorname{dom} T$. Let $y \in \operatorname{dom} T^*$. By (iii), there exists an element $x \in \operatorname{dom} T$ such that $(T^* - i)y = (T - i)x = (T^* - i)x$. Then $y - x \in \ker(T^* - i) = (\operatorname{ran}(T + i))^\perp = \{0\}$. Thus $y = x \in \operatorname{dom} T$. □

The spectrum and the resolvent set of unbounded operators are defined as follows.

Definition 2.15. Let $T : H \supset \operatorname{dom} T \to H$ be a linear operator. We define

$$\varrho(T) := \left\{\lambda \in \mathbb{C} \,\middle|\, T - \lambda : \operatorname{dom} T \to H \text{ is bijective and } (T-\lambda)^{-1} \in L(H)\right\},$$

and $\sigma(T) := \mathbb{C} \setminus \varrho(T)$. These are called the **resolvent set** and the **spectrum** of T, respectively.

Let us collect a few basic facts on the spectrum of a linear operator.

Corollary 2.16. *Let $T : H \supset \operatorname{dom} T \to H$ be a linear operator. Then the following statements hold.*

(i) *If T is not closed then $\sigma(T) = \mathbb{C}$.*

(ii) *If T is closed, then*

$$\varrho(T) := \{\lambda \in \mathbb{C} \mid T - \lambda : \operatorname{dom} T \to H \textit{ is bijective}\}.$$

(iii) $\sigma(T)$ *is closed.*

(iv) *If* T *is selfadjoint then* $\sigma(T) \subset \mathbb{R}$.

Proof. Obviously, T is closed if and only if $T - \lambda$ is if and only if $(T - \lambda)^{-1}$ is. This yields (i). (ii) follows from Theorem 2.6, and (iii) is proved similarly as in the bounded case. (iv) is a consequence of Theorem 2.14 and (ii). □

2.2. Integration with Respect to Spectral Measures

By Σ we denote the Borel sigma algebra on $\mathbb{R}$. Let $T \in L(H)$ be selfadjoint. Below, it will be shown that $E : \Sigma \to L(H)$, defined by $E(\Delta) := \chi_\Delta(T)$, is a spectral measure in virtue of the following definition.

Definition 2.17. A mapping $E : \Sigma \to L(H)$ is called a **spectral measure** if the following conditions are satisfied:

(i) $E(\Delta)$ is an orthogonal projection for each $\Delta \in \Sigma$.

(ii) $E(\varnothing) = 0$ and $E(\mathbb{R}) = I$.

(iii) For a sequence $(\Delta_j) \subset \Sigma$ of mutually disjoint sets, we have for all $x \in H$ that

$$E\left(\bigcup_{j=1}^{\infty} \Delta_j\right) x = \sum_{j=1}^{\infty} E(\Delta_j)x,$$

where the series converges in H.

E is said to have **compact support** if there exists a compact $K \subset \mathbb{R}$ such that $E(K) = I$.

In what follows, we shall often write E_Δ instead of $E(\Delta)$.

Lemma 2.18. *For a spectral measure* E *the following holds:*

(i) $I - E_\Delta = E_{\Delta^c}$ *for all* $\Delta \in \Sigma$.

(ii) *For* $A, B \in \Sigma$ *we have* $E_A E_B = E_B E_A = E_{A\cap B}$.

Proof. Exercise. □

Given a spectral measure E on $\mathbb{R}$, we can for every $x \in H$ define a measure μ_x by

$$\mu_x(\Delta) := \langle E_\Delta x, x\rangle = \|E_\Delta x\|^2, \quad \Delta \in \Sigma.$$

It is evident that μ_x is a positive finite Borel measure for every $x \in H$.

Theorem 2.19. *Let* $T = T^* \in L(H)$*. Then* $E : \Sigma \to L(H)$*, defined by*

$$E(\Delta) := \chi_{\Delta\cap\sigma(T)}(T), \quad \Delta \in \Sigma,$$

is a spectral measure with compact support. For $f \in B(\sigma(T))$ *and* $x \in H$*, we have*

$$\langle f(T)x, x\rangle = \int f \,\mathrm{d}\mu_x.$$

Proof. We write $\sigma := \sigma(T)$. From $\chi_{\Delta\cap\sigma} = \chi^2_{\Delta\cap\sigma} = \overline{\chi_{\Delta\cap\sigma}}$ we immediately conclude that $E_\Delta = E_\Delta^2 = E_\Delta^*$. Hence, E_Δ is an orthogonal projection. We have $E_\varnothing = 0$ and $E_\mathbb{R} = E_\sigma = \mathbb{1}(T) = I$. Let $(\Delta_j) \subset \Sigma$ be a sequence of mutually disjoint sets. Define

$$f_n := \sum_{j=1}^n \chi_{\Delta_j\cap\sigma} = \chi_{\bigcup_{j=1}^n(\Delta_j\cap\sigma)} \qquad \text{and} \qquad f := \chi_{\bigcup_{j=1}^\infty(\Delta_j\cap\sigma)}.$$

Hence $\|f_n\|_\infty = 1$ for all $n \in \mathbb{N}$ and $f_n \to f$ pointwise. By the properties of the measurable functional calculus, $f_n(T)x \to f(T)x$ for each $x \in H$, and hence

$$\sum_{j=1}^n E_{\Delta_j} x \overset{n\to\infty}{\longrightarrow} E_{\bigcup_{j=1}^\infty \Delta_j} x.$$

This shows that E is a spectral measure with compact support. The integral identity follows from the proof of Theorem 1.15 (see (1.7)). □

For the rest of this section, let E be a spectral measure. A function $f : \mathbb{R} \to \mathbb{C}$ is called **simple** if it has the form

$$f = \sum_{i=1}^n \alpha_i \chi_{\Delta_i},$$

where $\alpha_i \in \mathbb{C}$ and the $\Delta_i \in \Sigma$ are mutually disjoint. Define

$$\int f \,\mathrm{d}E := \sum_{i=1}^n \alpha_i E_{\Delta_i} \in L(H).$$

It is easily seen that this definition does not depend on the representation of f. For $x \in H$, we have

$$\begin{aligned}\left\|\left(\int f\,\mathrm{d}E\right)x\right\|^2 &= \left\|\sum_{i=1}^n \alpha_i E_{\Delta_i} x\right\|^2 = \sum_{i=1}^n |\alpha_i|^2 \|E_{\Delta_i}x\|^2 \\ &= \sum_{i=1}^n |\alpha_i|^2 \mu_x(\Delta_i) = \int |f|^2 \,\mathrm{d}\mu_x.\end{aligned} \tag{2.1}$$

By $\mathcal{E}$ denote the set of all simple functions on $\mathbb{R}$. For a measureable function $f : \mathbb{R} \to \mathbb{C}$, we set $D_f := \{x \in H : f \in L^2(\mu_x)\}$. Moreover, for $x \in H$, we put $L^2_x := L^2(\mu_x)$.

Lemma 2.20. *If $f : \mathbb{R} \to \mathbb{C}$ is measurable, then there exists a sequence $(f_n) \subset \mathcal{E}$ with $f_n \to f$ and $|f_n| \nearrow |f|$ pointwise. If, in addition, $f \in L^2_x$ for some $x \in H$, then we have $f_n \to f$ in L^2_x.*

Proof. It suffices to prove the first claim of the lemma for $\mathbb{R}$-valued f. But in this case,

$$f_n(t) := \begin{cases} n, & \text{if } f(t) \geq n, \\ k2^{-n}, & \text{if } k2^{-n} \leq f(t) < (k+1)2^{-n},\ k = 0, \dots, n2^n - 1, \\ -k2^{-n}, & \text{if } -(k+1)2^{-n} < f(t) \leq -k2^{-n},\ k = 0, \dots, n2^n - 1, \\ -n, & \text{if } f(t) \leq -n. \end{cases}$$

is as desired.

For the second claim, observe that $|f_n - f|^2 \to 0$ (pointwise) and $|f_n - f|^2 \leq (|f_n| + |f|)^2 \leq 4|f|^2 \in L^1(\mu_x)$. By Lebesgue's Theorem, it follows that $f_n \to f$ in L^2_x. $\square$

Let $f : \mathbb{R} \to \mathbb{C}$ be measurable, $x \in D_f$. Choose $(f_n) \subset \mathcal{E}$ such that $f_n \to f$ in L^2_x. Then, due to (2.1), we have

$$\left\| \int f_n \,\mathrm{d}Ex - \int f_m \,\mathrm{d}Ex \right\|^2 = \left\| \int (f_n - f_m) \,\mathrm{d}Ex \right\|^2 = \int |f_n - f_m|^2 \,\mathrm{d}\mu_x \to 0. \tag{2.2}$$

Therefore, we may define

$$\int f \,\mathrm{d}Ex := \lim_{n\to\infty} \int f_n \,\mathrm{d}Ex.$$

If $(f'_n) \subset \mathcal{E}$ is another sequence with $f'_n \to f$ in L^2_x, then (2.2) with f_m replaced by f'_n shows that $\int f \,\mathrm{d}Ex$ is well-defined.

Let $x \in H$. Due to the above definition, for $f, g \in L^2_x$ and $\alpha, \beta \in \mathbb{C}$, we have

$$\int (\alpha f + \beta g) \,\mathrm{d}Ex = \alpha \int f \,\mathrm{d}Ex + \beta \int g \,\mathrm{d}Ex.$$

Moreover,

$$\left\| \int f \,\mathrm{d}Ex \right\|^2 = \int |f|^2 \,\mathrm{d}\mu_x.$$

For measurable $f : \mathbb{R} \to \mathbb{C}$, we now define an operator f_E by $\operatorname{dom} f_E := D_f$ and

$$f_E x := \int f \,\mathrm{d}Ex, \quad x \in D_f.$$

Theorem 2.21. *Let E be a spectral measure and $f, g : \mathbb{R} \to \mathbb{C}$ measurable functions. Then the following statements hold.*

(i) *f_E is a closed and densely defined linear operator.*

(ii) *f_E is normal, i.e.* $\operatorname{dom} f_E = \operatorname{dom} f_E^*$.

(iii) $(\overline{f})_E = f_E^*$.

(iv) *If f is bounded, $f_E \in L(H)$ with $\|f_E\| \leq \|f\|_\infty$.*

(v) *$f_E + g_E \subset (f+g)_E$ and* $\operatorname{dom}(f_E + g_E) = D_{|f|+|g|}$.

(vi) *$f_E g_E \subset (fg)_E$ and* $\operatorname{dom}(f_E g_E) = D_g \cap D_{fg}$.

Proof. First of all, we prove that f_E is a linear operator and D_f is a vector space. For this, let $\lambda \in \mathbb{C}$ and $x \in H$. Then

$$\mu_{\lambda x}(\Delta) = \langle E_\Delta(\lambda x), \lambda x\rangle = |\lambda|^2 \mu_x(\Delta).$$

Hence, if $x \in D_f$ then

$$\int |f|^2 \, \mathrm{d}\mu_{\lambda x} = |\lambda|^2 \int |f|^2 \, \mathrm{d}\mu_x < \infty,$$

which implies $\lambda x \in D_f$. Choose a sequence of simple functions (f_n) with $f_n \to f$ and $|f_n| \nearrow |f|$. Then

$$f_E(\lambda x) = \int f \, \mathrm{d}E(\lambda x) = \lim_{n\to\infty} \int f_n \, \mathrm{d}E(\lambda x) = \lambda \lim_{n\to\infty} \int f_n \, \mathrm{d}Ex = \lambda \int f \, \mathrm{d}Ex = \lambda f_E x.$$

Let $x, y \in D_f$, i.e. $f \in L^2_x \cap L^2_y$ (where $L^2_x := L^2(\mu_x)$). Then

$$\begin{aligned}
\left(\int |f_n|^2 \, \mathrm{d}\mu_{x+y} \right)^{\frac{1}{2}} &= \|(f_n)_E(x+y)\| = \|(f_n)_E x + (f_n)_E y\| \\
&\le \|(f_n)_E x\| + \|(f_n)_E y\| = \left(\int |f_n|^2 \, \mathrm{d}\mu_x \right)^{\frac{1}{2}} + \left(\int |f_n|^2 \, \mathrm{d}\mu_y \right)^{\frac{1}{2}} \\
&\le \left(\int |f|^2 \, \mathrm{d}\mu_x \right)^{\frac{1}{2}} + \left(\int |f|^2 \, \mathrm{d}\mu_y \right)^{\frac{1}{2}} = \|f_E x\| + \|f_E y\| < \infty
\end{aligned}$$

By Beppo Levi's Theorem, we obtain

$$\left(\int |f|^2 \, \mathrm{d}\mu_{x+y} \right)^{\frac{1}{2}} = \left(\lim_{n\to\infty} \int |f_n|^2 \, \mathrm{d}\mu_{x+y} \right)^{\frac{1}{2}} \le \|f_E x\| + \|f_E y\| < \infty,$$

and therefore $x + y \in D_f$. Furthermore,

$$f_E(x+y) = \lim_{n\to\infty} (f_n)_E(x+y) = \lim_{n\to\infty} (f_n)_E x + (f_n)_E y = f_E x + f_E y.$$

The statement (iv) is evident. We will now prove (v) and (vi). For (v), let $x \in \mathrm{dom}(f_E + g_E) = D_f \cap D_g$. Then $f, g \in L^2_x$. Thus, $f + g \in L^2_x$, i.e., $x \in D_{f+g} = \mathrm{dom}((f+g)_E)$, and

$$(f+g)_E x = \int (f+g) \, \mathrm{d}Ex = \int f \, \mathrm{d}Ex + \int g \, \mathrm{d}Ex = f_E x + g_E x.$$

This implies $f_E + g_E \subset (f+g)_E$. The remainder follows from

$$x \in \mathrm{dom}(f_E + g_E) \iff f, g \in L^2_x \iff |f| + |g| \in L^2_x \iff x \in D_{|f|+|g|}.$$

For the proof of (vi), the following principles will be useful:

(P1) $f_n \to f$ in $L^2_x \implies (f_n)_E x \to f_E x$.

(P2) $f_n \in L^2_x$, $f_n \to f$, $|f_n| \nearrow |f|$, and $(f_n)_E x$ converges $\Longrightarrow x \in D_f$ and $(f_n)_E x \to f_E x$.

The principle (P1) is easy to prove and will be left to the reader. As to (P2), Beppo Levi's Theorem yields

$$\int |f|^2 \,\mathrm{d}\mu_x = \lim_{n\to\infty} \int |f_n|^2 \,\mathrm{d}\mu_x = \lim_{n\to\infty} \|(f_n)_E x\|^2 < \infty,$$

which implies $f \in L^2_x$, and from Lebesgue's Theorem we get $f_n \to f$ in L^2_x. The claim now follows from (P1).

We first prove $f_E g_E \subset (fg)_E$ when $f \in \mathcal{E}$. In this case, $f_E \in L(H)$. For $x \in \operatorname{dom}(f_E g_E) = \operatorname{dom}(g_E) = D_g$, we choose $(g_n) \subset \mathcal{E}$ with $g_n \to g$, $|g_n| \nearrow |g|$, and obtain

$$f_E g_E x = f_E \left(\lim_{n\to\infty} (g_n)_E x \right) = \lim_{n\to\infty} f_E (g_n)_E x = \lim_{n\to\infty} (f g_n)_E x \overset{(P2)}{=} (fg)_E x.$$

This shows $f_E g_E \subset (fg)_E$.

For general f, let $x \in \operatorname{dom}(f_E g_E)$, i.e. $x \in D_g$ and $g_E x \in D_f$. Let $(f_n) \subset \mathcal{E}$ with $f_n \to f$ and $|f_n| \nearrow |f|$. Then

$$f_E g_E x = \lim_{n\to\infty} (f_n)_E g_E x = \lim_{n\to\infty} (f_n g)_E x \overset{(P2)}{=} (fg)_E x,$$

and, in particular, $x \in D_{fg}$. This implies $f_E g_E \subset (fg)_E$. In particular, $\operatorname{dom}(f_E g_E) \subset D_g \cap D_{fg}$. For $x \in D_g \cap D_{fg}$ we have

$$(f_n)_E g_E x = (f_n g)_E x \overset{(P1)}{\longrightarrow} (fg)_E x$$

since we have $f_n g \overset{L^2_x}{\longrightarrow} fg$ by Lebesgue's Theorem. By (P2), $g_E x \in D_f$.

We will now show that D_f is dense in H. For this, set $\chi_n := \chi_{|f|\le n}$, and observe that

$$\operatorname{dom}(f_E (\chi_n)_E) = D_{\chi_n} \cap D_{f\chi_n} = H.$$

Therefore, for all $x \in H$ we have $(\chi_n)_E x \in D_f$. But as $\chi_n \to \mathbb{1}$ in L^2_x we obain from (P1) that $x = \lim_{n\to\infty} (\chi_n)_E x$.

(iii). This is true for $f \in \mathcal{E}$! Let f be arbitrary and $(f_n) \subset \mathcal{E}$ with $f_n \to f$ and $|f_n| \nearrow |f|$. Let $y \in D_{\overline{f}} = D_f$. Then for all $x \in D_f$ we have

$$\langle f_E x, y \rangle = \lim_{n\to\infty} \langle (f_n)_E x, y \rangle = \lim_{n\to\infty} \left\langle x, \left(\overline{f_n}\right)_E y \right\rangle = \left\langle x, \left(\overline{f}\right)_E y \right\rangle.$$

This shows that $y \in \operatorname{dom}(f_E^*)$ and $f_E^* y = \left(\overline{f}\right)_E y$ and hence $\left(\overline{f}\right)_E \subset f_E^*$. In particular, (iii) holds for bounded f. Therefore, as $f_E (\chi_n)_E = (f\chi_n)_E$ is bounded, we have $(f_E(\chi_n)_E)^* = (f\chi_n)_E^* = \left(\overline{f}\chi_n\right)_E$. For $y \in \operatorname{dom} f_E^*$ and $x \in H$, we therefore have

$$\langle (\chi_n)_E f_E^* y, x \rangle = \langle f_E^* y, (\chi_n)_E x \rangle = \langle y, f_E (\chi_n)_E x \rangle = \left\langle \left(\overline{f}\chi_n\right)_E y, x \right\rangle.$$

Therefore $(\chi_n)_E f_E^* y = \left(\overline{f}\chi_n\right)_E y$. Letting $n \to \infty$, we get $f_E^* y = \lim_{n\to\infty} (\overline{f}\chi_n)_E y$. By (P2), $y \in D_f$.

(ii). This now easily follows from what we have already proved:

$$\operatorname{dom} f_E^* = \operatorname{dom} \left(\overline{f}\right)_E = D_{\overline{f}} = D_f = \operatorname{dom} f_E,$$

and for $x \in D_f$ we have

$$\|f_E^* x\|^2 = \left\|\left(\overline{f}\right)_E x\right\|^2 = \int |f|^2 \,\mathrm{d}\mu_x = \|f_E x\|^2.$$

(i). It remains to show that f_E is closed. We prove that, if T is densely defined and normal, then T is closed. Let $(x_n) \subset \operatorname{dom} T$ with $x_n \to x$ and $Tx_n \to y$. Then

$$\|T^*(x_n - x_m)\| = \|T(x_n - x_m)\| \overset{n,m\to\infty}{\longrightarrow} 0.$$

Hence, there exists $z \in H$ such that $T^* x_n \to z$. Since T^* is closed, we know that $x \in \operatorname{dom} T^* = \operatorname{dom} T$ and $T^* x = z$. Moreover,

$$\|T(x_n - x)\| = \|T^*(x_n - x)\| \to 0,$$

and thus $Tx = \lim_{n\to\infty} Tx_n = y$. □

2.3. The Spectral Theorem

In this section, it is our aim to prove the following theorem.

Theorem 2.22 (The Spectral Theorem). *Let $T = T^*$ be a selfadjoint operator. Then there exists a unique spectral measure E such that $T = \mathrm{id}_E$, i.e.,*

$$T = \int t \,\mathrm{d}E.$$

The proof of Theorem 2.22 will be partitioned into three steps:

1. Observe that the so-called **Cayley transform**

$$U_T := (T - i)(T + i)^{-1}$$

   of T is unitary and prove that $-U_T = e^{iA}$, where $A = A^* \in L(H)$ with $\sigma(A) \subset [-\pi, \pi]$.

2. Use the spectral measure of the bounded operator A to define E, and prove that $T = \mathrm{id}_E$.

3. Prove Stieltjes' Inversion Formula and use it to prove uniqueness.

Note that the Cayley transform U_T is well-defined by Corollary 2.16(iv). Moreover, $U_T \in L(H)$. Indeed, we have

$$U_T = ((T + i) - 2i)(T + i)^{-1} = I - 2i(T + i)^{-1}.$$

Now, it is an easy task to prove the following lemma.

Lemma 2.23. *The* Caley *transform U_T of T is unitary i.e. $U_T^* U_T = U_T U_T^* = I$. Moreover, 1 is not an eigenvalue of U_T.*

Lemma 2.24. *Let $W, T \in L(H)$ be selfadjoint, such that $WT = TW$ and $W^2 = T^2$. Then the orthogonal projection P onto $\ker(W - T)$ has the following properties:*

(i) *If $L = L^* \in L(H)$ and $L(W - T) = (W - T)L$, then $LP = PL$.*

(ii) *$Wx = 0$ implies $Px = x$.*

(iii) *$W = (2P - I)T$.*

Proof. (i). $L(W-T) = (W-T)L$ shows that $\ker(W-T)$ and $\operatorname{ran}(W-T)$ are L-invariant and therefore also $\overline{\operatorname{ran}(W - T)} = \ker(W - T)^\perp$. Let $x \in H$, $x = u + v$, $u \in \ker(W - T)$, $v \in \ker(W - T)^\perp$. Then

$$PLx = P(Lu + Lv) = PLu = Lu = LPx.$$

(ii). We have $\|Wx\|^2 = \langle Wx, Wx \rangle = \langle W^2 x, x \rangle = \langle T^2 x, x \rangle = \|Tx\|^2$. Hence, $Tx = 0$, which implies $(W - T)x = 0$. Thus, $x \in \ker(W - T)$ and $Px = x$.

(iii). We have $(W - T)(W + T) = 0$. So, $\operatorname{ran}(W + T) \subset \ker(W - T)$. Hence, $P(W + T) = W + T$. Moreover, $(W - T)P = 0$. As P commutes with W and T, $W + T = P(W + T) = (W + T)P = (W - T + 2T)P = 2TP$. This shows $W = 2TP - T = (2P - I)T$. □

The next theorem is key to the construction of the spectral measure in the spectral theorem.

Theorem 2.25. *Let $U \in L(H)$ be unitary. Then there exists $A = A^* \in L(H)$ with $\sigma(A) \subset [-\pi, \pi]$ such that $U = e^{iA}$. If $-\pi$ or π is an eigenvalue of A, then -1 is an eigenvalue of U.*

Proof. Define $V := \frac{U+U^*}{2}$ and $W := \frac{U-U^*}{2}$. Then $U = V + iW$, $V = V^*$ and $W = W^*$. The unitarity of U implies that V and W commute and $V^2 + W^2 = I$. The latter implies $\|V\| \le 1$ and $\|W\| \le 1$. In particular, $\sigma(V) \subset [-1, 1]$. Let $f(t) := \sqrt{1 - t^2}$, $t \in [-1, 1]$. Then $f \in C(\sigma(V))$, and $T := f(V) = \Phi_V(f)$ (Φ_V denoting the continuous functional calculus of V) is defined. By Theorem 1.10, T commutes with V and W and is selfadjoint. Moreover,

$$T^2 = \Phi_V(f^2) = \Phi_V(\mathbb{1} - \mathrm{id}^2) = I - V^2 = W^2.$$

By Lemma 2.24, we get $W = (2P - I)T$ and $Px = x$ for $x \in \ker W$, where P is the orthogonal projection onto $\ker(W - T)$. Moreover $PV = VP$.

Now, define the operator

$$A := (2P - I)\arccos(V).$$

As $2P - I$ is unitary and $(2P - I)^2 = I$, we have

$$\|A\| = \|\arccos(V)\| = \|\arccos\|_\infty = \pi \qquad \text{and} \qquad A^2 = \arccos^2(V).$$

Now, we have that $\cos(z) = h_1(z^2)$ and $\sin(z) = z \cdot h_2(z^2)$ with entire functions h_1 and h_2. Therefore,

$$\cos(A) = h_1(A^2) = h_1(\arccos^2(V)) = \cos(\arccos(V)) = V,$$

and

$$\sin(A) = (2P - I)\arccos(V) h_2(\arccos^2(V)) = (2P - I)\sin(\arccos(V)).$$

Now, $\sin(\arccos(t)) = \sqrt{1-t^2} = f(t)$. Therefore, $\sin(A) = (2P - I)T = W$. Hence, we obtain $U = V + iW = \cos(A) + i\sin(A) = e^{iA}$.

The statement on the eigenvalues follows directly from Theorem 1.10(iii). □

For proving the uniqueness of the spectral measure in Theorem 1.13, we will make use of Stieltjes' Inversion Formula.

Lemma 2.26 (Stieltjes' Inversion Formula). *Let μ be a finite Borel measure on $\mathbb{R}$. Then*

$$\mu((-\infty, t]) = \lim_{\delta \searrow 0} \lim_{\varepsilon \searrow 0} \frac{1}{\pi} \int_{-\infty}^{t+\delta} \operatorname{Im} f(s + i\varepsilon)\, \mathrm{d}s,$$

where $f(z) := \int_{\mathbb{R}} \frac{1}{t-z}\, \mathrm{d}\mu(t)$, $z \in \mathbb{C} \setminus \mathbb{R}$.

Proof. For $\varepsilon > 0$ we have

$$\operatorname{Im} f(s + i\varepsilon) = \int_{\mathbb{R}} \operatorname{Im} \frac{1}{t - s - i\varepsilon}\, \mathrm{d}\mu(t) = \int_{\mathbb{R}} \frac{\varepsilon}{(t-s)^2 + \varepsilon^2}\, \mathrm{d}\mu(t).$$

By Fubini's theorem, we get for $r \in \mathbb{R}$:

$$\int_{-\infty}^{r} \operatorname{Im} f(s + i\varepsilon)\, \mathrm{d}s = \int_{\mathbb{R}} \int_{-\infty}^{r} \frac{\varepsilon}{(t-s)^2 + \varepsilon^2}\, \mathrm{d}s\, \mathrm{d}\mu(t) = \int_{\mathbb{R}} \left[\arctan\left(\frac{r-t}{\varepsilon}\right) + \frac{\pi}{2}\right] \mathrm{d}\mu(t).$$

We define $g_{\varepsilon,r}(t) := \arctan\left(\frac{r-t}{\varepsilon}\right) + \frac{\pi}{2}$ and note $|g_{\varepsilon,r}(t)| \leq \pi$. Then,

$$\lim_{\varepsilon \searrow 0} g_{\varepsilon,r}(t) = \begin{cases} \pi & \text{for } t < r \\ \frac{\pi}{2} & \text{for } t = r \\ 0 & \text{for } t > r. \end{cases}$$

Applying Lebesgue's theorem gives

$$\begin{aligned} \lim_{\varepsilon \searrow 0} \int_{-\infty}^{r} \operatorname{Im} f(s + i\varepsilon)\, \mathrm{d}s &= \pi\mu((-\infty, r)) + \frac{\pi}{2}\mu(\{r\}) \\ &= \pi\mu((-\infty, r)) + \frac{\pi}{2}\big[\mu((-\infty, r]) - \mu((-\infty, r))\big] \\ &= \frac{\pi}{2}\big[\mu((-\infty, r]) + \mu((-\infty, r))\big]. \end{aligned}$$

Putting $r = t + \delta$ and letting $\delta \searrow 0$ yields the desired result. □

Proof of Theorem 2.22. Let $A \in L(H)$ be selfadjoint with $\sigma(A) \subset [-\pi, \pi]$, such that $-U_T = e^{iA}$. By F denote the spectral measure of A as in Theorem 2.19, i.e., $F_\Delta = \chi_{\Delta\cap\sigma(A)}(A)$ for $\Delta \in \Sigma$. By Theorem 2.19 we get

$$\langle U_T x, x\rangle = -\int e^{it}\,\mathrm{d}\nu_x, \qquad x \in H,$$

where the measure ν_x is defined by $\nu_x(\Delta) := \langle F_\Delta x, x\rangle$, $\Delta \in \Sigma$. As shown above, $U_T = I - 2i(T+i)^{-1}$, which gives $(T+i)^{-1} = \frac{1}{2i}(I - U_T)$. Thus,

$$\left\langle (T+i)^{-1}x, x\right\rangle = \frac{1}{2i}\langle x - U_T x, x\rangle = \frac{1}{2i}\left(\|x\|^2 + \int e^{it}\,\mathrm{d}\nu_x\right) = \frac{1}{2i}\int_{(-\pi,\pi)} \left(1 + e^{it}\right)\mathrm{d}\nu_x.$$

Now, we define the spectral measure E by $E_\Delta := F_{2\arctan(\Delta)}$, $\Delta \in \Sigma$. As usual, for $x \in H$, we set $\mu_x(\Delta) := \langle E_\Delta x, x\rangle$, $\Delta \in \Sigma$. Then

$$\left\langle (T+i)^{-1}x, x\right\rangle = \frac{1}{2i}\int_{\mathbb{R}} \left(1 + e^{2i\arctan(t)}\right)\mathrm{d}\mu_x.$$

Taking into consideration that

$$\cos(\arctan(s)) = \frac{1}{\sqrt{1+s^2}}, \qquad \sin(\arctan(s)) = \frac{s}{\sqrt{1+s^2}},$$

and

$$\cos(2s) = \cos^2(s) - \sin^2(s), \qquad \sin(2s) = 2\sin(s)\cos(s),$$

we conclude

$$1 + e^{2i\arctan(s)} = 2\frac{1+is}{1+s^2}.$$

Therefore,

$$\left\langle (T+i)^{-1}x, x\right\rangle = \frac{1}{i}\int_{\mathbb{R}} \frac{1+is}{1+s^2}\,\mathrm{d}\mu_x = \int_{\mathbb{R}} \frac{s-i}{(s-i)(s+i)}\,\mathrm{d}\mu_x = \int_{\mathbb{R}} \frac{1}{s+i}\,\mathrm{d}\mu_x.$$

Now, let $S := \mathrm{id}_E$, $z \in \mathbb{C}\setminus\mathbb{R}$ (and thus $z \in \varrho(S)$ by Corollary 2.16(iv)), and define the bounded and measurable function $f : \mathbb{R} \to \mathbb{C}$ by $f(t) := \frac{1}{t-z}$. Then

$$f_E(S - z) = f_E(\mathrm{id} - z)_E \subset (f(\mathrm{id} - z))_E = \mathbb{1}_E = I.$$

This implies $f_E(S-z)x = x$ for all $x \in \operatorname{dom} S$. Thus, $f_E = (S-z)^{-1}$, which implies

$$\left\langle (S-z)^{-1}x, x\right\rangle = \int \frac{1}{t-z}\,\mathrm{d}\mu_x, \quad x \in H.$$

In particular, this shows $(S+i)^{-1} = (T+i)^{-1}$ and thus also $T = S = \mathrm{id}_E$.

It remains to prove the uniqueness of E. For this, let G be a spectral measure with $T = \mathrm{id}_G$, generating the measure $\gamma_x(\Delta) := \langle G_\Delta x, x\rangle$, $\Delta \in \Sigma$. Then, for $z \in \mathbb{C}\setminus\mathbb{R}$ we have

$$\left\langle (T-z)^{-1}x, x\right\rangle = \int_{\mathbb{R}} \frac{1}{t-z}\,\mathrm{d}\mu_x = \int_{\mathbb{R}} \frac{1}{t-z}\,\mathrm{d}\gamma_x, \quad x \in H.$$

Now, Lemma 2.26 implies that for $x \in H$ and $t \in \mathbb{R}$ the following holds:

$$\mu_x((-\infty, t]) = \lim_{\delta \searrow 0} \lim_{\varepsilon \searrow 0} \frac{1}{\pi} \int_{-\infty}^{t+\delta} \operatorname{Im} \left\langle (T - s - i\varepsilon)^{-1} x, x \right\rangle \mathrm{d}s = \gamma_x((-\infty, t]).$$

Thus, for all $x \in H$ we have $\gamma_x = \mu_x$, implying $G = E$. $\square$

2.4. The Spectrum of a Selfadjoint Operator

In Corollary 2.16(iv), we observed that the spectrum of a general selfadjoint operator is always real. In the finite-dimensional situation, the spectral values (in this case, eigenvalues) on the real line are related to eigenvectors and therefore to "parts" of the Hilbert space. A similar relation can be observed in the general setting. The connection is established by the spectral measure.

Definition 2.27. Let T be a selfadjoint operator and E its spectral measure, i.e., $T = \mathrm{id}_E$. For a measurable function $f : \mathbb{R} \to \mathbb{C}$, we define $f(T) := f_E$.

Theorem 2.28. *Let T be a selfadjoint operator, E its spectral measure, $f : \mathbb{R} \to \mathbb{C}$ a measurable function, and $\lambda \in \mathbb{C}$. Then the following statements hold.*

(i) $E_{\varrho(T) \cap \mathbb{R}} = 0$.

(ii) $\lambda \in \varrho(T)$ *if and only if* $E_U = 0$ *for some neighborhood* U *of* λ.

(iii) $Tx = \lambda x$ *implies* $x \in \operatorname{dom} f(T)$ *and* $f(T)x = f(\lambda)x$.

(iv) $\ker(T - \lambda) = E_{\{\lambda\}} H$.

(v) $\sigma(f(T)) \subset \overline{f(\sigma(T))}$.

Proof. (i). Let $[a, b] \subset \varrho(T)$ be a compact interval. Then, since $\lambda \mapsto (T - \lambda)^{-1}$ is continuous (even analytic) on a complex neighborhood of $[a, b]$, for $x \in H$ we have

$$\begin{aligned}
\mu_x((a, b]) &= \mu_x((-\infty, b]) - \mu_x((-\infty, a]) \\
&= \frac{1}{\pi} \lim_{\delta \searrow 0} \lim_{\varepsilon \searrow 0} \int_{a+\delta}^{b+\delta} \operatorname{Im} \left\langle (T - s - i\varepsilon)^{-1} x, x \right\rangle \mathrm{d}s \\
&= \frac{1}{\pi} \lim_{\delta \searrow 0} \int_{a+\delta}^{b+\delta} \underbrace{\operatorname{Im} \left\langle (T - s)^{-1} x, x \right\rangle}_{=0} \mathrm{d}s = 0.
\end{aligned}$$

Hence, for all $x \in H$ and all compact intervalls $\Delta \subset \varrho(T)$ we have $\mu_x(\Delta) = 0$. As measures are continuous from below[3], it follows that $\mu_x(J) = 0$ for each connected component J of $\varrho(T) \cap \mathbb{R}$. Since there are at most countably many of those, we obtain $\mu_x(\varrho(T) \cap \mathbb{R}) = 0$, which implies (i).

[3]Meaning that $A_n \nearrow A$ implies $\mu(A_n) \to \mu(A)$.

(ii). If $\lambda \in \varrho(T)$, then a neighborhood U as in (ii) exists due to (i) and the openness of $\varrho(T)$. Conversely, let U be a neighborhood of λ such that $E_U = 0$. Define a measurable function $f : \mathbb{R} \to \mathbb{C}$ by $f(t) := (t-\lambda)^{-1}$ if $t \notin U$ and $f(t) = 0$ for $t \in U$. Moreover, put $g(t) := t - \lambda$, $t \in \mathbb{R}$. Then $fg = \chi_{U^c}$, and f is bounded. Thus, $f_E g_E \subset (fg)_E = E_{U^c} = I$, $g_E f_E \subset I$, and $\operatorname{dom}(g_E f_E) = D_f \cap D_{gf} = H$. Consequently, $g_E f_E = I$. Hence, $g_E = T - \lambda$ is bijective, i.e., $\lambda \in \rho(T)$.

(iii). Let $Tx = \lambda x$ for some $x \in \operatorname{dom} T$. Assume that for all $\Delta \in \Sigma$

$$E_\Delta x = \chi_\Delta(\lambda) x. \tag{2.3}$$

Then the claim holds for $f \in \mathcal{E}$: If $f = \sum \alpha_i \chi_{\Delta_i}$, then

$$f(T)x = \sum \alpha_i E_{\Delta_i} x = \sum \alpha_i \chi_{\Delta_i}(\lambda) x = f(\lambda) x.$$

Moreover,

$$\mu_x(\Delta) = \langle E_\Delta x, x \rangle = \chi_\Delta(\lambda) \|x\|^2, \quad \Delta \in \Sigma.$$

Therefore, for each measurable f,

$$\int |f|^2 \, \mathrm{d}\mu_x = |f(\lambda)|^2 \|x\|^2 < \infty,$$

which implies $x \in D_f$ (or $f \in L^2_x$). Choose $(f_n) \subset \mathcal{E}$, $f_n \to f$ both pointwise and in L^2_x. Then

$$f(T)x = f_E x = \lim_{n\to\infty} (f_n)_E x = \lim_{n\to\infty} f_n(\lambda) x = f(\lambda) x.$$

So we only have to prove (2.3). For this, observe

$$\begin{aligned}
\left\langle E_{(-\infty,c]} x, x \right\rangle &= \lim_{\delta \searrow 0} \lim_{\epsilon \searrow 0} \frac{1}{\pi} \int_{-\infty}^{c+\delta} \operatorname{Im} \left\langle (T - s - i\varepsilon)^{-1} x, x \right\rangle \mathrm{d}s \\
&= \left(\lim_{\delta \searrow 0} \lim_{\epsilon \searrow 0} \frac{1}{\pi} \int_{-\infty}^{c+\delta} \operatorname{Im} \frac{1}{\lambda - s - i\varepsilon} \, \mathrm{d}s \right) \|x\|^2 \\
&= \lim_{\delta \searrow 0} \lim_{\epsilon \searrow 0} \frac{1}{\pi} \left(\arctan \left(\frac{c + \delta - \lambda}{\varepsilon} \right) + \frac{\pi}{2} \right) \|x\|^2 \\
&= \|x\|^2 \lim_{\delta \searrow 0} \begin{cases} 1, & \text{if } c - \lambda + \delta > 0, \\ \frac{1}{2}, & \text{if } c - \lambda + \delta = 0, \\ 0, & \text{if } c - \lambda + \delta < 0 \end{cases} \\
&= \|x\|^2 \begin{cases} 1, & \text{if } c \geq \lambda, \\ 0, & \text{if } c < \lambda \end{cases} = \chi_{(-\infty,c]}(\lambda) \|x\|^2.
\end{aligned}$$

Define $\nu_x := \chi_\Delta(\lambda)\|x\|^2$, $\Delta \in \Sigma$. Then $\mu_x = \nu_x$ on a generator of Σ, and thus also on all of Σ. This gives

$$\langle E_\Delta x, x \rangle = \chi_\Delta(\lambda) \|x\|^2, \quad \Delta \in \Sigma.$$

Note that for an orthogonal projection P the equality $\langle Px, x \rangle = 0$ implies $\|Px\|^2 = \langle Px, Px \rangle = \left\langle P^2 x, x \right\rangle = \langle Px, x \rangle = 0$ and hence $Px = 0$.

Let $\Delta \in \Sigma$. If $\lambda \notin \Delta$, we have $\langle E_\Delta x, x\rangle = 0$ and thus $E_\Delta x = 0$. If $\lambda \in \Delta$, then $\langle (I - E_\Delta)x, x\rangle = 0$ and thus $(I - E_\Delta)x = 0$, i.e., $\mathbb{E}_\Delta x = x$.

(iv). We have

$$(T-\lambda)E_{\{\lambda\}} = (\mathrm{id}-\lambda)_E(\chi_{\{\lambda\}})_E \subset \big((\mathrm{id}-\lambda)\cdot\chi_{\{\lambda\}}\big)_E = 0_E = 0$$

and

$$\mathrm{dom}\,\big((T-\lambda)E_{\{\lambda\}}\big) = \mathrm{dom}\,\Big((\mathrm{id}-\lambda)_E\,\big(\chi_{\{\lambda\}}\big)_E\Big) = D_{\chi_{\{\lambda\}}} \cap D_0 = H.$$

This shows $(T-\lambda)E_{\{\lambda\}} = 0$ and hence $E_{\{\lambda\}}H \subset \ker(T-\lambda)$. If $Tx = \lambda x$ with $x \in \mathrm{dom}\,T$, then $E_{\{\lambda\}}x = \chi_{\{\lambda\}}(T)x = \chi_{\{\lambda\}}(\lambda)x = x$. This shows $x \in E_{\{\lambda\}}H$.

(v). Let $\mu \notin \overline{f(\sigma(T))}$, and put $d := \mathrm{dist}\,\big(\mu, \overline{f(\sigma(T))}\big)$. Put $g(\lambda) := (\mu - f(\lambda))^{-1}$ for $\lambda \in \sigma(T)$, and $g(\lambda) := 0$ for $\lambda \in \varrho(T)$. This gives $|g(\lambda)| \le \frac{1}{d}$ for all $\lambda \in \mathbb{R}$. Moreover, set $h(\lambda) := \mu - f(\lambda)$. Then, since $hg = \chi_{\sigma(T)}$, we conclude $h_E g_E \subset (hg)_E = E_{\sigma(T)} = I$ and $\mathrm{dom}(h_E g_E) = D_g \cap D_{hg} = H$, i.e., $h_E g_E = I$. Since also $g_E h_E \subset (gh)_E = I$, we obtain that $h_E = \mu - f(T)$ is bijective, i.e., $\mu \in \varrho(T)$. □

Definition 2.29. Let S be a linear operator in H, and $M \subset H$ a closed subspace. We say that M **reduces** S if the following two conditions are satisfied:

(a) $\mathrm{dom}\,S = (M \cap \mathrm{dom}\,S) \oplus (M^\perp \cap \mathrm{dom}\,S)$.

(b) $S(M \cap \mathrm{dom}\,S) \subset M$ and $S(M^\perp \cap \mathrm{dom}\,S) \subset M^\perp$.

The next lemma is easy to see and is therefore left as an exercise.

Lemma 2.30. *M reduces S if and only if $P_M S \subset S P_M$, where P_M denotes the orthogonal projection onto M.*

Remark 2.31. Assume that M reduces S. Then, with respect to the decomposition $H = M \oplus M^\perp$, the operator S can be written as follows in operator matrix form:

$$S = \begin{pmatrix} S|_M & 0 \\ 0 & S|_{M^\perp} \end{pmatrix}.$$

It is easy to see that $\sigma(S) = \sigma(S|_M) \cup \sigma(S|_{M^\perp})$, but $\sigma(S|_M)$ and $\sigma(S|_{M^\perp})$ are not necessarily disjoint.

Theorem 2.32. *Let T be a selfadjoint operator, E its spectral measure, and $\Delta \in \Sigma$. Then the following holds:*

(i) *$E_\Delta H$ reduces T.*

(ii) $\sigma(T|_{E_\Delta H}) \subset \overline{\sigma(T) \cap \Delta}$.

Proof. (i). We have $(\mathrm{id})_E(\chi_\Delta)_E = (\mathrm{id}\cdot\chi_\Delta)_E$ since

$$\mathrm{dom}\,\big((\mathrm{id})_E(\chi_\Delta)_E\big) = D_{\chi_\Delta} \cap D_{\mathrm{id}\cdot\chi_\Delta} = D_{\mathrm{id}\cdot\chi_\Delta} = \mathrm{dom}\,((\mathrm{id}\cdot\chi_\Delta)_E)\,.$$

This shows

$$E_\Delta T = (\chi_\Delta)_E \operatorname{id}_E \subset (\chi_\Delta \cdot \operatorname{id})_E = (\operatorname{id} \cdot \chi_\Delta)_E = \operatorname{id}_E (\chi_\Delta)_E = T E_\Delta.$$

By Lemma 2.30, $E_\Delta H$ reduces T.

(ii). Let $\lambda \notin \overline{\sigma(T) \cap \Delta}$. If $\sigma(T) \cap \Delta = \varnothing$, then $E_\Delta = E_{\Delta \cap \sigma(T)} = 0$ and (ii) holds. Otherwise, let $\lambda_0 \in \sigma(T) \cap \Delta$, and define

$$f := \operatorname{id} \cdot \chi_{\sigma(T)\cap\Delta} + \lambda_0 \cdot \chi_{(\sigma(T)\cap\Delta)^c}.$$

In the proof of (i), we have seen that $(\operatorname{id} \cdot \chi_{\sigma(T)\cap\Delta})_E = T E_\Delta$. Moreover,

$$\left(\chi_{(\sigma(T)\cap\Delta)^c}\right)_E = E_{(\sigma(T)\cap\Delta)^c} = E_{\varrho(T)\cup\Delta^c} = E_{(\varrho(T)\cup\Delta^c)\cap\sigma(T)} = E_{\Delta^c\cap\sigma(T)} = E_{\Delta^c}.$$

Hence

$$f(T) = T E_\Delta + \lambda_0 E_{\Delta^c}.$$

We have

$$f(\sigma(T)) = \left\{ t \cdot \chi_{\sigma(T)\cap\Delta}(t) + \lambda_0 \chi_{(\sigma(T)\cap\Delta)^c}(t) : t \in \sigma(T) \right\} = (\sigma(T)\cap\Delta) \cup \{\lambda_0\} = \sigma(T)\cap\Delta.$$

This shows $\sigma(f(T)) \subset \overline{f(\sigma(T))} = \overline{\sigma(T)\cap\Delta}$ and thus $\lambda \in \varrho(f(T))$.

Put $T_\Delta := T|_{E_\Delta H}$. Let $x \in \ker(T_\Delta - \lambda)$. Then $x \in E_\Delta H$, $Tx = \lambda x$. By Theorem 2.28, we have $f(T)x = f(\lambda)x$. Since $x = E_\Delta x$, we get by the representation of $f(T)$ above that $Tx = \lambda_0 x$ and thus $(\lambda - \lambda_0)x = 0$. This shows $x = 0$

Let $y \in E_\Delta H$. Since $\lambda \in \varrho(f(T))$, there exists $x \in \operatorname{dom} f(T)$, such that

$$y = (f(T) - \lambda)x = T E_\Delta x + \lambda_0 E_{\Delta^c} x - \lambda x.$$

As $E_\Delta H$ reduces T, we have

$$y = E_\Delta y = T E_\Delta x - \lambda E_\Delta x = (T - \lambda) E_\Delta x = (T_\Delta - \lambda) E_\Delta x.$$

Thus, $\lambda \in \varrho(T_\Delta)$. ☐

3. Fourier Transform and Sobolev Spaces

The origin of functional analysis lies in the study of certain special function spaces and operators on and between them. Whereas having considered general Hilbert spaces in the last two chapters, we get back to these origins in this chapter, mainly operating in $L^p(\Omega)$, especially in $L^2(\Omega)$, where Ω is an open subset of $\mathbb{R}^n$.

Solutions of (ordinary and partial) differential equations have to be in some sense differentiable. However, many differential equations, especially those arising from applications, do not have solutions which are differentiable in the classical sense. A very simple example of this kind is $y' = \chi_{[0,1]}$. For this reason, Sobolev introduced the so-called weak or generalized derivative, a notion which is motivated by the formula of partial integration. The advantage of this approach is that the function is taken as a whole – in contrast to the local nature of classical differentiability. It turns out that many differential equations have solutions when the derivatives in the equation are understood in the weak sense.

Sobolev spaces are spaces of functions having generalized derivatives up to a certain order. A useful tool for studying those spaces is the Fourier transform as it turns differentiation into multiplication and vice versa. It is a unitary operator in $L^2(\mathbb{R}^n)$. But the Fourier transform also has a rather practical meaning as it gives information on the frequencies "inside a function". It can be seen as a continuous version of Fourier series. Therefore, it is a widely used object in engineering sciences. Finally, it should be mentioned that the Fourier transform is the basis of Harmonic Analysis where it is defined not only for functions on $\mathbb{R}^n$ but also on locally compact groups.

We start by setting some notation. In this chapter, Ω will always denote an open subset of $\mathbb{R}^n$. For $k \in \mathbb{N}_0$, we set

$$C^k(\Omega) := \{f \in C(\Omega) : f \text{ is } k\text{-times continuously differentiable}\},$$

and

$$C_c^k(\Omega) := \{f \in C^k(\Omega) : \operatorname{supp}(f) \text{ compact}\}.$$

For a function $f \in C^k(\Omega)$ and a *multiindex* $\alpha \in \mathbb{N}_0^n$ we set

$$|\alpha| := \sum_{i=1}^n \alpha_i \qquad \text{and} \qquad \partial^\alpha f := \partial_1^{\alpha_1} \cdot \ldots \cdot \partial_n^{\alpha_n} f.$$

The latter is, of course, only defined for $|\alpha| \le k$. Finally, for $p \in [1, \infty]$, we set $\|\cdot\|_p := \|\cdot\|_{L^p(\mathbb{R}^n)}$.

Theorem and Definition 3.1. *Let $p \in [1,\infty]$, $f \in L^1(\mathbb{R}^n)$, and $g \in L^p(\mathbb{R}^n)$.*

(i) *For almost every $x \in \mathbb{R}^n$, the function $y \mapsto f(x-y)g(y)$, $y \in \mathbb{R}^n$, belongs to $L^1(\mathbb{R}^n)$.*

(ii) *The **convolution** of f and g is defined by*

$$(f * g)(x) := \int_{\mathbb{R}^n} f(x-y)g(y)\,\mathrm{d}y, \quad x \in \mathbb{R}^n.$$

(iii) *We have* $f * g \in L^p(\mathbb{R}^n)$ *and* $\|f * g\|_p \le \|f\|_1 \|g\|_p$.

Proof. Exercise. $\square$

Lemma 3.2. *Let* $f \in C_c^k(\mathbb{R}^n)$ *and* $g \in L^p(\mathbb{R}^n)$. *Then* $f * g \in C^k(\mathbb{R}^n)$, *and*

$$\partial^\alpha (f * g) = (\partial^\alpha f) * g$$

for all multiindices α *with* $|\alpha| \le k$.

Proof. Let $\operatorname{supp}(f) \subset B_r(0)$ for some $r \in \mathbb{R}$. First, we consider $k = 0$ and note that f is uniformly continuous. Let $\varepsilon > 0$. Then there exists $\delta > 0$ such that for $w, z \in \mathbb{R}^n$ with $|w - z| < \delta$ we have $|f(w) - f(z)| < \varepsilon$. Without loss of generality, we can assume $\delta < 1$. Let $x, x' \in \mathbb{R}^n$ with $|x - x'| < \delta$. Then

$$\begin{aligned} |(f * g)(x) - (f * g)(x')| &\le \int_{\mathbb{R}^n} |f(x-y) - f(x'-y)| \cdot |g(y)| \, \mathrm{d}y \le \varepsilon \int_{B_{r+1}(x)} |g(y)| \, \mathrm{d}y \\ &\le \varepsilon \|g\|_p \cdot \operatorname{vol}\left(B_{r+1}(0)\right)^{\frac{1}{q}} . \end{aligned}$$

Now, let $k = 1$. For $u \in C(\mathbb{R}^n)$ define

$$\left(\Delta_{j,h} u\right)(x) := \frac{u(x + he_j) - u(x)}{h}, \quad j = 1, \dots, n, \;\; h \in \mathbb{R} \setminus \{0\}.$$

Obviously we have $\Delta_{j,h}(f * g) = (\Delta_{j,h} f) * g$. Thus,

$$\begin{aligned} |\Delta_{j,h}(f * g)(x) - (\partial_j f) * g(x)| &= \left|[(\Delta_{j,h} f) * g - (\partial_j f) * g](x)\right| \\ &= \big|[\underbrace{(\Delta_{j,h} f - \partial_j f)}_{\in C_c(\mathbb{R}^n) \subset L^q(\mathbb{R}^n)} * g](x)\big| \\ &\le \|\Delta_{j,h} f - \partial_j f\|_q \|g\|_p. \end{aligned}$$

The latter tends to zero as $h \to 0$ by Lebesgue's theorem, since $\Delta_{j,h} f \to \partial_j f$ pointwise and $|(\Delta_{j,h} f)(x) - \partial_j f(x)| = |\partial_j f(\xi) - \partial_j f(x)| \le 2\|\partial_j f\|_\infty$ with some $\xi \in [x, x + he_j]$. The claim for $k > 1$ follows by induction. $\square$

Define the function $g : \mathbb{R}_0^+ \to \mathbb{R}$ by

$$g(r) := \begin{cases} e^{-\frac{1}{1-r}}, & \text{for } r \in [0,1), \\ 0, & \text{for } r \in [1,\infty). \end{cases}$$

Obviously, $g \in C^\infty([0,\infty))$ and $\operatorname{supp}(g) = [0,1]$. Now, define

$$c := \int_{\mathbb{R}^n} g(|x|^2) \, \mathrm{d}x \qquad \text{and} \qquad \delta_k(x) := \frac{k^n}{c} g\left(|kx|^2\right), \quad x \in \mathbb{R}^n.$$

Then $\delta_k \in C_c^\infty(\mathbb{R}^n)$ and $\operatorname{supp}(\delta_k) = \overline{B_{\frac{1}{k}}(0)}$. Moreover,

$$\int_{\mathbb{R}^n} \delta_k(x) \, \mathrm{d}x = \frac{k^n}{c} \int_{\mathbb{R}^n} g(|kx|^2) \, \mathrm{d}x = \frac{1}{c} \int_{\mathbb{R}^n} g(|x|^2) \, \mathrm{d}x = 1.$$

Theorem 3.3. *Let $f \in L^p(\mathbb{R}^n)$, $p \in [1,\infty)$. Then the following statements hold.*

(i) *For all $k \in \mathbb{N}$ we have $\delta_k * f \in C^\infty(\mathbb{R}^n)$.*

(ii) $\|f - \delta_k * f\|_p \to 0$ *as $k \to \infty$.*

In particular, $C^\infty(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$.

Proof. (i) follows from Lemma 3.2. For (ii), let $f = \chi_Q$, where $Q = [a_1, b_1] \times \ldots \times [a_n, b_n] \subset \mathbb{R}^n$. Then

$$(\chi_Q - \delta_k * \chi_Q)(x) = \int (\chi_Q(x) - \chi_Q(x-y))\, \delta_k(y)\, \mathrm{d}y.$$

Let $q \in (1,\infty]$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Then we have

$$\begin{aligned}
\|\chi_Q - \delta_k * \chi_Q\|_p^p &= \int \left| \int (\chi_Q(x) - \chi_Q(x-y))\, \delta_k(y)\, \mathrm{d}y \right|^p \mathrm{d}x \\
&\leq \int \left(\int |\chi_Q(x) - \chi_Q(x-y)||\delta_k(y)|\, \mathrm{d}y \right)^p \mathrm{d}x \\
&\leq \int \left(\int |\chi_Q(x) - \chi_Q(x-y)|^p \delta_k(y)\, \mathrm{d}y \right)^{\frac{p}{p}} \left(\int \delta_k(y)\, \mathrm{d}y \right)^{\frac{p}{q}} \mathrm{d}x \\
&= \int \delta_k(y) \left(\int |\chi_Q(x) - \chi_Q(x-y)|^p\, \mathrm{d}x \right) \mathrm{d}y.
\end{aligned}$$

For fixed $y \in \mathbb{R}^n$, we put $Q_y := [Q \cap (y+Q)^c] \cup [Q^c \cap (y+Q)]$. Then $\chi_{Q_y}(x) = |\chi_Q(x) - \chi_Q(x-y)|$, and hence

$$\|\chi_Q - \delta_k * \chi_Q\|_p^p \leq \int \delta_k(y) \int \chi_{Q_y}(x)\, \mathrm{d}x\, \mathrm{d}y = \int_{\mathbb{R}^n} \delta_k(y) \operatorname{vol}(Q_y)\, \mathrm{d}y.$$

Let $\varepsilon > 0$. Then, there exists $r > 0$ such that $\operatorname{vol}(Q_y) < \varepsilon$ for $y \in \overline{B_r(0)} =: B$. Therefore,

$$\|\chi_Q - \delta_k * \chi_Q\|_p^p \leq \varepsilon \int_B \delta_k(y)\, \mathrm{d}y + 2 \operatorname{vol}(Q) \int_{B^c} \delta_k(y)\, \mathrm{d}y \leq \varepsilon$$

for k sufficiently large.

The theorem obviously also holds for stepfunctions. Let now $f \in L^p(\mathbb{R}^n)$ be arbitrary, $\varepsilon > 0$. Then there exists a step function φ such that $\|f - \varphi\|_p < \frac{\varepsilon}{3}$. Hence, for $k \geq K$ we have

$$\|f - (\delta_k * f)\|_p \leq \|f - \varphi\|_p + \|\varphi - (\delta_k * \varphi)\|_p + \|\delta_k * (\varphi - f)\|_p \leq \frac{2}{3}\varepsilon + \|\delta_k\|_1 \cdot \|\varphi - f\|_p \leq \varepsilon,$$

where we have used Theorem 3.1. □

Theorem 3.4. *$C_c^\infty(\Omega)$ is a dense subspace of $L^p(\Omega)$ for each $p \in [1,\infty)$.*

Proof. First, we assume that Ω is bounded. Let $f \in L^p(\Omega)$. For $\delta > 0$, define the set $\Omega_\delta := \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) \geq \delta\}$. For $m \in \mathbb{N}$, define $f_m : \mathbb{R}^n \to \mathbb{C}$ by

$$f_m(x) := \begin{cases} f(x), & \text{for } x \in \Omega_{\frac{1}{m}} \\ 0, & \text{otherwise.} \end{cases}$$

Then $f_m \in L^p(\mathbb{R}^n)$ and $\|f_m|_\Omega - f\|_{L^p(\Omega)} \to 0$ as $m \to \infty$. We have

$$(\delta_k * f_m)(x) = \int_{\mathbb{R}^n} f_m(x-y)\delta_k(y)\,\mathrm{d}y = \int_{B_{\frac{1}{k}}(0)} f_m(x-y)\delta_k(y)\,\mathrm{d}y.$$

For $k > m$ we have $\frac{1}{m} - \frac{1}{k} > 0$. If $x \notin \Omega_{\frac{1}{m}-\frac{1}{k}}$, then $x - y \notin \Omega_{\frac{1}{m}}$ for all $y \in B_{\frac{1}{k}}(0)$. That is, $\operatorname{supp}(\delta_k * f_m) \subset \Omega_{\frac{1}{m}-\frac{1}{k}}$. Hence, $(\delta_k * f_m)|_\Omega \in C_c^\infty(\Omega)$ and

$$\|f_m|_\Omega - (\delta_k * f_m)|_\Omega\|_{L^p(\Omega)} = \|f_m - (\delta_k * f_m)\|_{L^p(\mathbb{R}^n)} \to 0 \quad \text{as } k \to \infty.$$

The claim thus follows from

$$\|f - (\delta_k * f_m)|_\Omega\|_{L^p(\Omega)} \leq \|f - f_m|_\Omega\|_{L^p(\Omega)} + \|f_m|_\Omega - (\delta_k * f_m)|_\Omega\|_{L^p(\mathbb{R}^n)}.$$

For general Ω and $f \in L^p(\Omega)$, set $\Omega_k := \Omega \cap B_k(0)$, and choose $\varphi_k \in C_c^\infty(\Omega_k)$ such that $\|\varphi_k - f|\Omega_k\|_{L^p(\Omega_k)} \to 0$ as $k \to \infty$. Put $f_k(x) := \varphi_k(x)$ for $x \in \Omega_k$ and $f_k(x) := 0$ for $x \in \Omega \setminus \Omega_k$. Then $f_k \in C_c^\infty(\Omega)$ and

$$\|f - f_k\|_{L^p(\Omega)}^p = \|f|_{\Omega_k} - \varphi_k\|_{L^p(\Omega_k)}^p + \|f_{\Omega\setminus\Omega_k}\|_{L^p(\Omega\setminus\Omega_k)}^p,$$

which tends to zero as $k \to \infty$. □

Remark 3.5. Note that the approximating sequence in the above proof was defined independently of p. Thus, if, e.g., $f \in L^1(\Omega) \cap L^2(\Omega)$, then there exists $f_k \in C_c^\infty(\Omega)$ such that both $\|f - f_k\|_{L^1(\Omega)} \to 0$ and $\|f - f_k\|_{L^2(\Omega)} \to 0$ as $k \to \infty$.

Lemma 3.6 (Fundamental Lemma of Variational Calculus). *Let $f \in L^1_{loc}(\Omega)$ such that $\int_\Omega f\varphi\,\mathrm{d}x = 0$ for all $\varphi \in C_c^\infty(\Omega)$. Then $f = 0$.*

Proof. Let us first assume that Ω is bounded. Define Ω_δ and f_m as in the proof of Theorem 3.4. Note that $f_m \in L^1(\mathbb{R}^n)$ for each $m \in \mathbb{N}$. As before, for $k > m$ we have $(\delta_k * f_m)|_\Omega \in C_c^\infty(\Omega)$ with support in $\Omega_{\frac{1}{m}-\frac{1}{k}} \subset \Omega_{\frac{1}{m}-\frac{1}{m+1}} =: \Omega'_m$. Now, $\delta_k * f_m \to f_m$ in $L^1(\mathbb{R}^n)$ as $k \to \infty$. Thus, there exists (k_j) such that $\delta_{k_j} * f_m \to f_m$ pointwise a.e. as $j \to \infty$. Therefore, $\overline{f}(\delta_{k_j} * f_m)|_{\Omega'_m} \to \overline{f} f_m|_{\Omega'_m}$ and

$$\left|\overline{f}(\delta_{k_j} * f_m)|_{\Omega'_m}\right| \leq \left|f|_{\Omega'_m}\right| \cdot \|\delta_{k_j}\|_1 \cdot \|f_m\|_1 = \left|f|_{\Omega'_m}\right| \cdot \|f_m\|_1 \in L^1(\Omega'_m).$$

Hence, Lebesgue's theorem implies

$$0 = \int_\Omega \overline{f}(\delta_{k_j} * f_m)\,\mathrm{d}x = \int_{\Omega'_m} \overline{f}\left(\delta_{k_j} * f_m\right)\mathrm{d}x \to \int_{\Omega'_m} \overline{f} f_m\,\mathrm{d}x = \int_{\Omega_{\frac{1}{m}}} \overline{f} f_m\,\mathrm{d}x = \int_{\Omega_{\frac{1}{m}}} |f|^2\,\mathrm{d}x$$

as $j \to \infty$. Thus, $f|_{\Omega_{\frac{1}{m}}} = 0$ for each $m \in \mathbb{N}$. That is, $f = 0$.

For arbitrary Ω, set $\Omega_n := \Omega \cap B_n(0)$. Then, whenever $\varphi \in C_c^\infty(\Omega_n)$, we have that

$$\int_{\Omega_n} (f|_{\Omega_n})\varphi \,\mathrm{d}x = \int_\Omega f\widetilde{\varphi}\,\mathrm{d}x = 0,$$

where $\widetilde{\varphi} \in C_c^\infty(\Omega)$ is defined by $\widetilde{\varphi}|_{\Omega_n} := \varphi$ and $\widetilde{\varphi}|_{\Omega\setminus\Omega_n} := 0$. By the above, $f|_{\Omega_n} = 0$ for each $n \in \mathbb{N}$, and therefore $f = 0$. □

As a motivation for the next definition, let $f \in C^1([a,b])$ and $\varphi \in C_c^\infty((a,b))$. Then

$$\int_a^b f'\varphi\,\mathrm{d}x = [f\varphi]_a^b - \int_a^b f\varphi'\,\mathrm{d}x = -\int_a^b f\varphi'\,\mathrm{d}x.$$

Inductively, for $f \in C^k([a,b])$ and $\varphi \in C_c^\infty((a,b))$ we find

$$\int_a^b f^{(k)}\varphi\,\mathrm{d}x = (-1)^k \int_a^b f\varphi^{(k)}\,\mathrm{d}x.$$

But also, for an open set $\Omega \subset \mathbb{R}^n$, $f \in C^k(\overline{\Omega})$, and $\varphi \in C_c^\infty(\Omega)$, we have

$$\int_\Omega (\partial^\alpha f)\varphi\,\mathrm{d}x = (-1)^{|\alpha|}\int_\Omega f(\partial^\alpha\varphi)\,\mathrm{d}x.$$

Definition 3.7. Let $f \in L^1_{\mathrm{loc}}(\Omega)$, $\alpha \in \mathbb{N}_0^n$. A function $g \in L^1_{\mathrm{loc}}(\Omega)$ is called the **α-th weak (or generalized) derivative** of f, if

$$\int_\Omega g\varphi\,\mathrm{d}x = (-1)^{|\alpha|}\int_\Omega f(\partial^\alpha\varphi)\,\mathrm{d}x \qquad \forall\,\varphi \in C_c^\infty(\Omega).$$

Lemma 3.8. *The α-th weak derivative of f (if it exists) is unique.*

Proof. If $g, h \in L^1_{\mathrm{loc}}(\Omega)$ are two α-th weak derivatives of f, then

$$\int_\Omega g\varphi\,\mathrm{d}x = (-1)^{|\alpha|}\int_\Omega f(\partial^\alpha\varphi)\,\mathrm{d}x = \int_\Omega h\varphi\,\mathrm{d}x \qquad \forall\,\varphi \in C_c^\infty(\Omega).$$

Thus, Lemma 3.6 implies $g = h$. □

Definition 3.9. Let $\Omega \subset \mathbb{R}^n$ be open and $m \in \mathbb{N}_0$. The **Sobolev space** of m-th order on Ω is defined by

$$H^m(\Omega) := \left\{f \in L^2(\Omega) : \partial^\alpha f \text{ exists and } \partial^\alpha f \in L^2(\Omega)\,\forall\,\alpha \in \mathbb{N}_0^n, |\alpha| \le m\right\}.$$

On $H^m(\Omega)$, we consider the following inner product:

$$\langle f, g\rangle_{H^m} := \sum_{|\alpha|\le m} \langle \partial^\alpha f, \partial^\alpha g\rangle_{L^2}, \qquad f, g \in H^m(\Omega).$$

Moreover, we define

$$H_0^m(\Omega) := \overline{C_c^\infty(\Omega)}^{H^m}.$$

Remark 3.10. There are also Sobolev spaces in $L^p(\Omega)$, $p \neq 2$, defined by

$$W^{m,p}(\Omega) := \{f \in L^p(\Omega) : \delta^\alpha f \text{ exists and } \delta^\alpha f \in L^p(\Omega) \,\forall\, |a| \leq m\}.$$

On $W^{m,p}(\Omega)$ one defines the following norm:

$$\|f\|_{W^{m,p}} := \Big(\sum_{|a| \leq m} \|\partial^\alpha f\|_{L^p}^p \Big)^{1/p}, \qquad f \in W^{m,p}(\Omega).$$

Similarly as in the proof of the next theorem, one shows that these are Banach spaces. However – here, we will only deal with the H^m-Sobolev spaces in $L^2(\Omega)$.

Theorem 3.11. *$(H^m(\Omega), \langle\cdot,\cdot\rangle_{H^m})$ is a Hilbert space.*

Proof. Let (f_k) be a Cauchy sequence in $H^m(\Omega)$. From $\|f\|_{H^m}^2 = \sum_{|a| \leq m} \|\partial^\alpha f\|_2^2$, $f \in H^m$, we see that $(\partial^\alpha f_k)$ is a Cauchy sequence in $L^2(\Omega)$ for each $|\alpha| \leq m$. Let f_α be the limit of $(\partial^\alpha f_k)$ in $L^2(\Omega)$, $|\alpha| \leq m$. Set $f := f_{(0,\ldots,0)}$. We show that $\partial^\alpha f = f_\alpha$. Then $f \in H^m(\Omega)$ and $\|f_k - f\|_{H^m} \to 0$. And indeed, for $\varphi \in C_c^\infty(\Omega)$ and $|\alpha| \leq m$ we have

$$\langle f_\alpha, \varphi\rangle_{L^2} = \lim_{k\to\infty} \langle \partial^\alpha f_k, \varphi\rangle_{L^2} = \lim_{k\to\infty} (-1)^{|a|} \langle f_k, \partial^\alpha \varphi\rangle_{L^2} = (-1)^{|\alpha|} \langle f, \partial^\alpha\varphi\rangle_{L^2}.$$

Thus, $f_\alpha = \partial^\alpha f$. □

Definition 3.12. For $f \in L^1(\mathbb{R}^n)$, put

$$\widehat{f}(y) := (\mathcal{F}f)(y) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(x) e^{-i\langle x,y\rangle}\,\mathrm{d}x, \qquad y \in \mathbb{R}^n.$$

$\mathcal{F}f$ is called the **Fourier transform** of f.

For simplicity, we shall often write $xy := x \cdot y := \langle x,y\rangle_{\mathbb{R}^n}$ and $x^2 := |x|^2 = \langle x,x\rangle$, $x, y \in \mathbb{R}^n$. Moreover, the factor $(2\pi)^{-n/2}$ preceding the integral is often abbreviated by c.

Theorem 3.13. *Let $f \in L^1(\mathbb{R}^n)$. Then $\mathcal{F}f \in C(\mathbb{R}^n)$ and $\lim_{|y|\to\infty} (\mathcal{F}f)(y) = 0$. Moreover, $\|\mathcal{F}f\|_\infty \leq (2\pi)^{-\frac{n}{2}} \|f\|_1$.*

Proof. The "Moreover"-part is clear. If $y \to z$, then $f(x)e^{-ixy} \to f(x)e^{-ixz}$ for all $x \in \mathbb{R}^n$. By Lebesgue's theorem, we have $(\mathcal{F}f)(y) \to (\mathcal{F}f)(z)$, and thus $\mathcal{F}f \in C(\mathbb{R}^n)$. For the remainder, let $f \in C_c^\infty(\mathbb{R}^n)$, $R > 0$, and $|y| \geq R$. Then there exists a coordinate y_j of y with $|y_j| \geq \frac{R}{\sqrt{n}}$, and hence

$$\begin{aligned} |(\mathcal{F}f)(y)| &= c \cdot \left| \int_{\mathbb{R}^n} f(x) e^{-ix\cdot y}\,\mathrm{d}x \right| = c \cdot \left| \int_{\mathbb{R}^n} (\partial_j f)(x) \cdot \frac{1}{-iy_j} e^{-ix\cdot y}\,\mathrm{d}x \right| \\ &\leq c \cdot \int_{\mathbb{R}^n} \frac{|(\partial_j f)(x)|}{|y_j|}\,\mathrm{d}x \leq \frac{c\sqrt{n}}{R} \int_{\mathbb{R}^n} |(\partial_j f)(x)|\,\mathrm{d}x = \frac{c\sqrt{n}}{R} \|\partial_j f\|_1. \end{aligned}$$

This proves $\lim_{|y|\to\infty} (\mathcal{F}f)(y) = 0$. □

Definition 3.14. A function $f : \mathbb{R}^n \to \mathbb{C}$ is called **rapidly decreasing** if

$$\lim_{|x|\to\infty} x^\alpha f(x) = 0 \qquad \forall\, \alpha \in \mathbb{N}_0^n,$$

where $x^\alpha := x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n}$. The space

$$\mathcal{S}(\mathbb{R}^n) := \left\{ f \in C^\infty(\mathbb{R}^n) : \partial^\beta f \text{ rapidly decreasing } \forall\, \beta \in \mathbb{N}_0^n \right\}$$

is called **Schwartz space**, and the functions in $\mathcal{S}(\mathbb{R}^n)$ are called **Schwartz functions**.

The proofs of the following two lemmas, characterizing rapidly decreasing and Schwartz functions, are simple and are therefore left to the reader.

Lemma 3.15. *Let $f : \mathbb{R}^n \to \mathbb{C}$. Then the following statements are equivalent:*

(i) *f is rapidly decreasing.*

(ii) *For every polynomial $p : \mathbb{R}^n \to \mathbb{C}$ we have*

$$\lim_{|x|\to\infty} p(x) f(x) = 0.$$

(iii) *For all $m \in \mathbb{N}_0$ we have*

$$\lim_{|x|\to\infty} |x|^m f(x) = 0.$$

Lemma 3.16. *For $f : \mathbb{R}^n \to \mathbb{C}$, the following statements are equivalent.*

(a) *$f \in \mathcal{S}(\mathbb{R}^n)$.*

(b) *For all $m \in \mathbb{N}_0$ and $\beta \in \mathbb{N}_0^n$ we have*

$$\sup_{x\in\mathbb{R}^n} (1 + |x|^m) \left|(\partial^\beta f)(x)\right| < \infty.$$

Lemma 3.17. *For all $p \in [1, \infty]$, we have $\mathcal{S}(\mathbb{R}^n) \subset L^p(\mathbb{R}^n)$.*

Proof. For $p = \infty$, this is obvious. Let $p \in [1, \infty)$. Choose $m > \frac{n}{p}$ and let $f \in \mathcal{S}(\mathbb{R}^n)$. Then there exists $C > 0$ such that

$$(1 + |x|^m) |f(x)| \le C \quad \forall\, x \in \mathbb{R}^n.$$

If $mp = n + \varepsilon$, then

$$\begin{aligned}\int_{\mathbb{R}^n} |f|^p \,\mathrm{d}x &\le \int_{\mathbb{R}^n} \left(\frac{C}{1 + |x|^m}\right)^p \mathrm{d}x = C^p \cdot \mathrm{Vol}_{n-1}(S^{n-1}) \cdot \int_0^\infty \frac{r^{n-1}}{(1 + r^m)^p} \,\mathrm{d}r \\ &\le K \cdot \left(\int_0^1 1 \,\mathrm{d}r + \int_1^\infty r^{n-1-mp} \,\mathrm{d}r\right) = K \left(1 + \int_1^\infty r^{-1-\varepsilon} \,\mathrm{d}r\right) < \infty,\end{aligned}$$

where $K := C^p \cdot \mathrm{Vol}_{n-1}(S^{n-1})$. □

Remark 3.18. (a) Obviously, $C_c^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$, and thus $\mathcal{S}(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$.
(b) If $f \in \mathcal{S}(\mathbb{R}^n)$ and $\alpha \in \mathbb{N}_0^n$, then also $x^\alpha f, \partial^\alpha f \in \mathcal{S}(\mathbb{R}^n)$.

Lemma 3.17 allows us to apply the Fourier transform to Schwartz functions (and their derivatives). The following lemma reveals the important fact that the Fourier transform turns derivation into multiplication with the free variable and vice versa.

Lemma 3.19. *Let $f \in \mathcal{S}(\mathbb{R}^n)$ and $\alpha \in \mathbb{N}_0^n$. Then:*

(i) $\mathcal{F}f \in C^\infty(\mathbb{R}^n)$ *and* $\partial^\alpha(\mathcal{F}f) = (-i)^{|\alpha|}\mathcal{F}(x^\alpha f)$.

(ii) $(\mathcal{F}(\partial^\alpha f))(y) = i^{|\alpha|} y^\alpha (\mathcal{F}f)(y)$ *for all* $y \in \mathbb{R}^n$.

Proof. Let $y \in \mathbb{R}^n$. Then we have

$$\begin{aligned}\partial^\alpha(\mathcal{F}f)(y) &= c\frac{\partial^\alpha}{\partial y^\alpha}\int_{\mathbb{R}^n} f(x)e^{-ixy}\,\mathrm{d}x = c\int_{\mathbb{R}^n} f(x)\frac{\partial^\alpha}{\partial y^\alpha}e^{-ixy}\,\mathrm{d}x \\ &= c\int_{\mathbb{R}^n} f(x)(-i)^{|\alpha|}x^\alpha e^{-ixy}\,\mathrm{d}x = (-i)^{|\alpha|}\left(\mathcal{F}(x^\alpha f)\right)(y)\end{aligned}$$

as well as

$$\begin{aligned}(\mathcal{F}(\partial^\alpha f))(y) &= c\int_{\mathbb{R}^n}(\partial^\alpha f)(x)e^{-ixy}\,\mathrm{d}x = c(-1)^{|\alpha|}\int_{\mathbb{R}^n} f(x)\left(\frac{\partial^\alpha}{\partial x^\alpha}e^{-ixy}\right)\mathrm{d}x \\ &= ci^{|\alpha|}y^\alpha\int_{\mathbb{R}^n} f(x)e^{-ixy}\,\mathrm{d}x = i^{|\alpha|}y^\alpha(\mathcal{F}f)(y).\end{aligned}$$

The lemma is proved. □

Lemma 3.20. *If $f \in \mathcal{S}(\mathbb{R}^n)$ then also $\mathcal{F}f \in \mathcal{S}(\mathbb{R}^n)$.*

Proof. We already established by Lemma 3.19 that $\mathcal{F}f \in C^\infty(\mathbb{R}^n)$. We have to show that $y^\alpha(\partial^\beta\mathcal{F}f)(y) \to 0$ as $|y| \to \infty$ for all $\alpha, \beta \in \mathbb{N}_0^n$. By Lemma 3.19,

$$y^\alpha(\partial^\beta\mathcal{F}f)(y) = (-i)^{|\beta|}y^\alpha\mathcal{F}(x^\beta f)(y) = (-i)^{|\alpha|+|\beta|}\mathcal{F}(\partial^\alpha x^\beta f)(y).$$

Due to Remark 3.18, we know that $\partial^\alpha x^\beta f \in \mathcal{S}(\mathbb{R}^n) \subset L^1(\mathbb{R}^n)$. Hence, Lemma 3.13 implies the claim. □

Define $\gamma : \mathbb{R}^n \to \mathbb{R}$ by $\gamma(x) := e^{-\frac{x^2}{2}}$. If $n = 1$ then

$$\left(\int_{\mathbb{R}} e^{-\frac{x^2}{2}}\,\mathrm{d}x\right)^2 = \int_{\mathbb{R}}\int_{\mathbb{R}} e^{-\frac{x^2}{2}}e^{-\frac{y^2}{2}}\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}^2} e^{-\frac{x^2+y^2}{2}}\,\mathrm{d}(x,y).$$

Define the transformation $\varphi : [0,\infty) \times [0, 2\pi) \to \mathbb{R}^2$ by $\varphi(r,t) := r(\cos t, \sin t)^T$. Then

$$D\varphi(t) = \begin{pmatrix}\cos t & -r\sin t \\ \sin t & r\cos t\end{pmatrix},$$

and thus $\det D\varphi(t) = r$. Therefore,

$$\left(\int_{\mathbb{R}} e^{-\frac{x^2}{2}}\,\mathrm{d}x\right)^2 = \int_{r=0}^{\infty}\int_{t=0}^{2\pi} e^{-\frac{r^2}{2}} r\,\mathrm{d}t\,\mathrm{d}r = -2\pi\int_0^{\infty} -re^{-\frac{r^2}{2}}\,\mathrm{d}r = -2\pi(0-1) = 2\pi.$$

Thus, we know

$$\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-\frac{x^2}{2}}\,\mathrm{d}x = 1.$$

Let $n > 1$. Since $\gamma(x) = \prod_{k=1}^{n} e^{-\frac{x_k^2}{2}}$, we have

$$\frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n}\gamma(x)\,\mathrm{d}x = \left(\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-\frac{x^2}{2}}\,\mathrm{d}x\right)^n = 1.$$

We set $\gamma_a : \mathbb{R}^n \to \mathbb{R}$, $\gamma_a(x) := \gamma(ax)$.

Lemma 3.21. *For $a > 0$, we have $(\mathcal{F}\gamma_a)(y) = \frac{1}{a^n}\gamma\left(\frac{y}{a}\right)$, $y \in \mathbb{R}^n$. In particular $\mathcal{F}\gamma = \gamma$.*

Proof. Put $\varphi(x) := ax$, $x \in \mathbb{R}^n$. Then $D\varphi = aI$ and thus $\det D\varphi = a^n$. Hence,

$$(\mathcal{F}\gamma_a)(y) = \frac{c}{a^n}\cdot\int_{\mathbb{R}^n} a^n\gamma(\varphi(x))e^{-i\varphi(x)\frac{y}{a}}\,\mathrm{d}x = \frac{c}{a^n}\int_{\mathbb{R}^n}\gamma(x)e^{-ix\frac{y}{a}}\,\mathrm{d}x = \frac{1}{a^n}(\mathcal{F}\gamma)\left(\frac{y}{a}\right).$$

It remains to prove $\mathcal{F}\gamma = \gamma$. For this, let $n = 1$. Then γ satisfies the initial value problem

$$y' + xy = 0, \qquad y(0) = 1. \tag{3.1}$$

Taking the Fourier transform and using Lemma 3.19 gives

$$0 = \mathcal{F}(\gamma' + x\gamma) = iy(\mathcal{F}\gamma) + i(\mathcal{F}\gamma)'.$$

Moreover, we have

$$(\mathcal{F}\gamma)(0) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\gamma(x)\,\mathrm{d}x = 1.$$

So, both γ and $\mathcal{F}\gamma$ satisfy the same initial value problem and hence must be equal.

For the case $n > 1$, we observe that $\gamma(x_1, \ldots, x_n) = \gamma(x_1)\cdot\ldots\cdot\gamma(x_n)$. This and the above also imply $\mathcal{F}\gamma = \gamma$ for $n > 1$. □

Lemma 3.22. *For $f \in \mathcal{S}(\mathbb{R}^n)$, we have $\mathcal{F}^2 f = f(-\cdot)$.*

Proof. First of all, we observe that for $f, g \in \mathcal{S}(\mathbb{R}^n)$ the mapping $(x, y) \mapsto f(y)g(x)e^{-ixy}$ belongs to $L^1(\mathbb{R}^{2n})$ and that by Fubini's theorem we have

$$\int_{\mathbb{R}^n}(\mathcal{F}f)(y)g(y)\,\mathrm{d}y = c\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x)g(y)e^{-ixy}\,\mathrm{d}y\,\mathrm{d}x = \int_{\mathbb{R}^n} f(x)(\mathcal{F}g)(x)\,\mathrm{d}x. \tag{3.2}$$

Let $y_0 \in \mathbb{R}^n$ and $a > 0$ be arbitrary, and set $g(x) = e^{-ixy_0}\gamma_a(x)$, $x \in \mathbb{R}^n$. Then, since

$$(\mathcal{F}g)(y) = c\int_{\mathbb{R}^n} e^{-ixy_0}\gamma_a(x)e^{-ixy}\,\mathrm{d}x = (\mathcal{F}\gamma_a)(y + y_0),$$

we get, using (3.2),

$$\int_{\mathbb{R}^n} (\mathcal{F}f)(x)g(x)\,\mathrm{d}x = \int_{\mathbb{R}^n} f(x)(\mathcal{F}\gamma_a)(x+y_0)\,\mathrm{d}x = \frac{1}{a^n}\int_{\mathbb{R}^n} f(x)\gamma\left(\frac{x+y_0}{a}\right)\mathrm{d}x$$
$$= \int_{\mathbb{R}^n} f(a\varphi(x)-y_0)\gamma(\varphi(x))\cdot|\det D\varphi(x)|\,\mathrm{d}x = \int_{\mathbb{R}^n} f(ax-y_0)\gamma(x)\,\mathrm{d}x,$$

where $\varphi(x) := \frac{x+y_0}{a}$. Letting $a \to 0$ and using Lebesgue's theorem with majorants $|\mathcal{F}f|$ and $\|f\|_\infty\gamma$, respectively, implies $(\mathcal{F}^2 f)(y_0) = f(-y_0)$. □

Theorem 3.23. *The Fourier transform $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is a bijection, and its inverse is given by*

$$\mathcal{F}^{-1}g(x) = \frac{1}{\sqrt{2\pi}^n}\int_{\mathbb{R}^n} g(y)e^{ixy}\,\mathrm{d}y, \qquad x \in \mathbb{R}^n. \tag{3.3}$$

Moreover, we have

$$\langle \mathcal{F}f, \mathcal{F}g\rangle_{L^2} = \langle f, g\rangle_{L^2}, \qquad f, g \in \mathcal{S}(\mathbb{R}^n). \tag{3.4}$$

Proof. By Lemma 3.22, we have $\mathcal{F}^4 = I_{\mathcal{S}(\mathbb{R}^n)}$. Thus, $\mathcal{F}$ is bijective and $\mathcal{F}^3 = \mathcal{F}^{-1}$. So,

$$\mathcal{F}^{-1}g = \mathcal{F}^2(\mathcal{F}g) = (\mathcal{F}g)(-\cdot).$$

This shows (3.3). Recall that $\int(\mathcal{F}f)g\,\mathrm{d}x = \int f(\mathcal{F}g)\,\mathrm{d}x$ for all $f, g \in \mathcal{S}(\mathbb{R}^n)$ (cf. (3.2)). Thus for $f, g \in \mathcal{S}(\mathbb{R}^n)$ we get

$$\int_{\mathbb{R}^n} (\mathcal{F}f)(x)\overline{(\mathcal{F}g)(x)}\,\mathrm{d}x = \int_{\mathbb{R}^n} (\mathcal{F}f)(x)(\mathcal{F}^{-1}\overline{g})(x)\,\mathrm{d}x = \int_{\mathbb{R}^n} f(x)\overline{g(x)}\,\mathrm{d}x.$$

This is (3.4). □

Remark 3.24. Formula (3.4) is called the *Plancherel formula.* It in particular implies that $\|\mathcal{F}f\|_2 = \|f\|_2$ for $f \in \mathcal{S}(\mathbb{R}^n)$. It was already remarked in Remark 3.18, that $\mathcal{S}(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$. So, $\mathcal{F}$ extends uniquely to an isometry $\mathcal{F}_2 : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$. This extension is called the **Fourier-Plancherel transform**. It follows easily from the surjectivity of $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ that $\mathcal{F}_2$ is bijective. Thus, $\mathcal{F}_2$ is unitary. In particular,

$$\langle \mathcal{F}_2 f, \mathcal{F}_2 g\rangle_{L^2} = \langle f, g\rangle_{L^2}, \qquad f, g \in L^2(\mathbb{R}^n).$$

It is important to note that the integral $\int_{\mathbb{R}^n} f(x)e^{-ixy}\,\mathrm{d}x$ only exists if and only if $f \in L^1(\mathbb{R}^n)$. Thus, in general, $\mathcal{F}_2 f$ cannot be expressed by the formula for the Fourier transform in Definition 3.12 for arbitrary $f \in L^2(\mathbb{R}^n)$. However, the following theorem shows in particular that this is the case for $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$.

Theorem 3.25. *The following holds for the Fourier-Plancherel transform $\mathcal{F}_2$.*

(a) *For $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, we have a.e.*

$$(\mathcal{F}_2 f)(y) = c\int_{\mathbb{R}^n} f(x)e^{-ixy}\,\mathrm{d}x =: (\mathcal{F}_1 f)(y).$$

(b) *For $f \in L^2(\mathbb{R}^n)$, we have*

$$(\mathcal{F}_2 f)(y) = \mathrm{L}^2\text{-}\lim_{R\to\infty} \mathcal{F}_1\left(\chi_{B_R(0)} \circ f\right).$$

Proof. (a). By Remark 3.5, there exists a sequence $(f_k) \subset C_c^\infty(\mathbb{R}^n)$ such that $\|f_k - f\|_1 \to 0$ and $\|f_k - f\|_2 \to 0$ as $k \to \infty$. Note that $\mathcal{F}f_k = \mathcal{F}_1 f_k = \mathcal{F}_2 f_k$ for each $k \in \mathbb{N}$. By Theorem 3.13,

$$\|\mathcal{F}_1 f - \mathcal{F}_1 f_k\|_\infty = \|\mathcal{F}_1(f - f_k)\|_\infty \le (2\pi)^{-n/2}\|f - f_k\|_1 \to 0.$$

Thus, for all $R > 0$,

$$\int_{B_R(0)} |\mathcal{F}_1 f - \mathcal{F} f_k|^2 \,\mathrm{d}x \to 0 \qquad \text{as } k \to \infty.$$

On the other hand, $\|\mathcal{F}f_k - \mathcal{F}_2 f\|_2 = \|f_k - f\|_2 \to 0$, hence

$$\int_{B_R(0)} |\mathcal{F}_2 f - \mathcal{F} f_k|^2 \,\mathrm{d}x \to 0 \qquad \text{as } k \to \infty$$

for all $R > 0$. This implies $\mathcal{F}_1 f|_{B_R(0)} = \mathcal{F}_2 f|_{B_R(0)}$ a.e. for all $R > 0$, that is, $\mathcal{F}_1 f = \mathcal{F}_2 f$ a.e. on $\mathbb{R}^n$.

(b). By (a), for each $R > 0$, we have $\mathcal{F}_1\left(\chi_{B_R(0)} f\right) = \mathcal{F}_2\left(\chi_{B_R(0)} f\right)$ a.e. on $\mathbb{R}^n$. Now,

$$\left\|\chi_{B_R(0)} f - f\right\|_2 = \left\|\chi_{\mathbb{R}^n \setminus B_R(0)} f\right\|_2 \to 0$$

as $R \to \infty$. Thus, $\mathcal{F}_1\left(\chi_{B_R(0)} f\right) \to \mathcal{F}_2 f$ in $L^2(\mathbb{R}^n)$. □

In the sequel, we write $\mathcal{F}_1 f = \mathcal{F} f = \mathcal{F}_2 f = \widehat{f}$. This is justified by Theorem 3.25

Lemma 3.26. *Let $f \in H^m(\mathbb{R}^n)$. Then, for all $\alpha \in \mathbb{N}_0^n$, $|a| \le m$, we have*

$$\mathcal{F}(\partial^\alpha f) = i^{|\alpha|} y^\alpha \mathcal{F} f.$$

Proof. Let $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Then[4]

$$\begin{aligned}
\langle \mathcal{F}(\partial^\alpha f), \mathcal{F}\varphi\rangle_{L^2} &= \langle \partial^\alpha f, \varphi\rangle = (-1)^{|\alpha|}\langle f, \partial^\alpha \varphi\rangle = (-1)^{|\alpha|}\langle \mathcal{F}f, \mathcal{F}(\partial^\alpha\varphi)\rangle \\
&= (-1)^{|\alpha|}\left\langle \mathcal{F}f, i^{|\alpha|} y^\alpha \mathcal{F}\varphi\right\rangle = \int_{\mathbb{R}^n}\left[i^{|\alpha|} y^\alpha (\mathcal{F}f)(y)\right]\overline{(\mathcal{F}\varphi)(y)}\,\mathrm{d}y.
\end{aligned}$$

Thus, we have $h := \mathcal{F}(\partial^\alpha f) - i^{|\alpha|} y^\alpha \mathcal{F} f \in L^2_{\mathrm{loc}}(\mathbb{R}^n) \subset L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and $\int_{\mathbb{R}^n} h\psi \,\mathrm{d}x = 0$ for all $\psi \in C_c^\infty(\mathbb{R}^n)$. Therefore, $h = 0$ by Lemma 3.6. □

Theorem 3.27 (Sobolev's Lemma). *Let $\Omega \subset \mathbb{R}^n$ be open, $m, k \in \mathbb{N}_0$, $m > k + \frac{n}{2}$, and $f \in H^m(\Omega)$. Then there exists $g \in C^k(\Omega)$ such that $f = g$ a.e. on Ω.*

[4] The proof of the second equality is an exercise.

Proof. We assume first that $\Omega = \mathbb{R}^n$. We shall make use of the following implication which will be proved in an exercise

$$g \in L^1(\mathbb{R}^n) \text{ and } x^\alpha g \in L^1(\mathbb{R}^n)\,\forall\,|\alpha| \le k \implies \mathcal{F}g \in C^k(\mathbb{R}^n). \tag{3.5}$$

Let $f \in H^m(\mathbb{R}^n)$, $m > k + \frac{n}{2}$. By Lemma 3.26, $y^\alpha \mathcal{F}f \in L^2(\mathbb{R}^n)$ for all $\alpha \in \mathbb{N}_0^n$ with $|\alpha| \le m$. In particular, $y_j^{2m}|\mathcal{F}f|^2 \in L^1(\mathbb{R}^n)$ for each coordinate y_j of y. Setting $y_0 = 1$, we get

$$1 + y^2 = \sum_{j=0}^n y_j^2 \le \left(\sum_{j=0}^n 1\right)^{1/p} \left(\sum_{j=0}^n y_j^{2m}\right)^{1/m} = (n+1)^{1/p}\left(1 + \sum_{j=1}^n y_j^{2m}\right)^{1/m},$$

where $\frac{1}{p} + \frac{1}{m} = 1$. Thus,

$$\left(1+y^2\right)^m |\mathcal{F}f|^2 \in L^1(\mathbb{R}^n). \tag{3.6}$$

We have moreover

$$|y|^{2|\alpha|} = (y_1^2 + \ldots + y_n^2)^{\alpha_1+\ldots+\alpha_n} \ge y_1^{2\alpha_1} \cdot \ldots \cdot y_n^{2\alpha_n} = (y^\alpha)^2.$$

Thus, $|y^\alpha| \le |y|^{|\alpha|}$. This and Cauchy-Schwarz imply for $|\alpha| \le k$

$$\begin{aligned}\int_{\mathbb{R}^n} |y^\alpha(\mathcal{F}f)(y)|\,\mathrm{d}y &\le \int_{\mathbb{R}^n} |y|^{|\alpha|}\,|(\mathcal{F}f)(y)|\,\mathrm{d}y \le \int_{\mathbb{R}^n} \left(1+y^2\right)^{\frac{|\alpha|}{2}} |(\mathcal{F}f)(y)|\,\mathrm{d}y \\ &\le \int_{\mathbb{R}^n} \left(1+y^2\right)^{\frac{m}{2}} |(\mathcal{F}f)(y)| \left(1+y^2\right)^{\frac{k-m}{2}} \mathrm{d}y \\ &\le \left(\int_{\mathbb{R}^n} \left(1+y^2\right)^m |(\mathcal{F}f)(y)|^2\,\mathrm{d}y\right)^{1/2} \left(\int_{\mathbb{R}^n} \left(1+y^2\right)^{k-m} \mathrm{d}y\right)^{1/2}.\end{aligned}$$

Note that the first integral exists due to (3.6). Setting $w_{n-1} := \mathrm{Vol}_{n-1}(S^{n-1})$, we see that also the second integral exists:

$$\begin{aligned}\int_{\mathbb{R}^n} \left(1+y^2\right)^{k-m} \mathrm{d}y &= w_{n-1} \cdot \int_0^\infty (1+r^2)^{k-m} r^{n-1}\,\mathrm{d}r = w_{n-1} \int_0^\infty \frac{r^{n-1}}{(1+r^2)^{m-k}}\,\mathrm{d}r \\ &\le w_{n-1}\left(\int_0^1 1\,\mathrm{d}r + \int_1^\infty r^{n-1-2m+2k}\,\mathrm{d}r\right) < \infty.\end{aligned}$$

Thus, $y^\alpha \mathcal{F}f \in L^1(\mathbb{R}^n)$ for all multiindices α with $|\alpha| \le k$. In particular, $\mathcal{F}f \in L^1(\mathbb{R}^n)$. Thus we get by (3.5) $\mathcal{F}_1\mathcal{F}_2 f \in C^k(\mathbb{R}^n)$.

Define $U : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$, $g \mapsto g(-\cdot)$. Then $\|Ug\|_2^2 = \int |g(-x)|^2\,\mathrm{d}x = \|g\|_2^2$, i.e., U is unitary. By Lemma 3.19, we have $\mathcal{F}^2|_{\mathcal{S}(\mathbb{R}^n)} = U|_{\mathcal{S}(\mathbb{R}^n)}$, and due to the density of $\mathcal{S}(\mathbb{R}^n)$ in $L^2(\mathbb{R}^n)$, we get $\mathcal{F}^2 = U$. That is, we have $f(-\cdot) = \mathcal{F}^2 f$ a.e. on $\mathbb{R}^n$. Hence, f coincides a.e. with the C^k-function $(\mathcal{F}_1\mathcal{F}_2 f)(-\cdot)$.

Now, let $\Omega \subset \mathbb{R}^n$ be an arbitrary open set and $f \in H^m(\Omega)$. Let $K \subset \Omega$ be compact, and choose $\varphi \in C_c^\infty(\Omega)$ with $\varphi|_K = 1$. Put

$$\widetilde{f}(x) := \begin{cases} \varphi(x)f(x), & \text{for } x \in \Omega, \\ 0, & \text{for } x \in \mathbb{R}^n \setminus \Omega. \end{cases}$$

In an exercise, we show that $\varphi f \in H^m(\Omega)$ and $\widetilde{f} \in H^m(\mathbb{R}^n)$. By the first part, there exists $h \in C^k(\mathbb{R}^n)$ such that $\widetilde{f} = h$ a.e. on $\mathbb{R}^n$. Then $f|_K = \varphi f|_K = \widetilde{f}|_K = h|_K$ a.e. on K.

Now, choose an increasing sequence $(K_j) \subset \Omega$ of compact sets, such that $\Omega = \bigcup_{j=1}^\infty K_j$. Then, for all $j \in \mathbb{N}$ there exists a function $h_j \in C^k(\mathbb{R}^n)$ such that $f|_{K_j} = h_j|_{K_j}$ a.e. on K_j. For $i, j \in \mathbb{N}$ with $i < j$, we have $h_j|_{K_i} = \left(h_j|_{K_j}\right)|_{K_i} = \left(f|_{K_j}\right)|_{K_i} = f|_{K_i} = h_i|_{K_i}$. Therefore, the function $h : \Omega \to \mathbb{C}$, defined by $h(x) := h_j(x)$ if $x \in K_j$, is well-defined, contained in $C^k(\Omega)$, and $h = f$ a.e. on Ω. □

In the theory of differential equations, one usually has to deal with boundary conditions: in order to obtain a unique solution, it should not only satisfy the differential equation, but also conditions on its values at the boundary of its domain of definition Ω. These boundary conditions are typically integrated into the domain $\operatorname{dom} D$ of the differential operator D. That is, one considers $Df = g$, $f \in \operatorname{dom} D$, where $\operatorname{dom} D$ is the subspace of some Sobolev space. Often, one imposes so-called *Dirichlet boundary conditions* – meaning that the solution should vanish at the boundary of Ω. Therefore, the space $H_0^m(\Omega)$ plays a special role[5].

Theorem 3.28 (Rellich's Embedding Theorem). *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set and $m \in \mathbb{N}$. Then the operator $T : H_0^m(\Omega) \to H_0^{m-1}(\Omega)$, $f \mapsto f$, is compact.*

Proof. We shall first prove the theorem for $m = 1$. Then $T : H_0^1(\Omega) \to L^2(\Omega)$, since $H^0(\Omega) = L^2(\Omega)$, and thus

$$H_0^0(\Omega) = \overline{C_c^\infty(\Omega)}^{L^2(\Omega)} = L^2(\Omega).$$

Let $(f_k) \subset H_0^1(\Omega)$ be bounded, i.e. $\|f_k\|_{H^1} \leq K$ for all $k \in \mathbb{N}$ and some $K > 0$. We have to show, that there exists a subsequence (f_{k_j}) which converges in $L^2(\Omega)$. Therefore, it is obviously no restriction to assume, that $f_k \in C_c^\infty(\Omega)$ for each $k \in \mathbb{N}$.

For $f : \Omega \to \mathbb{C}$ let $\widetilde{f} : \mathbb{R}^n \to \mathbb{C}$ denote the trivial extension, i.e., $\widetilde{f}|_{\mathbb{R}^n \setminus \Omega} = 0$. Then $\widetilde{f}_k \in \mathbb{C}_c^\infty(\mathbb{R}^n)$. Since (f_k) is bounded in $L^2(\Omega)$, there exists a subsequence (f_{k_j}) such that f_{k_j} weakly converges to some f in $L^2(\Omega)$. Without loss of generality, we can assume that $f_k \to f$ weakly in $L^2(\Omega)$. We prove that (f_k) is a Cauchy sequence in $L^2(\Omega)$. For this, we observe that

$$\|f_k - f_j\|_{L^2(\Omega)}^2 = \|\widetilde{f}_k - \widetilde{f}_j\|_{L^2(\mathbb{R}^n)}^2 = \|\mathcal{F}\widetilde{f}_k - \mathcal{F}\widetilde{f}_j\|_{L^2(\mathbb{R}^n)}^2 = \int_{\mathbb{R}^n} |\mathcal{F}\widetilde{f}_k - \mathcal{F}\widetilde{f}_j|^2 \, \mathrm{d}y.$$

For $l = 1, \ldots, n$ and $g \in H^1(\mathbb{R}^n)$ we have

$$\|y_l \mathcal{F} g\|_2 = \|\mathcal{F}(\partial_l g)\|_2 = \|\partial_l g\|_2 \leq \|g\|_{H^1}.$$

[5] For "nice boundaries" $\partial\Omega$, one can define a *boundary operator* $B : H^m(\Omega) \to L^2(\partial\Omega)$ such that $Bf = f|_{\partial\Omega}$ for $f \in C^\infty(\overline{\Omega})$. It should be mentioned that $\ker B$ does not always coincide with $H_0^m(\Omega)$.

Hence, for each $R > 0$, we have

$$\int_{|y|<R} \left|\mathcal{F}\widetilde{f}_k - \mathcal{F}\widetilde{f}_j\right|^2 \mathrm{d}y \le \frac{1}{R^2} \int_{|y|>R} |y|^2 \left|\mathcal{F}\widetilde{f}_k - \mathcal{F}\widetilde{f}_j\right|^2 \mathrm{d}y$$
$$\le \frac{1}{R^2} \sum_{l=1}^{n} \int_{\mathbb{R}^n} y_l^2 \left|\mathcal{F}\left(\widetilde{f}_k - \widetilde{f}_j\right)\right|^2 \mathrm{d}y \le \frac{4nK^2}{R^2}.$$

For the integral over $B_R(0)$ we note that

$$(\mathcal{F}\widetilde{f}_k)(y) = c \int_{\Omega} f_k(x) e^{-ixy} \, \mathrm{d}x = c \langle f_k, e_y \rangle ,$$

where $e_y(x) := e^{-ixy}$, $x \in \Omega$, $y \in \mathbb{R}^n$. Thus, for each $y \in \mathbb{R}^n$, we have

$$(\mathcal{F}\widetilde{f}_k)(y) \to c \langle f, e_y \rangle = (\mathcal{F}\widetilde{f})(y) \quad \text{as } k \to \infty.$$

Moreover,

$$\left\|\mathcal{F}\left(\widetilde{f}_k - \widetilde{f}\right)\right\|_\infty \le c \left\|\widetilde{f}_k - \widetilde{f}\right\|_1 = c \, \|f_k - f\|_{L^1(\Omega)}$$
$$\le c \operatorname{Vol}(\Omega)^{\frac{1}{2}} \|f_k - f\|_{L^2(\Omega)} \le c \operatorname{Vol}(\Omega)^{\frac{1}{2}} (K + \|f\|_{L^2(\Omega)}).$$

Now, Lebesgue's theorem implies that $\int_{|y|\le R} |\mathcal{F}\widetilde{f}_k - \mathcal{F}\widetilde{f}|^2 \, \mathrm{d}y \to 0$ as $k \to \infty$. Thus, $\int_{|y|\le R} |\mathcal{F}\widetilde{f}_k - \mathcal{F}\widetilde{f}_j|^2 \to 0$ as $k, j \to \infty$ and thus (f_k) is a Cauchy sequence in $L^2(\Omega)$.

Let now $m \ge 2$. Let $(f_n) \subset H_0^m(\Omega)$ be a bounded sequence. Then for each $\alpha \in \mathbb{N}_0^n$ with $|\alpha| \le m-1$ we have that $(\partial^\alpha f_k) \subset H_0^1(\Omega)$ is bounded in $H_0^1(\Omega)$. Thus, from the first part of the proof it follows that there exist a subsequence (f_{k_j}) of (f_k) and $g_\alpha \in L^2(\Omega)$ such that $\|\partial^\alpha f_{k_j} - g_\alpha\|_2 \to 0$ as $j \to \infty$. The sequence (f_{k_j}) is therefore a Cauchy sequence in $H_0^{m-1}(\Omega)$. As $H_0^{m-1}(\Omega)$ is complete, (f_{k_j}) converges in $H_0^{m-1}(\Omega)$. □

Remark 3.29. The proof of Theorem 3.24 reveals that the theorem still holds for $H^{m-1}(\Omega)$ instead of $H_0^{m-1}(\Omega)$ if functions from $H^{m-1}(\Omega)$ can be extended to $H^{m-1}(\mathbb{R}^n)$-functions. This is possible, e.g., if $\partial\Omega$ is sufficiently smooth.

We close this chapter with a theorem which shows that Sobolev spaces on $\mathbb{R}^n$ can be characterized with the help of the Fourier transform.

Theorem 3.30. *For $m \in \mathbb{N}_0$ we have*

$$H^m(\mathbb{R}^n) = \{f \in L^2(\mathbb{R}^n) : (1 + y^2)^{m/2} \mathcal{F}f \in L^2(\mathbb{R}^n)\}.$$

Proof. Note that $(1+y^2)^{m/2}\mathcal{F}f \in L^2(\mathbb{R}^n)$ is equivalent to $y^\alpha \mathcal{F}f \in L^2(\mathbb{R}^n)$ for all $|\alpha| \le m$. If $f \in H^m(\mathbb{R}^n)$, then the latter follows directly from Lemma 3.26. Let $y^\alpha \mathcal{F}f \in L^2(\mathbb{R}^n)$ for all $|\alpha| \le m$, and put $f_\alpha := \mathcal{F}^{-1}(i^{|\alpha|} y^\alpha \mathcal{F}f) \in L^2(\mathbb{R}^n)$. Then for each $\varphi \in C_c^\infty(\mathbb{R}^n)$ and each $|\alpha| \le m$, we have

$$\langle f_\alpha, \varphi \rangle = \langle \mathcal{F}^{-1}(i^{|\alpha|} y^\alpha \mathcal{F}f), \varphi \rangle = \langle i^{|\alpha|} y^\alpha \mathcal{F}f, \mathcal{F}\varphi \rangle = (-1)^{|\alpha|} \langle \mathcal{F}f, i^{|\alpha|} y^\alpha \mathcal{F}\varphi \rangle$$
$$= (-1)^{|\alpha|} \langle \mathcal{F}f, \mathcal{F}(\partial^\alpha \varphi) \rangle = (-1)^{|\alpha|} \langle f, \partial^\alpha \varphi \rangle.$$

This proves that $\partial^\alpha f = f_\alpha$ exists and is an element of $L^2(\mathbb{R}^n)$. □

Remark 3.31. The above representation allows the definition of Sobolev spaces of fractional or even real order ($s \in \mathbb{R}$):

$$H^s(\mathbb{R}^n) := \{f \in L^2(\mathbb{R}^n) : (1 + y^2)^{s/2}\mathcal{F}f \in L^2(\mathbb{R}^n)\}.$$

4. Topological Vector Spaces

In Functional Analysis I (cf. [5]), we have introduced and studied the weak topology on normed spaces. When we forget about that there is a norm on the space, we are merely left with a topological space with a linear structure upon it. However, there is a certain connection between topological and linear structure: vector addition and scalar multiplication are continuous operations (with respect to the weak topology). In this chapter, we release the topology from its connection to a norm and generalize the above situation by introducing the so-called topological vector spaces. To get started, we recall some basic notions and facts from topology.

4.1. Topological Basics

Let X be a set. By $\mathfrak{P}(X)$, we denote the **power set** of X, i.e., the set of all subsets of X. For $\gamma \subset \mathfrak{P}(X)$ define

$$\mathfrak{I}(\gamma) := \left\{ \bigcap_{i=1}^{n} \gamma_i : \gamma_i \in \gamma, n \in \mathbb{N} \right\} \qquad \text{(finite intersections of sets in } \gamma\text{)},$$

and

$$\mathfrak{U}(\gamma) := \left\{ \bigcup_{B \in \mathcal{B}} B : \mathcal{B} \subset \gamma \right\} \qquad \text{(arbitrary unions of sets in } \gamma\text{)}.$$

Moreover, we set $\mathcal{O}(\gamma) := \mathfrak{U}(\mathfrak{I}(\gamma))$.

Definition 4.1. Let X be a set. A set system $\mathcal{O} \subset \mathfrak{P}(X)$ is called a **topology** on X if

(i) $\varnothing, X \in \mathcal{O}$,

(ii) $\mathfrak{I}(\mathcal{O}) \subset \mathcal{O}$ and

(iii) $\mathfrak{U}(\mathcal{O}) \subset \mathcal{O}$.

The pair $(X, \mathcal{O})$ is then called a **topological space**. The sets in $\mathcal{O}$ are called **open**. A set $A \subset X$ is called **closed**, if $X \setminus A \in \mathcal{O}$. The **closure** and the **interior** of $A \subset X$ are defined by

$$\overline{A} := \bigcap \{C \subset X \text{ closed} : A \subset C\} \quad \text{and} \quad \overset{\circ}{A} := \operatorname{int} A := \bigcup \{U \subset X \text{ open} : U \subset A\},$$

respectively. A set $A \subset X$ is called a **neighborhood** of $x \in X$ if there exists a set $U \in \mathcal{O}$ such that $x \in U$ and $U \subset A$.

Example 4.2. Simple examples of topological spaces are $(X, \mathcal{P}(X))$ and $(X, \{\varnothing, X\})$. The first topology is called the *discrete topology*, the second the *trivial topology*.

The following characterizations of the interior and the closure of a set will be frequently used in the sequel.

Lemma 4.3. *Let $(X, \mathcal{O})$ be a topological space, $B \subset X$, and $x \in X$. Then the following statements are equivalent.*

(i) $x \in \overline{B}$.

(ii) *For each neighborhood U of x we have $B \cap U \neq \varnothing$.*

In addition, also the next two conditions are equivalent.

(iii) $x \in \operatorname{int} B$.

(iv) *There exists a neighborhood U of x such that $U \subset B$.*

Definition 4.4. Let $\mathcal{O}$ be a topology on X. A system $B \subset \mathcal{O}$ is called a **basis** of $\mathcal{O}$ if $\mathfrak{U}(\mathcal{B}) = \mathcal{O}$. A system $\gamma \subset \mathcal{O}$ is called a **subbasis** of $\mathcal{O}$ if $\mathfrak{U}(\mathfrak{I}(\gamma)) = \mathcal{O}$, i.e. if $\mathfrak{I}(\gamma)$ is a basis of $\mathcal{O}$.

Example 4.5. Open balls $\{B_r(x) : r > 0,\, x \in X\}$ are a basis of the topology of a metric space (like $\mathbb{R}^n$).

Proposition 4.6. *Let X be a set, $\gamma \subset \mathfrak{P}(X)$ with $\varnothing, X \in \gamma$. Then $\mathcal{O}(\gamma) = \mathfrak{U}(\mathfrak{I}(\gamma))$ is a topology on X with basis $\mathfrak{I}(\gamma)$ and subbasis γ.*

Definition 4.7. Let $(X_1, \mathcal{O}_1)$ and $(X_2, \mathcal{O}_2)$ be topological spaces and $f : X_1 \to X_2$. f is called **continuous** at $x \in X_1$ if for every neighborhood V of $f(x)$ there exists a neighborhood U of x such that $f(U) \subset V$.

Lemma 4.8. *A function $f : X_1 \to X_2$ is continuous (at each $x \in X_1$) if and only if $f^{-1}(V) \in \mathcal{O}_1$ for all $V \in \mathcal{O}_2$.*

Lemma 4.9. *Let $(\mathcal{O}_i)_{i \in I}$ be a family of topologies on X. Then $\bigcap_{i \in I} \mathcal{O}_i$ is a topology on X.*

Note that the union of topologies is in general not a topology.

Definition 4.10. Let X be a set, $(X_i, \mathcal{O}_i)_{i \in I}$ a family of topological spaces, and $f_i : X \to X_i$, $i \in I$. The **initial topology** on X with respect to $\{f_i\}_{i \in I}$ is defined by

$$\bigcap \{\mathcal{O} \subset \mathfrak{P}(X) : \mathcal{O} \text{ topology on } X,\, f_i : (X, \mathcal{O}) \to (X_i, \mathcal{O}_i) \text{ continuous } \forall i \in I\}\,.$$

It is the smallest (coarsest) topology with respect to which all f_i are continuous.

Proposition 4.11. *Let X, $(X_i, \mathcal{O}_i)_{i \in I}$ and $\{f_i\}_{i \in I}$ be as in Definition 4.10. Then the initial topology with respect to $\{f_i\}_{i \in I}$ is given by $\mathcal{O}(\gamma)$, where*

$$\gamma = \left\{f_i^{-1}(V) : V \in \mathcal{O}_i,\, i \in I\right\}.$$

Example 4.12. (a) Let $X \subset Y$, $(Y, \mathcal{O})$ a topological space and $\nu : X \to Y, x \mapsto x$. Then $\mathcal{O}_X := \{\nu^{-1}(V) : V \in \mathcal{O}\} = \{X \cap V : V \in \mathcal{O}\}$ is called the **trace topology** (or **relative** or **induced** topology). For example $Y = \mathbb{R}$, $X = [0, 1]$. In this case $[0, 1]$ is open in the trace topology.
(b) Let $(X_i, \mathcal{O}_i)_{i \in I}$ be topological spaces and

$$X := \prod_{i \in I} X_i := \left\{ x : I \to \bigcup_{i \in I} X_i \;\middle|\; x(i) \in X_i \right\}$$

(the cartesian product). Set $\pi_i : X \to X_i$, $\pi_i(x) := x(i)$ (the canonical projection). The **product topology** on X is the initial topology with respect to $\{\pi_i\}_{i \in I}$.

Lemma 4.13. *Let X be a set, $(X_i, \mathcal{O}_i)_{i \in I}$ topological spaces and $f_i : X_i \to X$, $i \in I$. Then $\mathcal{O} := \left\{ V \subset X : f_i^{-1}(V) \in \mathcal{O}_i \, \forall i \in I \right\}$ is a topology on X.*

The topology in Lemma 4.13 is the largest (finest) topology with respect to which each f_i is continous. It is called the **final topology** with respect to $\{f_i\}_{i \in I}$. A special case is the following.

Example 4.14. Let $(Y, \mathcal{O})$ be a topological space, $\sim$ an equivalence relation on Y and $\pi : Y \to Y/\sim$, $y \mapsto [y]_\sim$. The finial topology on $Y/\sim$ with respect to π is called the **quotient topology**.

Definition 4.15. Let $(X, \mathcal{O})$ be a topological space.

(i) $(X, \mathcal{O})$ is called **Hausdorff** if for all $x, y \in X$, $x \neq y$, there exist two sets $U_x, U_y \in \mathcal{O}$ such that $x \in U_x$, $y \in U_y$ and $U_x \cap U_y = \varnothing$.

(ii) A set system $\mathcal{U} \subset \mathfrak{P}(X)$ is called a **cover** of X if $X = \bigcup_{U \in \mathcal{U}} U$. A cover $\mathcal{U}$ is called **open** if $\mathcal{U} \subset \mathcal{O}$. A **subcover** of the cover $\mathcal{U}$ is a subset of $\mathcal{U}$ that is still a cover of X.

(iii) $(X, \mathcal{O})$ is called **compact** if every open cover of X has a finite subcover. A set $A \subset X$ is called **compact** if it is compact with respect to its trace topology.

Lemma 4.16. *Let $(X, \mathcal{O})$ be a topological space.*

(i) *If $(X, \mathcal{O})$ is compact and $A \subset X$ closed, then A is compact.*

(ii) *If $(X, \mathcal{O})$ is Hausdorff and $A \subset X$ compact, then A is closed.*

Proof. (i) is easy to see. For (ii), let $x \in X \setminus A$. For all $y \in A$ there exist $U_y, V_y \in \mathcal{O}$ such that $y \in U_y$ and $x \in V_y$ with $U_y \cap V_y = \varnothing$. Then $A \subset \bigcup_{y \in A} U_y$. Thus, there exists a finite number of elements $y_1, \ldots, y_n \in A$ such that $A \subset \bigcup_{i=1}^n U_{y_i}$. Define $V_x := \bigcap_{i=1}^n V_{y_i}$. Obviously, we have $x \in V_x \in \mathcal{O}$ and

$$V_x \subset \bigcap_{i=1}^n U_{y_j}^c = \left(\bigcup_{i=1}^n U_{y_i} \right)^c \subset X \setminus A.$$

This shows that $X \setminus A$ is open. □

Example 4.17. An example for the necessity of the Hausdorff property in Lemma 4.16 (ii) is the topological space $X = \{a, b\}$ with topology $\{\{a, b\}, \varnothing, \{a\}\}$. Here, $\{a\}$ is compact but not closed.

Definition 4.18. A relation $\leq$ on a set P is called

(i) **transitive** if $a \leq b$ and $b \leq c$ imply $a \leq c$ for all $a, b, c \in P$,

(ii) **reflexive** if $a \leq a$ for all $a \in P$,

(iii) **antisymmetric** if $a \leq b$ and $b \leq a$ imply that $a = b$ for all $a, b \in P$.

A transitive, reflexive, and antisymmetric relation on P is called **partial order**, and the set is called **partially ordered**. A partial order on P is called a **total order** if for all $a, b \in P$ at least $a \leq b$ or $b \leq a$ holds. In this case, P is called **totally ordered**.

Theorem 4.19 (Hausdorff Maximal Principle). *Let $(P, \leq)$ be partially ordered. Then there exists a maximal totally ordered $Q \subset P$, i.e., if $Q' \subset P$ is totally ordered with $Q \subset Q'$, then $Q' = Q$.*

Remark 4.20. The Hausdorff Maximal Principle is equivalent to Zorn's Lemma and the axiom of choice.

4.2. Topological Vector Spaces

Definition 4.21. Let X be a vector space over $\mathbb{K}$ and $\mathcal{O}$ a topology on X. The pair $(X, \mathcal{O})$ is called a **topological vector space** (TVS) if

(i) $+ : X \times X \to X, (x, y) \mapsto x + y$, is continuous with respect to the product topology and the given topology,

(ii) $\cdot : \mathbb{K} \times X \to X, (\lambda, x) \mapsto \lambda x$ is continuous,

(iii) $(X, \mathcal{O})$ is Hausdorff.

Example 4.22. (a) A normed space is a TVS with the norm topology.
(b) A normed space with the weak topology is a TVS. Here, (iii) follows from the Hahn-Banach theorem.

Remark 4.23. Let $(X, \mathcal{O})$ be a TVS. Then the following holds.

(i) If $x_1 + x_2 \in V \in \mathcal{O}$, then there exist $U_1, U_2 \in \mathcal{O}$ with $x_1 \in U_1$, $x_2 \in U_2$ and $U_1 + U_2 \subset V$.

(ii) If $\lambda x \in V \in \mathcal{O}$ then there exist $\varepsilon > 0$ and $U \in \mathcal{O}$ such that $x \in U$ and for all $\mu \in \mathbb{K}$ with $|\mu - \lambda| < \varepsilon$ we have $\mu U \subset V$.

Lemma 4.24. *Let $(X, \mathcal{O})$ be a TVS, $a \in X$, and $\lambda \in \mathbb{K} \setminus \{0\}$. Then $T_a, M_\lambda : X \to X$, where $T_a x := x + a$ and $M_\lambda x := \lambda x$, $x \in X$, are homeomorphisms.*

Proof. T_a and M_λ are continuous by Definition 4.21. Moreover, $T_a^{-1} = T_{-a}$ and $M_\lambda^{-1} = M_{\lambda^{-1}}$ are also continuous. □

Since $a + U = T_a(U)$ and $\lambda U = M_\lambda(U)$, the next corollary follows immediately.

Corollary 4.25. *Let $(X, \mathcal{O})$ be a TVS. Then the following holds.*

(a) *If $U \in \mathcal{O}$, $x \in X$, and $\lambda \in \mathbb{K} \setminus \{0\}$, then $x + U \in \mathcal{O}$ and $\lambda U \in \mathcal{O}$.*

(b) *If $x \in V \in \mathcal{O}$, then $V = x + U$ where $U \in \mathcal{O}$ with $0 \in U$.*

Definition 4.26. Let X be a vector space over $\mathbb{K}$. A set $A \subset X$ is called

(a) **absorbing** if for all $x \in X$ there exists $t > 0$ such that $tx \in A$;

(b) **symmetric** if $A = -A$;

(c) **balanced**[6] if $\lambda A \subset A$ for all $\lambda \in \mathbb{K}$ with $|\lambda| \leq 1$;

(d) **convex** if for all $x, y \in A$ and all $t \in (0, 1)$ we have $tx + (1 - t)y \in A$.

Remark 4.27. Note that a balanced set always contains the zero vector. Moreover, if A is balanced and $|\lambda| = 1$, then $\lambda A \subset A$ and $\lambda^{-1} A \subset A$ so that $A = \lambda A$. In particular, a balanced set is always symmetric.

Lemma 4.28. *Let $(X, \mathcal{O})$ be a TVS and $U \in \mathcal{O}$ with $0 \in U$. Then the following statements hold.*

(i) *U is absorbing.*

(ii) *For each $n \in \mathbb{N}$ there exists a symmetric set $V \in \mathcal{O}$ with $0 \in V$ such that*

$$\underbrace{V + \ldots + V}_{n \text{ times}} \subset U.$$

(iii) *There exists a balanced set $V \in \mathcal{O}$ with $0 \in V$ such that $V \subset U$.*

(iv) *If U is convex, then there exists a convex balanced set $V \in \mathcal{O}$ with $0 \in V$ such that $V \subset U$.*

Proof. (i). Let $x \in X$. Define $\varphi : \mathbb{K} \to X, z \mapsto zx$. Then $\varphi^{-1}(U) \subset \mathbb{K}$ is an open set and $0 \in \varphi^{-1}(U)$. Thus there exists a $t > 0$ with $t \in \varphi^{-1}(U)$, meaning that $tx \in U$.

(ii). By continuity of addition, there exist $V_1, V_2 \in \mathcal{O}$ with $0 \in V_1 + V_2$ and $V_1 + V_2 \subset U$. Now put

$$V := V_1 \cap V_2 \cap (-V_1) \cap (-V_2).$$

Then $V \in \mathcal{O}$, $V = -V$, $0 \in V$, and $V + V \subset U$. This proves the claim for $n = 2$. Since $0 \in V \in \mathcal{O}$, there exists a symmetric $\widetilde{V} \in \mathcal{O}$ with $0 \in \widetilde{V}$ with $\widetilde{V} + \widetilde{V} \subset V$. Thus,

$$\widetilde{V} + \widetilde{V} + \widetilde{V} \subset \widetilde{V} + \widetilde{V} + \widetilde{V} + \widetilde{V} \subset V + V \subset U.$$

[6] In german: kreisförmig

The rest follows by induction.

(iii). By continuity of scalar multiplication, there exist $\delta > 0$ and $W \in \mathcal{O}$ with $0 \in W$ such that $zW \subset U$ for all $z \in \mathbb{K}$ with $|z| < \delta$. Put $V := \bigcup_{|z|<\delta} zW \subset U$. Then $V \in \mathcal{O}$ and $0 \in V$. If $\lambda \in \mathbb{K}$ with $|\lambda| \leq 1$, then

$$\lambda V = \{\lambda zw : |z| < \delta,\, w \in W\} = \{\xi w : |\xi| < |\lambda|\delta,\, w \in W\} = \bigcup_{|\xi|<|\lambda|\delta} \xi W \subset V.$$

This shows that V is balanced.

(iv). By (iii), there exists a balanced set $W \in \mathcal{O}$ with $0 \in W$ and $W \subset U$. Put $A := \bigcap_{|\alpha|=1} \alpha U$. Since $\alpha W \subset W$ for $|\alpha| = 1$ we have that $W \subset \alpha U$ and thus $W \subset A$. Thus, $0 \in A$, and as an intersection of convex sets, A is convex. To show that A is balanced, choose $\lambda \in \mathbb{K}$ with $|\lambda| \leq 1$. We can write $\lambda = r\beta$ with $r \in [0, 1]$ and $\beta \in \mathbb{K}$ with $|\beta| = 1$. Since U is convex with $0 \in U$, we have $rU \subset U$. Therefore,

$$\lambda A = r\beta A = \bigcap_{|\alpha|=1} r\beta\alpha U = \bigcap_{|\alpha|=1} r\alpha U \subset \bigcap_{|\alpha|=1} \alpha U = A.$$

Put $V := \operatorname{int} A$. V is convex since for $x, y \in V$ and arbitrary $t \in (0, 1)$ we have $tx + (1-t)y \in tx + (1-t)V \subset A$ (cf. Lemma 4.3). Further, for $\alpha \in \mathbb{K}$ with $0 < |\alpha| \leq 1$ we have that $\alpha V \subset \alpha A \subset A$. This and Lemma 4.3 imply $\alpha V \subset V$. Moreover, $0 \in V$ since $0 \in W \subset A$. Finally, $V \subset U$ follows from $A = \bigcap_{|\alpha|=1} U \subset U$. □

Theorem 4.29. *Let $(X, \mathcal{O})$ be a TVS, $K \subset X$ compact, and $C \subset X$ closed with $K \cap C = \varnothing$. Then there exists a set $V \in \mathcal{O}$ with $0 \in V$ such that $\overline{K + V} \cap (C + V) = \varnothing$.*

Proof. If $K = \varnothing$, then $K + V = \varnothing$ for all $V \in \mathcal{O}$, and nothing is to be proved. Therefore, we assume that $K \neq \varnothing$. Let $x \in K$. By Lemma 4.28, there exists a symmetric set $V_x \in \mathcal{O}$ with $0 \in V_x$ such that $V_x + V_x + V_x \subset -x + (X \setminus C)$. Obviously, $x + V_x + V_x + V_x \subset X \setminus C$, and thus

$$(x + V_x + V_x) \cap (C + V_x) = \varnothing.$$

As K is compact there exist $x_1, \ldots, x_n \in K$ with

$$K \subset \bigcup_{i=1}^{n} (x_i + V_{x_i}).$$

Put $V := \bigcap_{i=1}^{n} V_{x_i}$. Then $0 \in V$ and

$$K + V \subset \left(\bigcup_{i=1}^{n} (x_i + V_{x_i}) \right) + V = \bigcup_{i=1}^{n} (x_i + V_{x_i} + V) \subset \bigcup_{i=1}^{n} (x_i + V_{x_i} + V_{x_i}).$$

Hence $(K + V) \cap (C + V) = \varnothing$, since $C + V = \bigcap_{i=1}^{n} (C + V_{x_i})$.

If $x \in C + V$ then $C + V = \bigcup_{y \in C} (y + V)$ is an open neighborhood of x which has an empty intersection with $K + V$. By Lemma 4.3, we have $x \notin \overline{K + V}$. Hence, $\overline{K + V} \cap (C + V) = \varnothing$. □

Corollary 4.30. *For all $U \in \mathcal{O}$ with $0 \in U$ there exists $V \in \mathcal{O}$ with $0 \in V$ such that $\overline{V} \subset U$.*

Proof. Define $K := \{0\}$ and $C := X \setminus U$. Then there exists $V \in \mathcal{O}$ with $0 \in V$ such that $\overline{K+V} \cap (C+V) = \varnothing$, i.e. $\overline{V} \cap ((X \setminus U)+V) = \varnothing$. Let $x \in \overline{V}$. Then $x \notin (X \setminus U)+V$ and thus $x \notin X \setminus U$, i.e., $x \in U$. □

Lemma 4.31. *Let $(X, \mathcal{O}_X)$ and $(Y, \mathcal{O}_Y)$ be TVS's and $T : X \to Y$ a linear map. Then T is continuous if and only if T is continuous at 0.*

Proof. Assume that T is continuous at zero, and let $x_0 \in X$. Define mappings $T_X : X \to X$ and $T_Y : Y \to Y$ by $T_X(x) := x - x_0$, $x \in X$, and $T_Y(y) := y + Tx_0$, $y \in Y$. For $x \in X$, we have

$$T_Y T T_X(x) = T_Y T(x - x_0) = T_Y(Tx - Tx_0) = Tx - Tx_0 + Tx_0 = Tx.$$

Thus, $T = T_Y T T_X$. Now, T_X and T_Y are continuous and $T_X(x_0) = 0$. This implies that T is continous at x_0. □

Definition 4.32. Let $(X, \mathcal{O})$ be a TVS. Then

$$(X, \mathcal{O})' := \{f : X \to \mathbb{K} \,|\, f \text{ linear and continuous}\}$$

is called the **topological dual space** of X.

Theorem 4.33. *Let $(X, \mathcal{O})$ be a TVS and $f : X \to \mathbb{K}$ linear, $f \neq 0$. Then the following statements are equivalent:*

(i) *f is continous.*

(ii) $\ker f$ *is closed.*

(iii) $\ker f$ *is not dense in X.*

(iv) *There exists $U \in \mathcal{O}$ with $0 \in U$ such that $f(U) \subset \mathbb{K}$ is bounded.*

Proof. (i)⇒(ii). If f is continuous then $\ker f = f^{-1}(\{0\})$ is closed since $\{0\}$ is closed in $\mathbb{K}$.

(ii)⇔(iii). Since $f \neq 0$, there exists $x_0 \in X$ such that $fx_0 = 1$. If $x \in X$, then $x = (x - f(x)x_0) + f(x)x_0 \in \ker f \dotplus \operatorname{span}\{x_0\}$, where $\dotplus$ denotes the direct sum. Hence,

$$X = \ker f \dotplus \operatorname{span}\{x_0\}.$$

As $\ker f$ is a subspace, this implies that either $\ker f = \overline{\ker f}$ or $\overline{\ker f} = X$.

(ii)⇒(iv). Let $\ker f$ be closed, and let $x_0 \in X \setminus \ker f$. Applying Theorem 4.29 to $K := \{x_0\}$ and $C := \ker f$, we obtain some $W_0 \in \mathcal{O}$ such that $0 \in W_0$ and $(x_0 + W_0) \cap \ker f = \varnothing$. By Lemma 4.28(iii), there exists a balanced set $W \in \mathcal{O}$ with $0 \in W$ and $W \subset W_0$. Put $U := x_0 + W$. Then we have $U \cap \ker f = \varnothing$.

For $\lambda \in \mathbb{K}$, $|\lambda| \le 1$, we have $\lambda f(W) = f(\lambda W) \subset f(W)$, i.e., $f(W) \subset \mathbb{K}$ is balanced. Thus, if $z \in f(W)$ we have $B_{|z|}(0) \subset f(W)$. This shows that $f(W) = \mathbb{K}$ if $f(W)$ is unbounded. In this case, there exists $y \in W$ with $f(y) = -f(x_0)$, i.e., $y + x_0 \in \ker f \cap U$, a contradiction! Thus, $f(W)$ is bounded.

(iv)$\Leftrightarrow$(i). Let $U \in \mathcal{O}$, $0 \in U$, such that $f(U) \subset B_r(0)$, $r > 0$. Then, for each $\varepsilon > 0$, we have that $f(\frac{\varepsilon}{r}U) \subset B_\varepsilon(0)$. Hence, f is continuous at zero and therefore continuous. □

The proof of the following theorem is left to the reader.

Theorem 4.34. *Let $(X, \mathcal{O})$ be a TVS and $(X_i, \mathcal{O}_i)_{i\in I}$ a family of TVS's. Then the following statements hold.*

(a) *If $Y \subset X$ is a subspace. Then $(Y, \mathcal{O}_Y)$ is a TVS, where $\mathcal{O}_Y$ denotes the trace topology.*

(b) *$\left(\prod_{i\in I} X_i, \mathcal{O}\right)$ is a TVS, where $\mathcal{O}$ denotes the product topology.*

(c) *If $N \subset X$ is a closed subspace then $(X/N, \mathcal{O})$ is a TVS, where $\mathcal{O}$ denotes the quotient topology.*

Theorem 4.35. *Let $(X, \mathcal{O})$ be a TVS and $Y \subset X$ a subpace, $\dim Y = n$. Then the following statements hold.*

(i) *Every algebraic isomorphism $\mathbb{K}^n \to Y$, where $\mathbb{K}^n$ is equipped with the euclidian topology and Y with the trace topology, is a homeomorphism.*

(ii) *Y is closed in X.*

Proof. (i). Let $f : \mathbb{K}^n \to Y$ be an algebraic isomorphism, and let $(e_i)_{i=1,\dots,n}$ be the standard basis of $\mathbb{K}^n$. By $(\delta_i)_{i=1,\dots,n}$ we denote the dual basis with respect to (e_i), i.e., $\delta_i : \mathbb{K}^n \to \mathbb{K}$ is linear with $\delta_i(e_j) = \delta_{ij}$. Finally, define $F_i : \mathbb{K}^n \to \mathbb{K} \times Y$ by $F_i(x) := (\delta_i(x), f(e_i))$, $x \in \mathbb{K}^n$. Then F_i is continuous.

For $x \in \mathbb{K}^n$ we have

$$f(x) = f\left(\sum_{i=1}^n x_i e_i\right) = \sum_{i=1}^n \delta_i(x) f(e_i) = \sum_{i=1}^n \varphi\left(\delta_i(x), f(e_i)\right) = \sum_{i=1}^n \varphi(F_i(x)),$$

where $\varphi : \mathbb{K} \times Y \to Y$, $\varphi(\lambda, y) := \lambda y$. Thus, $f = \sum_{i=1}^n \varphi \circ F_i$. Hence, f is continuous.

In order to show that also f^{-1} is continous, define $S := \{x \in \mathbb{K}^n : \|x\|_2 = 1\}$. As S is compact and f continuous, also $f(S) \subset Y$ is compact. Moreover, $0 \notin f(S)$. By Theorem 4.29 and Lemma 4.28(iii), there exists a balanced $V \in \mathcal{O}$ with $0 \in V$ and $V \cap f(S) = \varnothing$. Put $E := f^{-1}(V \cap Y)$.

E is balanced since for $x \in E$ and $|\lambda| \le 1$ we have $x = f^{-1}(y)$ for some $y \in V \cap Y$, and thus $\lambda x = f^{-1}(\lambda y) \in E$, since $\lambda y \in V \cap Y$. Moreover, we have $E \cap S = \varnothing$. Indeed, if $x \in E \cap S$ then $f(x) \in V \cap Y \cap f(S) = \varnothing$, a contradiction. If $x \in E$ with $\|x\|_2 > 1$ then $\|x\|_2^{-1} x \in E \cap S = \varnothing$ (as E is balanced). Another contradiction! Thus, we have $\|x\|_2 < 1$ for all $x \in E$.

Now, let $\varepsilon > 0$. Then for all $x \in \varepsilon(V \cap Y)$ we have $f^{-1}(x) = \varepsilon e$ for some $e \in E$. This implies $\|f^{-1}(x)\|_2 = \varepsilon\|e\|_2 < \varepsilon$. Thus f^{-1} is continuous.

(ii). Let $z \in \overline{Y} \subset X$, let $f : \mathbb{K}^n \to Y$ be an algebraic isomorphism, and let V and E be as in the proof of (i). As V is absorbing by Lemma 4.28(i) (in X) there exists $t > 0$ such that $tz \in V$ and thus $z \in t^{-1}V$. For any $U \in \mathcal{O}$ with $z \in U$ we obtain $z \in U \cap t^{-1}V$ and thus $U \cap t^{-1}V \cap Y \neq \varnothing$ (as $z \in \overline{Y}$). This implies $z \in \overline{t^{-1}V \cap Y}$. Since $E \subset B := \{x \in \mathbb{K}^n : \|x\|_2 \leq 1\}$, we have

$$t^{-1}V \cap Y = t^{-1}(V \cap Y) = t^{-1}f(E) = f(t^{-1}E) \subset f(t^{-1}B).$$

As B is compact and f continuous, we get $z \in \overline{t^{-1}V \cap Y} \subset f(t^{-1}B) \subset Y$. □

Corollary 4.36. *Let* $(X, \mathcal{O})$ *be a TVS,* $Y, N \subset X$ *closed subspaces, and* $\dim N < \infty$*. Then* $Y + N$ *is closed.*

Proof. Let $\pi : X \to X/Y$ be the canonical projection. Then π is linear and continuous. Since $\dim \pi(N) < \infty$, $\pi(N)$ is closed in X/Y by Theorem 4.35. But

$$\begin{aligned}\pi^{-1}(\pi(N)) &= \{x \in X : \pi(x) \in \pi(N)\} = \{x \in X \mid \exists\, y \in N : \pi(x) = \pi(y)\} \\ &= \{x \in X \mid \exists\, y \in N : x - y \in Y\} = N + Y.\end{aligned}$$

Since π is continuous, we have that $N + Y$ is closed. □

5. Locally Convex Spaces

In the exercise class, we have seen that the space $X = L^p(0,1)$ with $p \in (0,1)$, equipped with the metric

$$d(f,g) := \int_0^1 |f-g|^p \,\mathrm{d}x,$$

is a topological vector space with the following properties:

(a) There are no non-trivial convex open sets in X.

(b) $X' = \{0\}$, which means that the only continuous linear functional on X is the zero functional.

In this chapter, we will analyze topological vector spaces X with "large" dual spaces X' in order to have "a lot of" information on X via X'.

Definition 5.1. The topology of a TVS $(X, \mathcal{O})$ is called **locally convex** if every neighborhood of zero contains a convex neighborhood of zero. In this case, $(X, \mathcal{O})$ is called a **locally convex space** (LCS).

Example 5.2. Any normed space is locally convex.

5.1. Characterization by Seminorms

A (sub)basis of the topology of a normed space is given by the system of open balls which are defined by the norm. In this section, it will turn out that locally convex topologies can be similarly characterized. In this characterization, the norm has to be replaced by a family of seminorms.

Definition 5.3. Let X be a linear space over $\mathbb{K}$. A mapping $p : X \to [0,\infty)$ is called a **seminorm** on X if it satisfies the following two conditions:

(i) $p(x+y) \le p(x) + p(y)$ for all $x, y \in X$.

(ii) $p(\lambda x) = |\lambda| p(x)$ for all $x \in X$ and $\lambda \in \mathbb{K}$.

Definition 5.4. A family P of seminorms on a linear space X is called **separating** if for all $x \in X \setminus \{0\}$ there exists a $p \in P$ such that $p(x) \neq 0$.

Example 5.5. (a) Each norm is separating.

(b) Let $X = C([0,1])$ and $P := \{p_t : t \in [0,1]\}$ with $p_t(f) := f(t)$, $f \in X$, $t \in [0,1]$. Then P is a separating family of seminorms.

(c) Let T be a topological space and $X = C(T) = \{f : T \to \mathbb{K} \,|\, f \text{ is continuous}\}$. Then $P = \{p_K : K \subset T \text{ compact}\}$ with $p_K(f) := \sup_{t\in K} |f(t)|$, $f \in X$, is a separating family of seminorms.

Lemma 5.6. *Let p be a seminorm on a linear space X and $N_p := \{x \in X : p(x) = 0\}$. Then the following statements hold.*

(i) $p(0) = 0$ *and* $p(x) = p(-x)$.

(ii) $|p(x) - p(y)| \le p(x - y)$.

(iii) N_p *is a subspace of* X.

(iv) $\widetilde{p} : X/N_p \to [0, \infty)$, $x + N_p \mapsto p(x)$, *is a norm on* X/N_p.

(v) *If* $(X, \mathcal{O})$ *is a TVS then* $p : (X, \mathcal{O}) \to [0, \infty)$ *is continuous if and only if there exists* $U \in \mathcal{O}$, $0 \in U$, *such that* $p(U)$ *is bounded.*

Proof. (i)–(iv) are trivial. For the proof of (v), let p be continuous. Then $U := p^{-1}([0, 1)) \in \mathcal{O}$, $0 \in U$, and $p(U) \subset [0, 1]$. For the converse, assume that there exists $U \in \mathcal{O}$, $0 \in U$, such that $p(U)$ is bounded. Put $M := \sup\{p(u) : u \in U\} < \infty$. Let $x \in X$, $\varepsilon > 0$, and define $V := x + \frac{\varepsilon}{M+1} U \in \mathcal{O}$. If $y \in V$ then $\frac{M+1}{\varepsilon}(y - x) \in U$, and thus, by (ii) we get

$$|p(y) - p(x)| \le p(y - x) = \frac{\varepsilon}{M+1} \cdot p\left(\frac{M+1}{\varepsilon}(y - x)\right) \le \frac{M}{M+1}\varepsilon < \varepsilon.$$

Thus, $p(V) \subset B_\varepsilon(p(x))$, and p is continuous. $\square$

Theorem 5.7. *Let* X *be a linear space and* P *a separating family of seminorms on* X. *For* $p \in P$ *and* $\varepsilon > 0$ *define*

$$V(p, \varepsilon) := \{x \in X : p(x) < \varepsilon\}.$$

Then the topology $\mathcal{O} := \mathfrak{U}(\mathfrak{I}(\gamma))$ *where*

$$\gamma := \{x + V(p, \varepsilon) : x \in X,\, p \in P,\, \varepsilon > 0\},$$

turns X *into a LCS. Moreover, each* $p \in P$ *is continuous.*

Proof. It is left to the reader to prove that $(X, \mathcal{O})$ is indeed a TVS. As $V(p, \varepsilon)$ is convex for $p \in P$ and $\varepsilon > 0$, (finite) intersections of sets in γ are convex. Therefore, each $U \in \mathcal{O}$ is a union of convex open sets. Hence, $\mathcal{O}$ is locally convex. $\square$

Remark 5.8. We say that P induces or generates the topology $\mathcal{O}$ on X. In the sequel, this topology will often be denoted by $\mathcal{O}_P$.

The above theorem shows that separating families of seminorms generate locally convex topologies. In what follows, we will show that in fact *every* locally convex topology is generated in this way.

Definition 5.9. Let X be a linear space and $A \subset X$ absorbing. The mapping $\mu_A : X \to [0, \infty)$, defined by

$$\mu_A(x) := \inf\{t > 0 : x \in tA\}$$

is called the **Minkowski functional** of A.

Lemma 5.10. *Let X be a linear space and $A \subset X$ convex and absorbing. Then the following statements hold.*

(i) $\mu_A(x+y) \le \mu_A(x) + \mu_A(y)$ *for all* $x, y \in X$.

(ii) $\mu_{t^{-1}A}(x) = \mu_A(tx) = t\mu_A(x)$ *for all* $x \in X$ *and* $t > 0$.

If, in addition, A is balanced, then the following holds:

(iii) $\mu_A(\lambda x) = |\lambda|\mu_A(x)$ *for all* $x \in X$ *and* $\lambda \in \mathbb{K}$.

(iv) $\{x \in X : \mu_A(x) < 1\} \subset A \subset \{x \in X : \mu_A(x) \le 1\}$.

In particular, μ_A is a seminorm on X if A is absorbing, convex, and balanced.

Proof. (i). Let $x, y \in X$. Then there exist $s, t > 0$ such that $x \in sA$ and $y \in tA$. Hence, $x = sa$, $y = tb$ with $a, b \in A$. Since A is convex,

$$x + y = (s+t)\left(\frac{s}{s+t}a + \frac{t}{s+t}b\right) \in (s+t)A.$$

Therefore, $\mu_A(x+y) \le s + t$. Taking infima proves (i).

(ii)–(iv). These are straightforward. We only note that for proving (iii), one uses that for balanced A we have $\lambda A = A$ if $|\lambda| = 1$. □

Theorem 5.11. *Let $(X, \mathcal{O})$ be a LCS. Define the system of sets*

$$B := \{U \in \mathcal{O} : 0 \in U,\ U \text{ convex and balanced}\}.$$

Then $P = \{\mu_U : U \in B\}$ is a separating family of seminorms on X, generating the topology $\mathcal{O}$.

Proof. Let $x \in X \setminus \{0\}$. As $\mathcal{O}$ is Hausdorff, we find $\widetilde{U} \in \mathcal{O}$ with $0 \in \widetilde{U}$ and $x \notin \widetilde{U}$. Since X is locally convex, we find a convex subset $U' \in \mathcal{O}$ of $\widetilde{U}$ with $0 \in U'$. By Lemma 4.28(iv) we get the existence of a convex and balanced subset $U \in \mathcal{O}$ of U' with $0 \in U$. Thus $U \in B$ and $x \notin U$. Thus $\mu_U(x) \ge 1$. Thus, P is separating. By $\mathcal{O}_1$ we denote the topology generated by P. Let us show that $\mathcal{O}_1 = \mathcal{O}$.

$\mathcal{O}_1 \subset \mathcal{O}$: It suffices to prove that $V(\mu_U, \varepsilon) \in \mathcal{O}$ for all $U \in B$ and $\varepsilon > 0$. Let $x \in V(\mu_U, \varepsilon)$. Thus, $\inf\{t > 0 : x \in tU\} < \varepsilon$. Hence, there exists a $t \in (0, \varepsilon)$ such that $x \in tU$. If $y \in tU$, then we have $\mu_U(y) = t\mu_U(t^{-1}y) \le t < \varepsilon$, i.e., $y \in V(\mu_U, \varepsilon)$. Thus, $tU \subset V(\mu_U, \varepsilon)$ and hence $V(\mu_U, \varepsilon) \in \mathcal{O}$.

$\mathcal{O} \subset \mathcal{O}_1$: Let $\widehat{U} \in \mathcal{O}$ and $x \in \widehat{U}$. Put $V := \widehat{U} - x$ and choose $U \subset V$, $0 \in U$, U convex and balanced. By Lemma 5.10, we have $V(\mu_U, 1) \subset U$. Hence,

$$x + V(\mu_U, 1) \subset x + U \subset x + V = \widehat{U}.$$

This proves the theorem. □

Example 5.12. (a) Let X and Y be normed spaces. For $x \in X$ and $T \in L(X,Y)$, set $p_x(T) := \|Tx\|$. Then p_x is a seminorm on $L(X,Y)$ and $P := \{p_x : x \in X\}$ is separating and induces the so called *strong operator topology* on $L(X,Y)$.
(b) Let $H(\mathbb{D})$ denote the space of holomorphic functions on $\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$. For $K \subset \mathbb{D}$ compact, set $p_K(f) := \sup_{z\in K} |f(z)|$ for $f \in H(\mathbb{D})$. Then $H(\mathbb{D})$ becomes a LCS with $P := \{p_K : K \subset \mathbb{D} \text{ compact}\}$. We have $f_n \to f$ in $\mathcal{O}_P$ if and only if $f_n \to f$ locally uniformly.

We conclude this section with two statements on the continuity of seminorms on and linear maps between locally convex spaces.

Lemma 5.13. *Let $(X, \mathcal{O})$ be a LCS, where $\mathcal{O}$ is generated by a family P of seminorms. A seminorm q on X is continuous if and only if there exist $M \geq 0$ and a finite set $F \subset P$ such that for all $x \in X$ we have*

$$q(x) \leq M \cdot \max_{p\in F} p(x). \tag{5.1}$$

Proof. By Lemma 5.6 we have that q is continuous if and only if there exists a neighborhood $U \in \mathcal{O}$ of 0 such that $q(U)$ is bounded. This is equivalent to the existence of $\varepsilon > 0$ and a finite subset $F \subset P$ such that

$$q\left(\bigcap_{p\in F} V(p,\varepsilon)\right) \text{ is bounded.}$$

That is, there exist $\varepsilon > 0$, $F \subset P$ finite, and $\widetilde{M} \geq 0$ such that

$$p(x) \leq \varepsilon \;\forall p \in F \implies q(x) \leq \widetilde{M}. \tag{5.2}$$

Clearly, (5.1) implies (5.2) with $\varepsilon = 1$ and $\widetilde{M} = M$. Now, assume that (5.2) holds, and define the seminorm $p_0(x) := \max_{p\in F} p(x)$, $x \in X$. Fix $x \in X$. If $p_0(x) = 0$, then also $p_0(tx) = tp_0(x) = 0$ for all $t > 0$. Hence, $tq(x) = q(tx) \leq \widetilde{M}$ for all $t > 0$ and thus $q(x) = 0$, so that (5.1) is satisfied. Let $p_0(x) > 0$. Then, since

$$p\left(\frac{\varepsilon x}{p_0(x)}\right) = \varepsilon\frac{p(x)}{p_0(x)} \leq \varepsilon$$

for all $p \in F$, we obtain from (5.2),

$$q(x) = \frac{p_0(x)}{\varepsilon} \cdot q\left(\frac{\varepsilon x}{p_0(x)}\right) \leq \frac{\widetilde{M}}{\varepsilon} p_0(x).$$

Setting $M := \varepsilon^{-1}\widetilde{M}$ completes the proof. □

Theorem 5.14. *Let $(X, \mathcal{O}_P)$ and $(Y, \mathcal{O}_Q)$ be two LCS's generated by the families of seminorms P and Q, respectively. Further, let $T : X \to Y$ be linear. Then the following statements are equivalent.*

(i) *T is continuous.*

(ii) *For every $q \in Q$ we have that $q \circ T$ is continous on X.*

(iii) *For each continuous seminorm q on Y we have that $q \circ T$ is continous on X.*

Proof. (i)⇒(ii) is clear. To prove (ii)⇒(iii), let q_0 be a continuous seminorm on Y. By Lemma 5.13, there exist a finite set $G \subset Q$ and $M_0 > 0$ such that $q_0(y) \le M_0 \max_{q\in G} q(y)$ for all $y \in Y$. Hence, $(q_0 \circ T)(x) \le M_0 \max_{q\in G}(q \circ T)(x)$ for all $x \in X$. Now, apply Lemma 5.13 to each $q \circ T$, $q \in G$, to obtain that $q_0 \circ T$ is a continuous seminorm on X.

(iii)⇒(i). We prove that T is continuous at zero. For this, it suffices to show that for $\varepsilon > 0$ and finitely many $q_1, \dots, q_n \in Q$, there exist $U \in \mathcal{O}_P$ with $0 \in U$ such that

$$T(U) \subset V := \bigcap_{i=1}^{n} V(q_i, \varepsilon) = \{y \in Y : q_i(y) < \varepsilon \,\forall\, i = 1, \dots, n\}.$$

By (iii), there exist $M_1, \dots, M_n \ge 0$ and finite subsets $F_1, \dots, F_n \subset P$ such that for all $x \in X$ and all $i \in \{1, \dots, n\}$ we have

$$q_i(Tx) \le M_i \max_{p\in F_i} p(x).$$

Define $F := \bigcup_{i=1}^{n} F_i$ and $M := \max_{i=1,\dots,n} M_i$. Then for all $x \in X$ we have

$$\max_{i=1,\dots,n} q_i(Tx) \le M \max_{p\in F} p(x).$$

If $x \in \bigcap_{p\in F} V\left(p, \frac{\varepsilon}{M}\right)$ then $\max_{p\in F} q_i(Tx) < \varepsilon$, i.e., $Tx \in V$ and thus

$$T\left(\bigcap_{p\in F} V\left(p, \frac{\varepsilon}{M}\right)\right) \subset V.$$

The theorem is proved. □

5.2. The Hahn-Banach Theorem

First of all, let us recall the linear algebra version of the Hahn-Banach theorem.

Theorem 5.15. *Let X be a linear space over $\mathbb{K}$, $M \subset X$ a subspace, $p : X \to \mathbb{R}$ sublinear, and $\ell : M \to \mathbb{K}$ a linear functional satisfying $\operatorname{Re}\ell(x) \le p(x)$ for all $x \in M$. Then there exists a linear functional $L : X \to \mathbb{K}$, such that*

$$L|_M = \ell \quad \textit{and} \quad \operatorname{Re} L(x) \le p(x) \textit{ for all } x \in X.$$

Remark 5.16. (a) $p : X \to \mathbb{R}$ is sublinear if $p(x + y) \le p(x) + p(y)$ and $p(\lambda x) = \lambda p(x)$ for all $x, y \in X$ and $\lambda \ge 0$.

(b) If $\mathbb{K} = \mathbb{C}$, note that a linear functional f on X is uniquely determined by its real part $\operatorname{Re} f$ since $-\operatorname{Re} f(ix) = -\operatorname{Re}(if(x)) = \operatorname{Im} f(x)$. Moreover, if $A \subset X$ is balanced, we have that

$$\sup_{x \in A} |f(x)| = \sup_{x \in A} \operatorname{Re} f(x).$$

Indeed, for $x \in X$ we have $f(x) = |f(x)|\lambda^{-1}$, where $|\lambda| = 1$, and hence $|f(x)| = \operatorname{Re}(\lambda f(x)) = \operatorname{Re} f(\lambda x)$. The claim now follows from the fact that $\lambda A = A$.

We will now make use of Theorem 5.15 to prove the following separation theorem.

Theorem 5.17. *Let X be a TVS and $A, B \subset X$ non-empty convex sets with $A \cap B = \varnothing$. Then the following statements hold.*

(i) *If A is open, then there exist $f \in X'$ and $\gamma \in \mathbb{R}$ such that*

$$\operatorname{Re} f(x) < \gamma \le \operatorname{Re} f(y) \qquad \text{for all } x \in A \text{ and } y \in B.$$

(ii) *If X is a LCS, A is compact, and B is closed, then there exist $f \in X'$ and $\gamma_1, \gamma_2 \in \mathbb{R}$ such that*

$$\operatorname{Re} f(x) < \gamma_1 < \gamma_2 < \operatorname{Re} f(y) \qquad \text{for all } x \in A \text{ and } y \in B.$$

Proof. Below, we will prove the theorem for $\mathbb{K} = \mathbb{R}$. Let us see that this implies the general case. For this, assume that the theorem is true for $\mathbb{K} = \mathbb{R}$, and let $\mathbb{K} = \mathbb{C}$. Then X is also a linear space over $\mathbb{R}$. Denote this space by $X_{\mathbb{R}}$. Then $(X_{\mathbb{R}}, \mathcal{O})$ is a TVS, and $(X_{\mathbb{R}}, \mathcal{O})$ is a LCS if $(X, \mathcal{O})$ is a LCS. Hence, in each of the cases (i) and (ii), we obtain a continuous $\mathbb{R}$-linear functional $g : X \to \mathbb{R}$ with the respective property. Now, put $f(x) := g(x) - ig(ix)$, $x \in X$. Then $\operatorname{Re} f = g$, and f is $\mathbb{C}$-linear. Moreover, $\varphi : X \to X$ with $\varphi(x) := ix$, $x \in X$, is continuous, and so $g \circ \varphi : X \to \mathbb{R}$ is a continuous $\mathbb{R}$-linear functional. Hence, $\ker f = \ker g \cap \ker(g \circ \varphi)$ is closed, and thus $f \in X'$ by Theorem 4.33. Let us now prove the theorem for $\mathbb{K} = \mathbb{R}$.

(i). Let $a_0 \in A$, $b_0 \in B$, and put $x_0 := b_0 - a_0$ and $C := A - B - x_0$. Then C is a convex open neighborhood of 0 and $x_0 \notin C$. Thus, C is absorbing (as an open neighborhood of 0) and $\mu_C(x_0) \ge 1$ (cf. Lemma 5.10). On $M := \operatorname{span}\{x_0\}$, define $g : M \to \mathbb{R}$ by $g(tx_0) := t$, $t \in \mathbb{R}$. For $t \ge 0$, we have $g(tx_0) = t \le t\mu_C(x_0) = \mu_C(tx_0)$, and if $t < 0$ then $g(tx_0) = t < 0 \le \mu_C(tx_0)$. Thus, $g(x) \le \mu_C(x)$ for all $x \in M$.

By Theorem 5.15, there exists a linear functional $f : X \to \mathbb{R}$ such that $f|_M = g$ and $f(x) \le \mu_C(x)$ for all $x \in X$. For all $x \in C$ we thus have $f(x) \le 1$. Also, $f(x) \ge -1$ for $x \in -C$. The set $C \cap -C$ is an open neighborhood of $0 \in X$ and $|f(x)| \le 1$ for all $x \in C \cap -C$. From Theorem 4.33 we conclude that $f \in X'$.

For $x \in C$ we have $\mu_C(x) < 1$. Otherwise, $x \notin tC$ for all $t \in (0,1)$. But then $t^{-1}x \in X \setminus C$ for all $t \in (0,1)$. Hence, $x = \lim_{n\to\infty}(1 - \frac{1}{n})^{-1}x \in X \setminus C$, a contradiction.

Now, for $a \in A$ and $b \in B$, we have $f(a) - f(b) + 1 = f(a - b + x_0) \le \mu_C(a - b + x_0) < 1$. Thus, $f(a) < f(b)$ for all $a \in A$ and $b \in B$. In an exercise it was shown that $f(A)$ is open. This implies $f(a) < \gamma := \sup f(A) \le f(b)$ for all $a \in A$ and $b \in B$.

(ii). By Theorem 4.29, there exists $V \in \mathcal{O}$ with $0 \in V$ such that $(A+V) \cap B = \varnothing$. As X is a LCS, we may assume that V is convex and balanced. Then $A+V$ is convex and open. Hence, by (i), there exist $f \in X'$ and $\gamma \in \mathbb{R}$ such that $f(a+v) < \gamma \le f(b)$ for all $a \in A$, $v \in V$, and $b \in B$. Let $y \in A$ be such that $f(y) = \max_{a \in A} f(a)$. Then

$$\sup f(A) = f(y) < f(y) + \sup f(V) = \sup f(y+V) \le \sup f(A+V) \le \gamma.$$

Thus, there exist $\gamma_1, \gamma_2 \in \mathbb{R}$ such that $\sup f(A) < \gamma_1 < \gamma_2 < \gamma \le f(b)$ for all $b \in B$. □

Corollary 5.18. *Let X be a LCS. Then the following statements hold.*

(i) *X' is separating points, i.e., for all $x, y \in X$ with $x \neq y$ there exists $f \in X'$ with $f(x) \neq f(y)$.*

(ii) *If $M \subset X$ is a closed subspace and $x_0 \in X \setminus M$, then there exists $f \in X'$ with $f|_M = 0$ and $f(x_0) = 1$.*

(iii) $\dim X = \dim X'$.

Proof. (i). Let $x \neq y$ and put $A := \{x\}$, $B := \{y\}$. By Lemma 5.17 there exists $f \in X'$ with $\operatorname{Re} f(x) < \operatorname{Re} f(y)$.

(ii). Set $A := \{x_0\}$ and $B := M$. By Lemma 5.17, there exists $g \in X'$ with $\operatorname{Re} g(x_0) < \operatorname{Re} g(x)$ for all $x \in M$. Suppose $\operatorname{Re} g(x) \neq 0$ for some $x \in M \setminus \{0\}$. Then $\operatorname{Re} g(tx) = t \operatorname{Re} g(x) \to -\infty$ for $t \to \infty$ or $t \to -\infty$. A contradiction. Thus, $\operatorname{Re} g(x) = 0$ for all $x \in M$. Therefore, $g(x) = \operatorname{Re} g(x) - i \operatorname{Re} g(ix) = 0$ for all $x \in M$, and also $g(x_0) < 0$. Hence, the functional $f : x \mapsto g(x)/g(x_0)$ is as desired.

(iii). If $\dim X < \infty$, then $\dim X = \dim X'$ is a well known fact from linear algebra. Let now $\dim X = \infty$. In this case, there exists a sequence of subspaces $Y_n \subset X$ with $Y_n \subset Y_{n+1}$ and $\dim Y_n = n$, $n \in \mathbb{N}$. For each $n \in \mathbb{N}$, choose a vector $y_n \in Y_n \setminus Y_{n-1}$. As each Y_n is closed (cf. Theorem 4.25), there exist $f_n \in X'$ with $f_n|_{Y_{n-1}} = 0$ and $f_n(y_n) = 1$. The f_n are linearly independent, since

$$\sum_{n=1}^{N} \lambda_n f_n = 0 \quad \Longrightarrow \quad 0 = \sum_{n=1}^{N} \lambda_n f_n(y_1) = \lambda_1 f_1(y_1) = \lambda_1,$$

and iteratively, $\lambda_k = 0$ for all $k = 1, \ldots, N$. Thus, we have constructed an infinite sequence of linear independent elements of X', i.e., $\dim X' = \infty$. □

5.3. The Krein-Milman Theorem

Definition 5.19. Let X be a linear space and $K \subset X$. The **convex hull** of K is defined by

$$\operatorname{co}(K) := \bigcap \{C \supset K : C \text{ convex}\}.$$

We write $\overline{\operatorname{co}}(K) := \overline{\operatorname{co}(K)}$ if X is a TVS.

Lemma 5.20. *Let X be a linear space and $K \subset X$. Then*

$$\operatorname{co}(K) = \left\{ \sum_{i=1}^{n} \lambda_i x_i : n \in \mathbb{N},\, \lambda_i \in [0,1],\, x_i \in K\, \forall i,\, \sum_{i=1}^{n} \lambda_i = 1 \right\}.$$

Proof. Exercise □

Definition 5.21. Let X be a linear space and $K \subset X$ convex. A point $z \in K$ is called an **extreme point** of K, if whenever $x, y \in K$ and $t \in (0,1)$ such that $tx + (1-t)y = z$, then $x = y = z$. By $\operatorname{ex}(K)$ we denote the set of all extreme points of K.

In other words, $z \in K$ is an extreme point of K if it is not in the relative interior of a segment in K.

Example 5.22. (a) $X = \mathbb{R}$, $K = [a,b]$. Then $\operatorname{ex}(K) = \{a, b\}$.
(b) $X = \mathbb{C}$, $K = \{z \in \mathbb{C} : |z| \leq 1\}$. Then $\operatorname{ex}(K) = \{z \in \mathbb{C} : |z| = 1\}$.
(c) $X = L^\infty((a,b))$, $K = \{f \in X : f(x) \in [0,1] \text{ a.e.}\}$. Then $\operatorname{ex}(K) = \{\chi_E : E \in \Sigma\}$, where Σ denotes the Borel σ-algebra on (a,b).
(d) $X = L^1((0,1))$, $K = \{f \in X : \|f\|_{L^1} \leq 1\}$. Then $\operatorname{ex}(K) = \varnothing$. To see this, let $f \in K$ be arbitrary, i.e., $\int_0^1 |f| \,\mathrm{d}x \leq 1$. By the intermediate value theorem, there exists $x_0 \in (0,1)$ such that

$$\int_0^{x_0} |f| \,\mathrm{d}x = \frac{1}{2}\|f\|_{L^1}.$$

Define $g := 2\chi_{[0,x_0]} f$ and $h := 2\chi_{[x_0,1]} f$. Then

$$\|g\|_{L^1} = 2\int_0^{x_0} |f| \,\mathrm{d}x = \|f\|_{L^1} \leq 1$$

and similarly $\|h\|_{L^1} = \|f\|_{L^1} \leq 1$. Thus $g, h \in K$ and $f = \frac{1}{2}g + \frac{1}{2}h$. Thus $f \notin \operatorname{ex}(K)$.

Lemma 5.23. *Let X be a LCS and $K_1, \ldots, K_n \subset X$ compact convex sets. Then $K := \operatorname{co}(K_1 \cup \ldots \cup K_n)$ is compact and $\operatorname{ex}(K) \subset K_1 \cup \ldots \cup K_n$.*

Proof. Put

$$\Delta_n := \left\{ (\lambda_1, \ldots, \lambda_n) \in [0,1]^n : \sum_{i=1}^{n} \lambda_i = 1 \right\},$$

and define

$$f : \Delta_n \times K_1 \times \ldots \times K_n \to X, \quad (\lambda_1, \ldots, \lambda_n, x_1, \ldots, x_n) \mapsto \sum_{i=1}^{n} \lambda_i x_i.$$

Obviously, f is continuous. Let us show that $\operatorname{im}(f) = K$. Then it follows, that K is compact. The inclusion $\operatorname{im}(f) \subset K$ is clear. Let $x \in K$. Then there exist $\mu_1, \ldots, \mu_m \in [0,1]$ with $\sum_{i=1}^{m} \mu_i = 1$ and $y_1, \ldots, y_m \in C := K_1 \cup \ldots \cup K_n$ such that $x = \sum_{j=1}^{m} \mu_j y_j$.

Fore each $j \in \{1, \ldots, m\}$ there exists $i(j) \in \{1, \ldots, n\}$ such that $y_j \in K_{i(j)}$. Put $\lambda_i := \sum_{j:i(j)=i} \mu_j$ (empty sums are to be treated as zero). Then

$$x = \sum_{i=1}^{n} \sum_{j:i(j)=i} \mu_j y_j = \sum_{i=1}^{n} \lambda_i \underbrace{\sum_{j:i(j)=i} \frac{\mu_j}{\lambda_i} y_j}_{\in K_i} \qquad \text{and} \qquad \sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} \sum_{j:i(j)=i} \mu_j = 1.$$

Thus, $x \in \operatorname{im}(f)$.

It remains to prove that $\operatorname{ex}(K) \subset C$. For this, let $x \in \operatorname{ex}(K) \subset K = \operatorname{co}(C)$. Then $x = \sum_{j=1}^{m} \lambda_j y_j$ with $\lambda_j \in [0,1]$, $\sum_{j=1}^{m} \lambda_j = 1$, and $y_j \in C$. Without loss of generality, we may assume that $\lambda_1 \in (0,1)$. Thus,

$$x = \lambda_1 y_1 + (1 - \lambda_1) \sum_{j=2}^{m} \frac{\lambda_j}{1 - \lambda_1} y_j.$$

We have $y_1 \in K$, and since

$$\sum_{j=2}^{m} \frac{\lambda_j}{1 - \lambda_1} = \frac{1}{1 - \lambda_1}(1 - \lambda_1) = 1,$$

we have $\sum_{j=2}^{m} \frac{\lambda_j}{1-\lambda_1} y_j \in K$. Hence, as $x \in \operatorname{ex}(K)$, we obtain $x = y_1 \in C$. □

Definition 5.24. Let X be a linear space and $K \subset X$ convex. A set $S \subset K$ is called **extreme** with respect to K if $x, y \in K$, $t \in (0,1)$ and $tx + (1-t)y \in S$ implies $x, y \in S$.[7]

Theorem 5.25 (Krein-Milman Theorem). *Let X be a LCS and $K \subset X$ convex and compact, $K \neq \varnothing$. Then the following statements hold.*

(i) $\operatorname{ex}(K) \neq \varnothing$.

(ii) $K = \overline{\operatorname{co}}(\operatorname{ex}(K))$.

(iii) *If $K = \overline{\operatorname{co}}(B)$, then $\operatorname{ex}(K) \subset \overline{B}$.*

Proof. Define the following system of sets:

$$P := \{S \subset K : S \text{ compact and extreme with respect to } K,\ S \neq \varnothing\}.$$

Then $P \neq \varnothing$, since $K \in P$. The proof is devided into four parts.

1. In this first part of the proof, let us show that the following two claims hold true:

(a) If $U \subset P$ and $S = \bigcap_{T \in U} T$, then we have $S = \varnothing$ or $S \in P$.

(b) If $S \in P$, $f \in X'$, $\mu = \max\{\operatorname{Re} f(x) : x \in S\}$ and $S_f := \{x \in S : \operatorname{Re} f(x) = \mu\}$, then $S_f \in P$.

[7] A less formal way to write this would be $x, y \in K$, $(x,y) \cap S \neq \varnothing$ implies $[x,y] \subset S$.

(a). As an intersection of compact sets, S is compact. Assume that $S \neq \varnothing$. If $x, y \in K$ and $t \in (0,1)$, such that $tx + (1-t)y \in S$, then $tx + (1-t)y \in T$ for each $T \in U$ and thus $x, y \in T$ for each $T \in U$, i.e., $x, y \in S$. Hence, S is extreme with respect to K, and $S \in P$ follows.

(b). $S_f \neq \varnothing$ since S is compact and $\operatorname{Re} f$ is continuous. Moreover, $S_f = (\operatorname{Re} f)^{-1}(\{\mu\})$ is closed. From $S_f \subset S$ it follows that S_f is compact. In order to see that S_f is extreme with respect to K, let $x, y \in K$, $t \in (0,1)$, and $z := tx + (1-t)y \in S_f$. Since $S_f \subset S$ and S is extreme, we have $x, y \in S$. Thus, $\operatorname{Re} f(x) \leq \mu$ and $\operatorname{Re} f(y) \leq \mu$. But $z \in S_f$ implies $\mu = \operatorname{Re} f(z) = t \operatorname{Re} f(x) + (1-t) \operatorname{Re} f(y) \leq \mu$. Therefore, $\operatorname{Re} f(x) = \operatorname{Re} f(y) = \mu$, and we obtain $x, y \in S_f$. Hence, S_f is extreme, and $S_f \in P$ follows.

2. Let $S \in P$. Let us show that $S \cap \operatorname{ex}(K) \neq \varnothing$. Then (i) follows. For this, define $P_S := \{T \in P : T \subset S\}$. We have $P_S \neq \varnothing$, since $S \in P_S$. Order P_S by inclusion, which is a partial order on P_S. By the Hausdorff Maximal Principle (Theorem 4.19) there exists a maximal totally ordered $\Omega \subset P_S$. Put $M := \bigcap_{T \in \Omega} T$. Suppose that $M = \varnothing$. Then

$$K = K \setminus M = K \setminus \bigcap_{T \in \Omega} T = \bigcup_{T \in \Omega} (K \setminus T).$$

Since K is compact, there exists a finite subset $\Omega' \subset \Omega$ such that $K = \bigcup_{T \in \Omega'} (K \setminus T)$ which implies $\bigcap_{T \in \Omega'} T = \varnothing$. We have $\Omega' = \{T_1, \ldots, T_n\}$ with $T_1 \subset T_2 \subset \ldots \subset T_n$. Thus, $T_1 = \varnothing$, which contradicts $T_1 \in \Omega' \subset \Omega \subset P_S \subset P$. Hence, $M \neq \varnothing$. By (a), we obtain $M \in P$. And since $M \subset S$, we conclude that $M \in P_S$.

Now, if $M' \subset M$ with $M' \in P_S$, define $\widetilde{\Omega} := \Omega \cup \{M'\}$. Then $\widetilde{\Omega} \supset \Omega$ is totally ordered and $\widetilde{\Omega} \subset P_S$. But Ω is maximal, and thus $\widetilde{\Omega} = \Omega$. This implies $M' \in \Omega$. Thus, $M \subset M'$, and hence $M = M'$.

We have $M \in P$. Let $f \in X'$, and set $\mu_f := \max\{\operatorname{Re} f(x) : x \in M\}$, $M_f := \{x \in M : \operatorname{Re} f(x) = \mu_f\}$. Then, by (b), $M_f \in P$. Moreover, $M_f \subset M \subset S$, and thus $M_f \in P_S$, $M_f \subset M$, i.e., $M_f = M$. This shows that $\operatorname{Re} f = \mathit{const.}$ on M. Since X' is separating points, it follows that $M = \{x_0\}$ with some $x_0 \in S$.

Let $x, y \in K$ and $t \in (0,1)$ such that $x_0 = tx + (1-t)y$. Then $tx + (1-t)y \in T$ for all $T \in \Omega \subset P_S \subset P$. As each $T \in \Omega$ is extreme with respect to K, we have $x, y \in T$ for each $T \in \Omega$, and thus $x, y \in \bigcap_{T \in \Omega} T = M = \{x_0\}$. This proves that $x_0 = x = y$ and hence $x_0 \in \operatorname{ex}(K)$. Hence, $S \cap \operatorname{ex}(K) \neq \varnothing$.

3. In this step, we prove (ii). Put $H := \operatorname{co}(\operatorname{ex}(K))$. Our aim is to show that $\overline{H} = K$. We have $H \subset \operatorname{co}(K) = K$, and since K is compact, $\overline{H} \subset K$. Suppose that there exists $x_0 \in K \setminus \overline{H}$. By Theorem 5.17 there exists a functional $f \in X'$ such that $\operatorname{Re} f(x) < \operatorname{Re} f(x_0)$ for all $x \in \overline{H}$. Put $\mu = \max\{\operatorname{Re} f(x) : x \in K\}$, $K_f := \{x \in K : \operatorname{Re} f(x) = \mu\}$. By (b), we have $K_f \in P$. Thus, $K_f \cap \operatorname{ex}(K) \neq \varnothing$, which implies $K_f \cap H \neq \varnothing$. But for $x \in K_f \cap H$, we have $\mu = \operatorname{Re} f(x) < \operatorname{Re} f(x_0) \leq \mu$, a contradiction. (ii) is proved.

4. It remains to prove (iii). First of all, we observe that $B \subset \operatorname{co}(B) \subset \operatorname{co}(\overline{B})$. Hence, $\overline{B} \subset \overline{\operatorname{co}}(B) \subset \overline{\operatorname{co}}(\overline{B})$, and thus $\overline{\operatorname{co}}(\overline{B}) \subset \overline{\operatorname{co}}(B) \subset \overline{\operatorname{co}}(\overline{B})$. Therefore, we may assume that B is closed.

Let $x \in \operatorname{ex}(K)$, and let U be a convex and symmetric closed neighborhood of 0. We shall show that $(x + U) \cap B \neq \varnothing$. Then it follows from Corollary 4.30 that $x \in \overline{B} = B$.

As $K = \overline{\text{co}}(B) \supset B$, B is compact, so that there exist $x_1, \ldots, x_n \in B$ such that $B \subset \bigcup_{i=1}^n (x_i + U)$. Set $K_i := \overline{\text{co}}\,((x_i + U) \cap B) \subset \overline{\text{co}}(B) = K$. Thus, the K_i are compact and convex. Furthermore,

$$B \subset \bigcup_{i=1}^n [(x_i + U) \cap B] \subset \bigcup_{i=1}^n K_i.$$

By Lemma 5.23, $\text{co}\,(\bigcup_{i=1}^n K_i)$ is compact, which implies

$$K = \overline{\text{co}}(B) \subset \overline{\text{co}} \left(\bigcup_{i=1}^n K_i \right) = \text{co} \left(\bigcup_{i=1}^n K_i \right) \subset \text{co}(K) = K.$$

This shows $K = \text{co}\,(\bigcup_{i=1}^n K_i)$. Again, by Lemma 5.23, as $x \in \text{ex}(K)$, we have $x \in \bigcup_{i=1}^n K_i$, and hence $x \in K_i = \overline{\text{co}}((x_i + U) \cap B)$ for some $i \in \{1, \ldots, n\}$. Further, as U is closed and convex, $K_i \subset \overline{\text{co}}(x_i + U) = x_i + U$. Thus, $x = x_i + u$ for some $u \in U$. Since U is symmetric, we have $x_i = x - u \in x + U$ and thus $x_i \in (x + U) \cap B$. □

5.4. Weak Topologies

Definition 5.26. Let X, Y be linear spaces and $\langle \cdot, \cdot \rangle : X \times Y \to \mathbb{K}$ a bilinear form. (X, Y) is called a **dual pair** if for all $x \in X \setminus \{0\}$ there exists $y \in Y$ such that $\langle x, y \rangle \neq 0$, and for all $y \in Y \setminus \{0\}$ there exists $x \in X$ such that $\langle x, y \rangle \neq 0$.

Note that if (X, Y) is a dual pair, then also (Y, X) is a dual pair.

Example 5.27. Let X be a LCS. Then (X, X') is a dual pair via $\langle x, f \rangle = f(x)$, since X' is separating points.

Definition 5.28. Let (X, Y) be a dual pair. For $y \in Y$, let $p_y : X \to [0, \infty)$, $x \mapsto |\langle x, y \rangle|$. Then $P = \{p_y : y \in Y\}$ is a separating family of seminorms. We define $\sigma(X, Y) := \mathcal{O}_P$. The topology $\sigma(Y, X)$ is defined analogously.

Note that $(X, \sigma(X, Y))$ and $(Y, \sigma(Y, X))$ are locally convex spaces.

Example 5.29. (a) The **weak topology** $\sigma(X, X')$ on a LCS X: We have $\sigma(X, X') \subset \mathcal{O}_X$, i.e., $\sigma(X, X')$ is a coarser locally convex topology on X than $\mathcal{O}_X$.
(b) The **weak* topology** $\sigma(X', X)$ on the dual space X' of a LCS. Note that if X is a normed space, we have the norm topology $\mathcal{O}_{X'}$ on X', and again, $\sigma(X', X)$ is coarser than $\mathcal{O}_{X'}$. It is even coarser than $\sigma(X', X'')$. For a general LCS, however, there clearly is no norm topology on X'.

Remark 5.30. (a) Let X be a LCS, $\dim X = \infty$. If $f_1, \ldots, f_n \in X'$ then

$$\bigcap_{i=1}^n \ker f_i \subset \{x \in X : |f_i(x)| < \varepsilon \; \forall \, i \in \{1, \ldots, n\}\}.$$

Thus, every weakly open neighborhood of zero contains a whole non-trivial subspace. If X is a normed space, this in particular implies that $\sigma(X, X') \subsetneq \mathcal{O}_X$.

(b) The space $L^1(0,1)$ is not (isometrically isomorphic to) the dual space of a normed space. Indeed, otherwise, there would be a locally convex topology on $L^1(0,1)$ with respect to which $\overline{B_1(0)}$ is compact (Alaoglu). But by Theorem 5.25, $\mathrm{ex}(\overline{B_1(0)}) \neq \varnothing$, contradicting Example 5.22(d).

Our next aim is to determine the dual space of $(X, \sigma(X, X'))$.

Lemma 5.31. *Let X be a linear space over $\mathbb{K}$ and $f, f_1, \dots, f_n : X \to \mathbb{K}$ linear functionals. Put $N := \bigcap_{i=1}^n \ker f_i$. Then the following are equivalent:*

(i) $f \in \operatorname{span}\{f_1, \dots, f_n\}$.

(ii) *There exists $M > 0$ such that $|f(x)| \leq M \cdot \max_{i=1,\dots,n} |f_i(x)|$ for all $x \in X$.*

(iii) $N \subset \ker f$.

Proof. The implications (i)⇒(ii)⇒(iii) are immediate. To prove (iii)⇒(i), define $V := \{(f_1(x), \dots, f_n(x)) : x \in X\} \subset \mathbb{K}^n$. V is obviously a subspace of $\mathbb{K}^n$. Define $\phi : V \to \mathbb{K}$ by $\phi((f_1(x), \dots, f_n(x))) := f(x)$, $x \in X$. This is well defined, since for two $x, \widehat{x} \in X$ with $(f_1(x), \dots, f_n(x)) = (f_1(\widehat{x}), \dots, f_n(\widehat{x}))$, we obviously have $x - \widehat{x} \in N \subset \ker f$, and thus $f(x) = f(\widehat{x})$. The functional ϕ is linear. Indeed, for $\lambda, \mu \in \mathbb{K}$ and $r = (f_1(x), \dots, f_n(x))$, $s = (f_1(y), \dots, f_n(y))$, $x, y \in X$, we have

$$\begin{aligned}
\phi(\lambda r + \mu s) &= \phi\left(\lambda(f_1(x), \dots, f_n(x)) + \mu(f_1(y), \dots, f_n(y))\right) \\
&= \phi\left(f_1(\lambda x + \mu y), \dots, f_n(\lambda x + \mu y)\right) \\
&= f(\lambda x + \mu y) = \lambda f(x) + \mu f(y) = \lambda \phi(r) + \mu \phi(s).
\end{aligned}$$

Let $\widehat{\phi} : \mathbb{K}^n \to \mathbb{K}$ be a linear extension of ϕ. Then $\widehat{\phi}(\xi_1, \dots, \xi_n) = \sum_{i=1}^n \alpha_i \xi_i$ for suitable $\alpha_i \in \mathbb{K}$, $i = 1, \dots, n$. Hence,

$$f(x) = \widehat{\phi}((f_1(x), \dots, f_n(x))) = \sum_{i=1}^n \alpha_i f_i(x)$$

for all $x \in X$, and thus $f = \sum_{i=1}^n \alpha_i f_i$. ☐

Corollary 5.32. *Let (X, Y) be a dual pair. A linear functional on X is $\sigma(X, Y)$-continuous if and only if it has the form $x \mapsto \langle x, y \rangle$ for some $y \in Y$.*

Proof. Let $f : X \to \mathbb{K}$ be linear. Then f is $\sigma(X, Y)$-continuous if and only if $|f|$ is. By Lemma 5.13, this is equivalent to the fact that there exist $y_1, \dots, y_m \in Y$ and $M \geq 0$ such that for all $x \in X$ we have

$$|f(x)| \leq M \cdot \max_{i=1,\dots,m} |\langle x, y_i \rangle|.$$

By Lemma 5.31, this is equivalent to the fact that $f \in \operatorname{span}\{\langle \cdot, y_i \rangle : i \in \{1, \dots, m\}\}$ for some $y_1, \dots, y_m \in Y$. This occurs if and only if $f = \langle \cdot, y \rangle$ for some $y \in Y$. ☐

Corollary 5.33. *Let X be a LCS. Then*

(a) *A linear functional $f : X \to \mathbb{K}$ is weakly continuous if and only if it is continuous with respect to $\mathcal{O}_X$. That is, $(X, \sigma(X, X'))' = X'$, although $\sigma(X, X')$ is coarser than $\mathcal{O}_X$.*

(b) *A linear functional $F : X' \to \mathbb{K}$ is weak-$*$-continuous if and only if it is an evaluation functional, i.e., $F(f) = f(x)$ for some fixed $x \in X$.*

Proposition 5.34. *Let $\mathcal{O}_1$ and $\mathcal{O}_2$ be locally convex topologies on a linear space X, such that $(X, \mathcal{O}_1)' = (X, \mathcal{O}_2)'$. Let further $C \subset X$ be convex. Then C is $\mathcal{O}_1$-closed if and only if it is $\mathcal{O}_2$-closed.*

Proof. Let C be $\mathcal{O}_2$-closed and $x_0 \in X \setminus C$. By the Hahn-Banach Theorem 5.17 there exists $f \in (X, \mathcal{O}_2)' = (X, \mathcal{O}_1)'$ and $\varepsilon > 0$ such that $\operatorname{Re} f(x_0) + \varepsilon < \operatorname{Re} f(y)$ for all $y \in C$. Put $U := \{x \in X : \operatorname{Re} f(x) < \varepsilon\} \in \mathcal{O}_1 \cap \mathcal{O}_2$. If $x \in (x_0 + U) \cap C$, then $x = x_0 + u$ with some $u \in U$, and thus $\operatorname{Re} f(x) = \operatorname{Re} f(x_0) + \operatorname{Re} f(u) < \operatorname{Re} f(x_0) + \varepsilon < \operatorname{Re} f(x)$. This contradiction gives $(x_0 + U) \cap C = \varnothing$. Hence, $x_0 \notin \overline{C}^{\mathcal{O}_1}$. This shows $\overline{C}^{\mathcal{O}_1} \subset C \subset \overline{C}^{\mathcal{O}_1}$. □

Corollary 5.35. *Let $(X, \mathcal{O})$ be a LCS and $C \subset X$ convex. Then $\overline{C}^{\mathcal{O}} = \overline{C}^{\sigma(X,X')}$.*

Proof. We have $(X, \sigma(X, X'))' = (X, \mathcal{O})'$. As $\overline{C}^{\sigma(X,X')}$ is convex and $\sigma(X, X')$-closed, it is $\mathcal{O}$-closed. Hence, $C \subset \overline{C}^{\sigma(X,X')}$ implies $\overline{C}^{\mathcal{O}} \subset \overline{C}^{\sigma(X,X')}$. The converse inclusion follows analogously. □

Corollary 5.36. *Let X be a normed space and $C \subset X$ convex. Then the following statements are equivalent.*

(i) *C is closed.*

(ii) *C is weakly closed.*

(iii) *C is weakly sequentially closed, i.e., if $(x_n) \subset C$, $x_n \xrightarrow{w} x$, then $x \in C$.*

Proof. (i)⇔(ii) follows from Proposition 5.34, and (ii)⇒(iii) is trivial. To prove (iii)⇒(i), let $(x_n) \subset C$ with $x_n \to x \in X$. Then $x_n \xrightarrow{w} x$ and thus $x \in C$. □

We mention that the above corollary also holds in the more general situation when X is a metrizable LCS. The proof remains the same since a set in a metric space is closed if and only if it is sequentially closed.

Corollary 5.37. *Let $T : X \supset \operatorname{dom} T \to Y$ be a linear operator between two normed spaces X and Y. Then T is closed if and only if for any $(x_n) \subset \operatorname{dom} T$ with $x_n \xrightarrow{w} x \in X$ and $Tx_n \xrightarrow{w} y \in Y$ we have $x \in \operatorname{dom} T$ and $Tx = y$.*

Proof. T is closed if and only if its graph $G(T)$ is closed. By Corollary 5.36, this is the case if and only if $G(T)$ is weakly sequentially closed, which is equivalent to the condition in the corollary. □

Definition 5.38. Let (X, Y) be a dual pair, $A \subset X$ and $B \subset Y$. The **polar set** of A is defined by

$$A^\circ := \{y \in Y : \operatorname{Re}\langle x, y\rangle \le 1 \,\forall\, x \in A\},$$

and the polar set of B is given by

$$B^\circ := \{x \in X : \operatorname{Re}\langle x, y\rangle \le 1 \,\forall\, y \in B\}.$$

If in the above definitions $\operatorname{Re}(\cdot)$ is replaced by $|\cdot|$, we call the corresponding sets **absolute polar set** of A and B, respectively. We will denote them by $A^{|\circ|}$ and $B^{|\circ|}$.

In sequel, we denote the closed unit ball of a normed space X by B_X, i.e., $B_X := \{x \in X : \|x\| \le 1\}$.

Example 5.39. Let X be a normed space and $Y = X'$. Then

$$B_X^{|\circ|} = \{f \in X' : |f(x)| \le 1 \,\forall x \in B_X\} = B_{X'}.$$

Moreover, as $\|x\| = \sup\{|f(x)| : f \in B_{X'}\}$ for $x \in X$,

$$B_{X'}^{|\circ|} = \{x \in X : |f(x)| \le 1 \,\forall f \in B_{X'}\} = B_X.$$

In the following lemma, we collect some useful properties of polar sets.

Lemma 5.40. *Let (X, Y) be a dual pair, $A \subset X$, and $(A_i)_{i\in I} \subset \mathfrak{P}(X)$. Then the following statements hold.*

(i) *A° is convex and $\sigma(Y, X)$-closed, and $A^\circ = (\overline{\operatorname{co}} A)^\circ$.*

(ii) *$0 \in A^\circ$, $A \subset A^{\circ\circ} = (A^\circ)^\circ$ and if $A_1 \subset A_2$, then $A_2^\circ \subset A_1^\circ$.*

(iii) *If A is balanced, then $A^\circ = A^{|\circ|}$.*

(iv) *$(\lambda A)^\circ = \lambda^{-1} A^\circ$ for all $\lambda > 0$.*

(v) *$\left(\bigcup_{i\in I} A_i\right)^\circ = \bigcap_{i\in I} A_i^\circ$.*

(vi) *$\left(\bigcap_{i\in I} A_i\right)^\circ \supset \overline{\operatorname{co}}\left(\bigcup_{i\in I} A_i^\circ\right)$.*

Proof. We leave the proofs of (ii), (iv), and (v) to the reader as they are straightforward.

(i). We have $A^\circ = \bigcap_{x\in A} A_x$, where $A_x = \{y \in Y : \operatorname{Re}\langle x, y\rangle \le 1\}$, $x \in A$. It is easy to see that each A_x is convex and $\sigma(Y, X)$-closed. This also implies these properties for A°.

We will now show, that $A^\circ = \overline{A}^\circ$ and $A^\circ = (\operatorname{co} A)^\circ$. This implies $A^\circ = (\operatorname{co} A)^\circ = (\overline{\operatorname{co}} A)^\circ$. First, since $A \subset \overline{A}$ and $A \subset \operatorname{co} A$, we have $\overline{A}^\circ \subset A^\circ$ and $(\operatorname{co} A)^\circ \subset A^\circ$ by (ii). Further, it is easy to see that $A^\circ \subset (\operatorname{co} A)^\circ$. Let $y \in A^\circ$. Put $U := \{u \in X : \operatorname{Re}\langle u, y\rangle > 1\}$. Then U is $\sigma(X, Y)$-open and $U \cap A = \varnothing$. Hence, $x \in \overline{A}$ implies $x \notin U$ and thus $\operatorname{Re}\langle x, y\rangle \le 1$. This yields $y \in \overline{A}^\circ$.

(iii). It is clear that $A^{|\circ|} \subset A^\circ$. Let $y \in A^\circ$. If $x \in A$, then, by writing $\langle x, y\rangle = e^{it}|\langle x, y\rangle|$, we can rewrite $|\langle x, y\rangle| = \operatorname{Re}\left\langle e^{-it}x, y\right\rangle \le 1$, since by balancedness of A, we have $e^{-it}x \in A$.

(vi). Put $A := \bigcap_{i\in I} A_i$. Then $A \subset A_i$ for all $i \in I$. Thus $A_i^\circ \subset A^\circ$ for all $i \in I$, and thus, by (v), $\bigcup_{i\in I} A_i^\circ \subset A^\circ$. Hence,

$$\overline{\text{co}}\left(\bigcup_{i\in I} A_i^\circ\right) \subset \overline{\text{co}} A^\circ = A^\circ = \left(\bigcap_{i\in I} A_i\right)^\circ .$$

The lemma is proved. □

Theorem 5.41 (Bipolar Theorem). *Let (X, Y) be a dual pair and $A \subset X$. Then*

$$A^{\circ\circ} = \overline{\text{co}}\,(A \cup \{0\}) .$$

Proof. It is evident that $A^\circ = (A \cup \{0\})^\circ$. By Lemma 5.40, we have $\overline{\text{co}}(A \cup \{0\}) \subset (\overline{\text{co}}(A \cup \{0\}))^{\circ\circ} = (A \cup \{0\})^{\circ\circ} = A^{\circ\circ}$.

We write $V := \overline{\text{co}}(A \cup \{0\})$. Suppose that there exists $x_0 \in A^{\circ\circ} \setminus V$. By the Hahn-Banach Theorem 5.25, there exists some $f \in (X, \sigma(X, Y))'$ and $\gamma \in \mathbb{R}$ such that $\operatorname{Re} f(x) \le \gamma < \operatorname{Re} f(x_0)$ for all $x \in V$. As $0 \in V$, we have $\gamma \ge 0$ and thus $\operatorname{Re} f(x_0) > 0$. It is no restriction to assume that $\gamma > 0$ or even $\gamma = 1$:

$$\operatorname{Re} f(x) \le 1 < \operatorname{Re} f(x_0) \quad \text{for all } x \in V.$$

By Corollary 5.32, $f = \langle \cdot, y_0\rangle$ for some $y_0 \in Y$. Thus $\operatorname{Re}\langle x, y_0\rangle \le 1 < \operatorname{Re}\langle x_0, y_0\rangle$ for all $x \in A \subset V$. The first inequality yields $y_0 \in A^\circ$. Thus, the second gives that $x_0 \notin A^{\circ\circ}$, a contradiction. □

Corollary 5.42. *Let (X, Y) be a dual pair, $C \subset X$ convex, and $0 \in C$. Then C is $\sigma(X, Y)$-closed if and only if $C = B^\circ$ for some $B \subset Y$.*

Proof. Set $B := C^\circ$. Then $B^\circ = C^{\circ\circ} = \overline{\text{co}}(C \cup \{0\}) = \overline{\text{co}}(C) = \overline{C}$. □

The following Banach-Alaoglu Theorem is the generalization to locally convex spaces of Alaoglu's theorem on normed spaces (cf. [5]).

Theorem 5.43 (Banach-Alaoglu). *Let X be a LCS and U a neighborhood of zero. Then U° is $\sigma(X', X)$-compact.*

Proof. First of all, we choose a balanced, convex open neighborhood V of zero with $V \subset U$. Then $V^\circ = V^{|\circ|}$ (see Lemma 5.40) and $U^\circ \subset V^\circ$. So, U° is compact if V° is. Hence, let us show that $V^{|\circ|}$ is $\sigma(X', X)$-compact.

A neighborhood of zero is absorbing, so that for all $x \in X$ there exists $\gamma(x) > 0$ such that $x \in \gamma(x)V$. Let $x \in X$ and $f \in V^{|\circ|}$. Then we have $x = \gamma(x)v$ for some $v \in V$, and thus $|f(x)| = \gamma(x)|f(v)| \le \gamma(x)$. Put $D_x := \{\alpha \in \mathbb{K} : |\alpha| \le \gamma(x)\}$, $x \in X$. These D_x are compact in $\mathbb{K}$ and $f(x) \in D_x$ whenever $f \in V^{|\circ|}$ and $x \in X$.

Define $P := \prod_{x\in X} D_x$. By Tychonov's Theorem (cf. [5]), P is compact with respect to the product topology. Note that P can be seen as the space of functions $f : X \to \mathbb{K}$ with $f(x) \in D_x$ for each $x \in X$. Hence, by the above, we have $V^{|\circ|} \subset X' \cap P$.

Let us prove the following claims:

(i) The trace topologies of $\mathcal{O}_P$ and $\sigma(X', X)$ on $V^{|\circ|}$ coincide.

(ii) $V^{|\circ|}$ is $\mathcal{O}_P$-closed.

Then, since P is compact, $V^{|\circ|}$ is $\mathcal{O}_P$-compact. By (i), it follows that $V^{|\circ|}$ is also $\sigma(X', X)$-compact.

(i). For $x_0 \in X$, $z_0 \in \mathbb{K}$, and $\varepsilon > 0$, define $W_1 := \{f \in X' : |f(x_0) - z_0| < \varepsilon\}$ and $W_2 := \{h \in P : |h(x_0) - z_0| < \varepsilon\}$. Note that the sets of the form W_1 form a subbasis of $\sigma(X', X)$, and the sets of the form W_2 form a subbasis of $\mathcal{O}_P$. Moreover, $W_1 \cap V^{|\circ|} = W_2 \cap V^{|\circ|}$ since $V^{|\circ|} \subset X' \cap P$. This implies (i).

(ii). Let f_0 be an element of the $\mathcal{O}_P$-closure of $V^{|\circ|}$. Let $x, y \in X$, $\alpha, \beta \in \mathbb{K}$, $\varepsilon > 0$, and define

$$U := \{h \in P : |h(z) - f_0(z)| < \varepsilon \; \forall z \in \{x, y, \alpha x + \beta y\}\}.$$

U is an $\mathcal{O}_P$-open neighborhood of f_0. Thus, $V^{|\circ|} \cap U \neq \varnothing$. Any $f \in V^{|\circ|} \cap U$ is linear, hence

$$f_0(\alpha x + \beta y) - \alpha f_0(x) - \beta f_0(y) = (f_0 - f)(\alpha x + \beta y) - \alpha(f_0 - f)(x) - \beta(f_0 - f)(y).$$

This implies $|f_0(\alpha x + \beta y) - \alpha f_0(x) - \beta f_0(y)| \leq (1 + |\alpha| + |\beta|)\varepsilon$. This shows that f_0 is in fact linear. Let $x \in V$, $\varepsilon > 0$. As $U \cap V^{|\circ|} \neq \varnothing$, there exists $f \in V^{|\circ|}$ such that $|f(x) - f_0(x)| < \varepsilon$. Thus, $|f_0(x)| \leq 1 + \varepsilon$ which further implies $|f_0(x)| \leq 1$. By Theorem 4.33, $f_0 \in X'$ and thus $f_0 \in V^{|\circ|}$. Thus, $V^{|\circ|}$ is $\mathcal{O}_P$-closed. □

Alaoglu's Theorem is now a simple corollary of the Banach-Alaoglu Theorem.

Corollary 5.44 (Alaoglu). *Let X be a normed space. Then $B_{X'} = \{f \in X' : \|f\| \leq 1\}$ is weak-$*$-compact.*

Proof. Since B_X is balanced, we have $B_X^{\circ} = B_X^{|\circ|} = B_{X'}$, see Example 5.39. □

Theorem 5.45. *Every Banach Space X is isometrically isomorphic to a closed subspace of $C(T)$, where T is some compact topological space.*

Proof. Put $T = (B_{X'}, \sigma(X', X))$. This space is compact. Now, define $\Phi : X \to C(T)$ by $(\Phi x)(f) := f(x)$, $x \in X$, $f \in T = B_{X'}$. Let us compute the norm of Φ:

$$\|\Phi x\|_\infty = \sup\{|(\Phi x)(f)| : f \in B_{X'}\} = \sup\{|f(x)| : f \in B_{X'}\} = \|x\|.$$

Thus, Φ is isometric and thus $X \cong \operatorname{ran} \Phi$, which is a closed subspace of $C(T)$ since X is complete. □

Let X be a normed space. It was proved in [5] that the canonical embedding $i : X \to X''$, defined by

$$(ix)f := f(x), \quad x \in X,\ f \in X',$$

is an isometry. Therefore, if X is a Banach space, we have $X \cong i(X)$.

Theorem 5.46 (Goldstine). *If X is a normed space then $i(B_X)$ is $\sigma(X'', X')$-dense in $B_{X''}$. As a consequence, $i(X)$ is $\sigma(X'', X')$-dense in X''.*

Remark 5.47. If X is a Banach space, then $i(X)$ and $i(B_X)$ are $\|\cdot\|$-closed. They are in general not closed in the weak-$*$-topology.

Proof of Theorem 5.46. As i is an isometry, we have $i(B_X) \subset B_{X''}$. We will now prove that with respect to the dual pair (X'', X') we have

$$(i(B_X))^\circ = B_{X'} \quad \text{and} \quad B_{X'}^\circ = B_{X''}.$$

Then, as $0 \in i(B_X)$ and $i(B_X)$ is convex, the Bipolar Theorem 5.41 yields

$$\overline{i(B_X)}^{\sigma(X'',X')} = \overline{\operatorname{co}}^{\sigma(X'',X')}(i(B_X) \cup \{0\}) = (i(B_X))^{\circ\circ} = B_{X''}.$$

Indeed, $B_{X'}^\circ = B_{X''}$ is proved as in Example 5.39 and

$$\begin{aligned}(i(B_X))^\circ &= \{f \in X' : |\varphi f| \leq 1 \ \forall \varphi \in i(B_X)\} = \{f \in X' : |f(x)| \leq 1 \ \forall x \in B_X\} \\ &= \{f \in X' : \|f\| \leq 1\} = B_{X'}.\end{aligned}$$

Finally, $i(X)$ is $\sigma(X'', X')$-dense in X'' since $B_{X''}$ is absorbing in X''. □

Lemma 5.48. *The mapping* $i : (X, \sigma(X, X')) \to (i(X), \sigma(X'', X'))$ *is a linear homeomorphism.*

Proof. Let q be a $\sigma(X'', X')$-continuous seminorm on X''. By Lemma 5.13, there exist $M \geq 0$ and $f_1, \ldots, f_n \in X'$ such that for all $\varphi \in X''$ we have $q(\varphi) \leq M \cdot \max_{i=1,\ldots,n} |\varphi f_i|$. This implies

$$q(ix) \leq M \cdot \max_{i=1,\ldots,n} |f_i(x)|, \quad x \in X.$$

Hence, $q \circ i$ is continuous, and so is i by Theorem 5.14. Now, let $f \in X'$, $\varepsilon > 0$, and $U := \{x \in X : |f(x)| < \varepsilon\}$. Then U is weakly open, and

$$i(U) = \{ix : |(ix)f| < \varepsilon\} = \{\varphi \in X'' : |\varphi f| < \varepsilon\} \cap i(X),$$

is (relatively) $\sigma(X'', X')$-open in $i(X)$. Thus, i^{-1} is also continuous. □

Theorem 5.49. *A Banach space* X *is reflexive if and only if* B_X *is* $\sigma(X, X')$*-compact.*

Proof. If X is reflexive, then $i(B_X) = B_{X''}$ and thus $B_X = i^{-1}(B_{X''})$. By Alaoglu's Theorem 5.44, $B_{X''}$ is $\sigma(X'', X')$-compact.

Now, let B_X be $\sigma(X, X')$-compact. Thus, $i(B_X)$ is $\sigma(X'', X')$-compact, and by Goldstine's Theorem 5.46 we get that $i(B_X) = B_{X''}$, implying that X is reflexive. □

6. An Introduction to Distribution Theory

Let $\Omega \subset \mathbb{R}^n$ be open. Throughout this chapter, we shall use the notation $\mathcal{D}(\Omega) := C_c^\infty(\Omega)$. For $f \in L^1_{\mathrm{loc}}(\Omega)$, define the linear functional $T_f : \mathcal{D}(\Omega) \to \mathbb{C}$ by

$$T_f\varphi := \int_\Omega \varphi f \,\mathrm{d}x.$$

Note that $f \mapsto T_f$ is injective by Lemma 3.6. Hence, every L^1_{loc}-function on Ω can be identified with a linear functional on $\mathcal{D}(\Omega)$.

A Problem from Physics: Let $\varphi \in \mathcal{D}(\Omega)$. Due to Theorem 3.3, we have that $\|\delta_k * \varphi - \varphi\|_p \to 0$ as $k \to \infty$ for any $p \in [1, \infty)$. It can be shown, that this extends to $p = \infty$ (because has φ has "good properties"). In particular, $\delta_k * \varphi \to \varphi$ pointwise. Since

$$(\delta_k * \varphi)(0) = \int_{\mathbb{R}^n} \delta_k(-y)\varphi(y)\,\mathrm{d}y = \int_{\mathbb{R}^n} \delta_k\varphi\,\mathrm{d}y,$$

we have for all $\varphi \in \mathcal{D}(\Omega)$:

$$\int_{\mathbb{R}^n} \delta_k\varphi\,\mathrm{d}y \to \varphi(0), \qquad k \to \infty.$$

Having in mind that δ_k tends pointwise to the "function" δ taking on the value ∞ at zero and vanishing everywhere else, many physicists write $\int \delta\varphi\,\mathrm{d}x = \varphi(0)$ for all $\varphi \in \mathcal{D}(\Omega)$. This is of course not correct since there is no function with the latter property. In fact, $\int \delta\varphi\,\mathrm{d}x = 0$ as the integrand is only non-zero on a set of measure zero.

However, physicists continue calculating with these kind of integrals. What they actually mean by $\int \delta\varphi\,\mathrm{d}x$ is the value $\varphi(0)$. Hence, they are dealing with the evaluation *functional* $\varphi \mapsto \varphi(0)$, called the δ-functional. [8]

In what follows, we are going to consider functionals on $\mathcal{D}(\Omega)$. Then we have both generalized functions and given a meaning to the weird δ-function. In order to apply methods from functional analysis to such functionals, we shall define a topology on $\mathcal{D}(\Omega)$. The functionals T_f and the δ-functional will turn out to be continuous with respect to that topology.

A Topology on $\mathcal{D}(\Omega)$: By $\mathrm{Comp}(\Omega)$ denote the set of all compact subsets of Ω. For $K \in \mathrm{Comp}(\Omega)$, define

$$\mathcal{D}_K(\Omega) := \{f \in \mathcal{D}(\Omega) : \operatorname{supp} f \subset K\}.$$

Note that each $\mathcal{D}_K(\Omega)$ is a subspace of $\mathcal{D}(\Omega)$. Now, for $m \in \mathbb{N}_0$, define the norm $p_m : \mathcal{D}(\Omega) \to [0, \infty)$ by

$$p_m(\varphi) := \max_{|\alpha| \le m} \sup_{x \in \Omega} |(\partial^\alpha\varphi)(x)|, \quad \varphi \in \mathcal{D}(\Omega).$$

[8] Another approach to get along with that integral is to treat $\delta\,\mathrm{d}x$ as the differential $\mathrm{d}\mu$ of the point measure μ with value 1 at zero. That is, $\int \delta\varphi\,\mathrm{d}x = \int \varphi\,\mathrm{d}\mu = \varphi(0)$. But this is only meaningful if μ is absolutely continuous with respect to the Lebesgue measure, which it is not.

Then, for each $K \in \mathrm{Comp}(\Omega)$, the family $P_K := \{p_m|_{\mathcal{D}_K(\Omega)} : m \in \mathbb{N}_0\}$ is separating on $\mathcal{D}_K(\Omega)$ (why?). Thus, $(D_K(\Omega), \mathcal{O}_{P_K})$ is a locally convex space. Theorem B.8 shows that $\mathcal{D}_K(\Omega)$ is even metrizable.

To generate a locally convex topology on $\mathcal{D}(\Omega)$, we now define the following family of seminorms:

$$P := \left\{p : \mathcal{D}(\Omega) \to [0, \infty) \,\middle|\, p \text{ seminorm, } p|_{\mathcal{D}_K(\Omega)} \text{ is } O_{P_K}\text{-continuous } \forall K \in \mathrm{Comp}(\Omega)\right\}.$$

Since for each $m \in \mathbb{N}_0$ and each $K \in \mathrm{Comp}(\Omega)$, the norm $p_m|\mathcal{D}_K(\Omega)$ is $\mathcal{O}_{P_K}$-continuous (as it generates $\mathcal{O}_{P_K}$), we have that $p_m \in P$ for all $m \in \mathbb{N}_0$. Hence, also P is separating. Thus, $(\mathcal{D}(\Omega), \mathcal{O}_P)$ is a locally convex space. We set $\mathcal{O}(\Omega) := \mathcal{O}_P$.

Remark 6.1. (a) As $p_0 \leq p_1 \leq \ldots$, we have $p \in P$ if and only if (cf. Lemma 5.13) for all $K \in \mathrm{Comp}(\Omega)$ there exist $M > 0$ and $m \in \mathbb{N}_0$ such that for all $\varphi \in \mathcal{D}_K(\Omega)$ we have $p(\varphi) \leq M p_m(\varphi)$.
(b) With (a) and Lemma 5.13 it is easily seen that a seminorm p on $\mathcal{D}(\Omega)$ is continuous if and only if $p \in P$.

Lemma 6.2. *For $K \in \mathrm{Comp}(\Omega)$, the following statements hold.*

(i) *The trace topology of $\mathcal{O}(\Omega)$ on $\mathcal{D}_K(\Omega)$ is $\mathcal{O}_{P_K}$.*

(ii) *$\mathcal{D}_K(\Omega)$ is an $\mathcal{O}(\Omega)$-closed subspace of $\mathcal{D}(\Omega)$.*

Proof. (i). Let $V \in \mathcal{O}(\Omega)$ with $0 \in V$. Then there exist $\varepsilon > 0$, $q_1, \ldots, q_r \in P$ such that $U_\varepsilon := \{\varphi \in \mathcal{D}(\Omega) : \max_i q_i(\varphi) < \varepsilon\} \subset V$. Let $K \in \mathrm{Comp}(\Omega)$. Then for each $i = 1, \ldots, r$, there exist $M_i > 0$ and $m_i \in \mathbb{N}_0$ such that for all $\varphi \in \mathcal{D}_K(\Omega) : q_i(\varphi) \leq M_i p_{m_i}(\varphi)$. Set $M := \max_{i=1,\ldots,r} M_i$ and $m := \max_{i=1,\ldots,r} m_i$. Then $M_i p_{m_i} \leq M p_m$ $\forall i$. Thus,

$$U_\varepsilon \cap \mathcal{D}_K(\Omega) \supset \left\{\varphi \in \mathcal{D}_K(\Omega) : p_m(\varphi) < \frac{\varepsilon}{M}\right\} =: U \in \mathcal{O}_{P_K}.$$

Hence, every relatively open set is $\mathcal{O}_{P_K}$-open.

Conversely, let $U \in \mathcal{O}_{P_K}$ with $0 \in U$. Then there exist $\varepsilon > 0$ and $m \in \mathbb{N}_0$ such that

$$U_0 := \{\varphi \in \mathcal{D}_K(\Omega) : p_m(\varphi) < \varepsilon\} \subset U.$$

But as $p_m \in P$, U_0 is relatively open in $\mathcal{D}_K(\Omega)$. This completes the proof of (i).

(ii). For $x \in \Omega$ define the seminorm $p_x : \mathcal{D}(\Omega) \to [0, \infty)$ by $p_x(\varphi) := |\varphi(x)|$, $\varphi \in \mathcal{D}(\Omega)$. Then $p_x(\varphi) = |\varphi(x)| \leq p_0(\varphi)$ holds for all $\varphi \in \mathcal{D}(\Omega)$. From Remark 6.1 we conclude that $p_x \in P$ for all $x \in \Omega$. Thus,

$$D_K(\Omega) = \{\varphi \in \mathcal{D}(\Omega) : p_x(\varphi) = 0 \;\forall x \notin K\} = \bigcap_{x \in \Omega \setminus K} p_x^{-1}(\{0\})$$

is closed. □

Lemma 6.3. *Let Y be a LCS and $L : \mathcal{D}(\Omega) \to Y$ linear. Then L is continuous if and only if $L|_{\mathcal{D}_K(\Omega)}$ is continuous for each $K \in \mathrm{Comp}(\Omega)$.*

Proof. If L is continuous, then for an open set $V \subset Y$, $\left(L|_{\mathcal{D}_K(\Omega)}\right)^{-1}(V) = L^{-1}(V) \cap \mathcal{D}_K(\Omega)$ is open in $\mathcal{D}_K(\Omega)$ by Lemma 6.2(i). Let $L|_{D_K(\Omega)}$ be continuous for all $K \in \mathrm{Comp}(\Omega)$, and let q be a continuous seminorm on Y. Then $(q \circ L)|_{\mathcal{D}_K(\Omega)} = q \circ \left(L|_{\mathcal{D}_K(\Omega)}\right)$ is continuous for all $K \in \mathrm{Comp}(\Omega)$. Thus, $q \circ L \in P$. From Theorem 5.14 we conclude that L is continuous. □

Theorem 6.4. *Let $(\varphi_n) \subset \mathcal{D}(\Omega)$. Then the following statements are equivalent.*

(i) $\varphi_n \to 0$ *in* $\mathcal{D}(\Omega)$.

(ii) *There exists* $K \in \mathrm{Comp}(\Omega)$ *such that* $\varphi_n \in \mathcal{D}_K(\Omega)$ *for all* $n \in \mathbb{N}$, *and* $\varphi_n \to 0$ *in* $\mathcal{D}_K(\Omega)$.

Proof. (ii)⇒(i) follows directly from Lemma 6.2(i). To prove (i)⇒(ii), assume that $\varphi_n \to 0$ in $\mathcal{D}(\Omega)$, and suppose there is no K as in (ii). Let $(K_n) \subset \mathrm{Comp}(\Omega)$ with $K_n \subset K_{n+1}$ for $n \in \mathbb{N}$, and $\Omega = \bigcup_{n=1}^{\infty} \mathrm{int}(K_n)$.

Now, choose subsequences (ψ_n) and $(\widetilde{K}_n)$ of (φ_n) and (K_n), respectively, such that $\psi_n \in \mathcal{D}_{\widetilde{K}_n}(\Omega) \setminus \mathcal{D}_{\widetilde{K}_{n-1}}(\Omega)$ holds for all $n \in \mathbb{N}$. Choose $x_n \in \operatorname{supp} \psi_n \setminus \widetilde{K}_{n-1} \subset \widetilde{K}_n \setminus \widetilde{K}_{n-1}$. We may assume that $x_n \in \mathrm{int}(\operatorname{supp} \psi_n)$ (why?), so that $\alpha_n := |\psi_n(x_n)| > 0$. For $n \in \mathbb{N}$, define $q_n : \mathcal{D}(\Omega) \to [0, \infty]$ by $q_n(\varphi) := \alpha_n^{-1} |\varphi(x_n)|$, which is a continuous seminorm, and put $q := \sum_{n=1}^{\infty} q_n$. By construction, only a finite number of summands will be non-zero in this sum.

If $K \in \mathrm{Comp}(\Omega)$, then there exists $N \in \mathbb{N}$ such that $K \subset \widetilde{K}_N$. Thus, $q|_{\mathcal{D}_K(\Omega)} = \sum_{i=1}^{N} q_n|_{\mathcal{D}_K(\Omega)}$ is continuous on $\mathcal{D}_K(\Omega)$, which implies $q \in P$. Further, $q(\psi_n) \geq q_n(\psi_n) = 1$. But $\varphi_n \to 0$ implies $\psi_n \to 0$ and thus $q(\psi_n) \to q(0) = 0$. A contradiction. □

Definition 6.5. The dual space $\mathcal{D}'(\Omega) := (\mathcal{D}(\Omega), \mathcal{O}(\Omega))'$ is called the **space of distributions** on Ω. The elements of $\mathcal{D}'(\Omega)$ are called **distributions**.

Theorem 6.6. *For a linear functional $T : \mathcal{D}(\Omega) \to \mathbb{C}$, the following statements are equivalent.*

(i) $T \in \mathcal{D}'(\Omega)$.

(ii) *For each* $K \in \mathrm{Comp}(\Omega)$ *we have* $T|_{\mathcal{D}_K(\Omega)} \in (D_K(\Omega))'$.

(iii) *For each* $K \in \mathrm{Comp}(\Omega)$ *there exist* $m \in \mathbb{N}_0$ *and* $C \geq 0$, *such that for all* $\varphi \in \mathcal{D}_K(\Omega)$ *we have*

$$T\varphi \leq C p_m(\varphi).$$

(iv) *For any sequence* $(\varphi_n) \subset \mathcal{D}(\Omega)$ *with* $\varphi_n \to 0$ *we have* $T\varphi_n \to 0$ *in* $\mathbb{C}$.

Proof. (i)⇔(iii) is due to Lemma 6.3. (ii)⇔(iii) follows from Theorem 5.14, since seminorms on $\mathbb{C}$ have form $p(z_0) = p_0|z|$ with some $p_0 \in [0, \infty)$. (ii)⇒(iv) is Theorem 6.4.

(iv)⇒(ii). Let $K \in \mathrm{Comp}(\Omega)$. From (iv) and Lemma 6.2(i) we conclude that $T|_{\mathcal{D}_K(\Omega)}$ is sequentially continuous. But $\mathcal{D}_K(\Omega)$ is metrizable (cf. Theorem B.8). Hence, $T|_{\mathcal{D}_K(\Omega)}$ is continuous. □

Definition 6.7. If in (iii) of Theorem 6.6 $m \in \mathbb{N}_0$ can be chosen independently of K, then T is said to be **of finite order**. The smallest such m is called the **order** of T and will be denoted by $\operatorname{ord}(T)$. Otherwise, we say that T is of **infinite order** and write $\operatorname{ord}(T) = \infty$.

Example 6.8. (a) As in the introduction of this chapter, for $f \in L^1_{\text{loc}}(\Omega)$, define the linear functional

$$T_f\varphi := \int_\Omega f\varphi \,\mathrm{d}x, \qquad \varphi \in \mathcal{D}(\Omega).$$

If $K \in \operatorname{Comp}(\Omega)$ and $\varphi \in \mathcal{D}_K(\Omega)$, then $|T_f\varphi| \leq \int_K |f||\varphi| \,\mathrm{d}x \leq \left(\int_K |f| \,\mathrm{d}x\right) \cdot p_0(\varphi)$. Thus, by Theorem 6.6 we have $T_f \in \mathcal{D}'(\Omega)$ and $\operatorname{ord}(T) = 0$. The distributions of type T_f are called *regular distributions*.

(b) Let $x_0 \in \Omega$ and define the *δ-functional* at x_0 by

$$\delta_{x_0}\varphi := \varphi(x_0), \qquad \varphi \in \mathcal{D}(\Omega).$$

If $K \in \operatorname{Comp}(\Omega)$ and $\varphi \in \mathcal{D}_K(\Omega)$, then $|\delta_{x_0}\varphi| = |\varphi(x_0)| \leq p_0(\varphi)$. This shows $\delta_{x_0} \in \mathcal{D}'(\Omega)$ and $\operatorname{ord}(\delta_{x_0}) = 0$.

(c) Let μ be a Borel measure on Ω such that $\mu(K) < \infty$ for every compact $K \in \operatorname{Comp}(\Omega)$. Put $T_\mu\varphi := \int_\Omega \varphi \,\mathrm{d}\mu$, $\varphi \in \mathcal{D}(\Omega)$. Then for $K \in \operatorname{Comp}(\Omega)$ and $\varphi \in \mathcal{D}_K(\Omega)$ we have

$$|T_\mu\varphi| \leq \int_K |\varphi| \,\mathrm{d}\mu \leq \mu(K)p_0(\varphi).$$

Thus $T_\mu \in \mathcal{D}'(\Omega)$ and $\operatorname{ord}(T) = 0$.

(d) Let $x_0 \in \Omega \subset \mathbb{R}$, and set $T\varphi := \varphi'(x_0)$ for $\varphi \in \mathcal{D}(\Omega)$. For $K \in \operatorname{Comp}(\Omega)$ and $\varphi \in \mathcal{D}_K(\Omega)$ we have $|T\varphi| = |\varphi'(x_0)| \leq p_1(\varphi)$. This shows that $T \in \mathcal{D}'(\Omega)$ and $\operatorname{ord}(T) = 1$.

(e) Let $\Omega = \mathbb{R}$. We define $T : \mathcal{D}(\Omega) \to \mathbb{C}$ by

$$T\varphi := \sum_{i=0}^{\infty} \varphi^{(n)}(n), \qquad \varphi \in \mathcal{D}(\Omega).$$

This sum is always finite, since the elements of $\mathcal{D}(\Omega)$ have compact support. For any $K \in \operatorname{Comp}(\Omega)$ we find $N \in \mathbb{N}$ such that $K \subset [-N, N]$. Thus, for $\varphi \in \mathcal{D}_K(\Omega)$,

$$|T\varphi| \leq \sum_{n=0}^{N-1} \left|\varphi^{(n)}(n)\right| \leq Np_{N-1}(\varphi).$$

Thus, $T \in \mathcal{D}'(\Omega)$ and $\operatorname{ord}(T) = \infty$.

Definition 6.9. For $T \in \mathcal{D}'(\Omega)$ and $\alpha \in \mathbb{N}_0^n$, define the α-th **derivative** of T by

$$(\partial^\alpha T)\varphi := (-1)^{|\alpha|}T(\partial^\alpha\varphi), \qquad \varphi \in \mathcal{D}(\Omega).$$

It is clear that T is well defined as a linear functional on $\mathcal{D}(\Omega)$. The next lemma shows that T is continuous.

Lemma 6.10. *For $T \in \mathcal{D}'(\Omega)$ and $\alpha \in \mathbb{N}_0^n$ we have $\partial^\alpha T \in \mathcal{D}'(\Omega)$.*

Proof. If $K \in \mathrm{Comp}(\Omega)$, then there exist $C > 0$ and $m \in \mathbb{N}_0$ such that $|T\varphi| \leq Cp_m(\varphi)$ for all $\varphi \in \mathcal{D}_K(\Omega)$. Hence,

$$|(\partial^\alpha T)\,\varphi| = |T\,(\partial^\alpha \varphi)| \leq Cp_m\,(\partial^\alpha \varphi) \leq Cp_{m+|\alpha|}(\varphi).$$

Thus, $\partial^\alpha T \in \mathcal{D}'(\Omega)$. □

Example 6.11. (a) If the α-th weak derivative of $f \in L^1_{\mathrm{loc}}(\Omega)$ exists, then for all $\varphi \in \mathcal{D}(\Omega)$ we have

$$(\partial^\alpha T_f)\varphi = (-1)^{|\alpha|}T_f(\partial^\alpha\varphi) = (-1)^{|\alpha|}\int_\Omega f(\partial^\alpha\varphi)\,\mathrm{d}x = \int_\Omega (\partial^\alpha f)\varphi\,\mathrm{d}x = T_{\partial^\alpha f}\varphi,$$

i.e. $\partial^\alpha T_f = T_{\partial^\alpha f}$. Hence, a derivative of a regular distribution is again a regular distribution. That is, the distributional derivative generalizes the weak derivative. Note that the distributional derivative always exists.
(b) Let $H = \chi_{[0,\infty)}$ be the Heavyside function on $\Omega = \mathbb{R}$. Then $H \in L^1_{\mathrm{loc}}(\Omega)$. We have

$$T_H'\varphi = -T_H\varphi' = -\int_0^\infty \varphi'\,\mathrm{d}x = \varphi(0) = \delta_0\varphi$$

for all $\varphi \in \mathcal{D}(\Omega)$. Thus $T_H' = \delta_0$.
(c) Let $\alpha, \beta \in \mathbb{N}_0^n$, $T \in \mathcal{D}'(\Omega)$, and $\varphi \in \mathcal{D}(\Omega)$. Then

$$\begin{aligned}(\partial^\alpha\partial^\beta T)\varphi &= (-1)^{|\alpha|}(\partial^\beta T)(\partial^\alpha\varphi) = (-1)^{|\alpha|+|\beta|}T(\partial^\beta\partial^\alpha\varphi)\\ &= (-1)^{|\alpha+\beta|}T(\partial^{\alpha+\beta}\varphi) = (\partial^{\alpha+\beta}T)\varphi,\end{aligned}$$

i.e. $\partial^\alpha\partial^\beta T = \partial^{\alpha+\beta}T$.

Appendix

A. The Riesz Representation Theorem

In this section, we are going to prove the following version of the Riesz Representation Theorem.

Theorem A.1 (Riesz Representation Theorem). *Let $\sigma \subset \mathbb{R}$ be a compact set and*[9] *$\ell \in C(\sigma,\mathbb{R})'$ with $\ell \geq 0$ (i.e., $f \geq 0$ implies $\ell f \geq 0$). Then there exists a finite positive Borel measure μ on σ with $\mu(\sigma) = \|\ell\|$ and*

$$\ell f = \int f \,\mathrm{d}\mu, \qquad f \in C(\sigma,\mathbb{R}).$$

In the proof, we will mainly follow Riesz' original ideas for intervals $\sigma = [a,b]$, cf. [6]. The main observation is that the integral on the right hand side above is also meaningful for characteristic functions. Since the left hand side is not, one extends ℓ to a functional L on $L^\infty(\sigma,\mathbb{R})$. An approach to define the measure μ is then $\mu(\Delta) = L(\chi_\Delta)$. But this is in general not a measure. Since a measure is at least continuous from above, one at first only defines $\mu(\Delta)$ as above only for intervals $\Delta = (s,t]$ where s,t are points of continuity of $t \mapsto L(\chi_{[a,t]})$. The goal is then almost reached.

Lemma A.2. *Let $f : [a,b] \to \mathbb{R}$ be non-decreasing. Then the set of points of discontinuity of f is countable.*

Proof. For each $t \in [a,b]$, the limits

$$f(t+) := \lim_{s\searrow t} f(s) = \inf_{s\geq t} f(s) \qquad \text{and} \qquad f(t-) := \lim_{s\nearrow t} f(s) = \sup_{s\leq t} f(s)$$

obviously exist. Moreover, it is no restriction to assume that $0 \leq f(x) \leq K$ for all $x \in [a,b]$. We claim that $X_k := \left\{t \in [a,b] : f(t+) - f(t-) > \frac{1}{k}\right\}$ is finite for each $k \in \mathbb{N}$. Indeed if $x_i \in X_k$, $i = 1,\ldots,kK$, with $x_i < x_{i+1}$ for all i, then

$$\begin{aligned} K &< \sum_{i=1}^{kK}(f(x_i+) - f(x_i-)) = -f(x_1-) + \sum_{i=1}^{kK-1}(f(x_i+) - f(x_{i+1}-)) + f(x_{kK}+) \\ &\leq f(x_{kK}+) - f(x_1-) \leq f(x_{kk}+) \leq f(b). \end{aligned}$$

A contradiction! Hence, $\bigcup_{k\in\mathbb{N}} X_k$ is countable. □

Let $f : [a,b] \to \mathbb{R}$. For $s,t \in [a,b]$, $s < t$, the **variation** of f on $[s,t]$ is defined by

$$V_f(s,t) := \sup\left\{\sum_{i=1}^{n} |f(t_i) - f(t_{i-1})| : s = t_0 < t_1 < \ldots < t_n = t,\ n \in \mathbb{N}\right\}.$$

[9] We emphasize that $C(\sigma,\mathbb{R})$ is a normed space over $\mathbb{R}$.

The set of all $f : [a,b] \to \mathbb{R}$ of bounded variation, i.e., $V_f(a,b) < \infty$, is denoted by $BV[a,b]$. We note that V_f is additive, i.e.,

$$V_f(s,t) = V_f(s,\xi) + V_f(\xi,t), \qquad \xi \in (s,t).$$

This proof of this is straightforward.

Proposition A.3. *For any $f \in BV[a,b]$ there exist non-decreasing functions $g, h : [a,b] \to \mathbb{R}$ such that $f = g - h$. In particular, the points of discontinuity of f are countable.*

Proof. Set $g(t) := V_f(a,t)$ and $h(t) := V_f(a,t) - f(t)$. Due to the additivity of V_f, g is non-decreasing, and for $s < t$ we have

$$V_f(a,t) - V_f(a,s) = V_f(s,t) \geq f(t) - f(s),$$

i.e., $V_f(a,s) - f(s) \leq V_f(a,t) - f(t)$. □

Lemma A.4. *Let $f \in BV[a,b]$ be continuous at both $s \in (a,b]$ and $t \in [a,b)$. Then*

$$\lim_{\delta \searrow 0} V_f(s-\delta, s) = \lim_{\delta \searrow 0} V_f(t, t+\delta) = 0.$$

Proof. Let $\varepsilon > 0$. Choose a partition of $[t,b]$, $t = t_0 < t_1 < \ldots < t_n = b$, such that

$$V_f(t,b) - \frac{\varepsilon}{2} \leq \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})|. \tag{A.1}$$

There exists $\delta > 0$ such that $|f(t) - f(s)| < \varepsilon/2$ if $|s-t| < \delta$. If $t_1 - t \geq \delta$, choose $t_1' \in [t, t_1)$ such that $t_1' - t < \delta$. If we add this t_1' to the partition, the sum in (A.1) becomes at least larger. Therefore, we may assume that $t_1 - t < \delta$. Then, for $s \in [t, t_1]$,

$$V_f(t,b) - \frac{\varepsilon}{2} \leq \frac{\varepsilon}{2} + \sum_{i=2}^{n} |f(t_i) - f(t_{i-1})| \leq \frac{\varepsilon}{2} + V_f(s,b).$$

This implies $V_f(t,s) = V_f(t,b) - V_f(s,b) \leq \varepsilon$. □

In the proof of the following lemma, we make use of the well known correspondence between cumulative distribution functions and Borel probability measures on $\mathbb{R}$ which is actually the only somewhat deeper result from measure theory that we shall leverage to prove Theorem A.1.

Theorem A.5. *Let $f : [a,b] \to \mathbb{R}$ be non-decreasing and right-continuous, i.e., $f(t) = f(t+)$ for all $t \in [a,b)$, and $c \geq 0$. Then there exists a Borel measure μ on $[a,b]$ with*

$$\mu((s,t]) = f(t) - f(s) \quad \text{for } s,t \in [a,b],\ s < t,$$

and $\mu(\{a\}) = c$.

Proof. If f is a constant we define μ by $\mu(\Delta) := c \cdot \chi_\Delta(a)$ for Borel sets $\Delta \subset [a,b]$, which then is a measure as desired. Otherwise, $d := f(b) - f(a) > 0$. Define the function $F : \mathbb{R} \to [0,1]$ by

$$F(t) := \frac{1}{c+d} \begin{cases} 0 & \text{if } t < a, \\ f(t) - f(a) + c & \text{if } a \le t < b, \\ c + d & \text{if } t \ge b. \end{cases}$$

Then F is a cumulative distribution function (i.e., a non-decreasing, right-continuous function with $F(-\infty) = 0$ and $F(+\infty) = 1$). Hence, there exists a Borel probability measure ν on $\mathbb{R}$ such that $\nu((s,t]) = F(t) - F(s)$ for all $s, t \in \mathbb{R}$. Now it can be easily checked that the measure $\mu := (c+d) \cdot \nu|_{[a,b]}$ is as desired. □

Before we start the proof of the Riesz Representation Theorem, let us note that for a non-negative functional ℓ on $C(\sigma)$ we have $\ell f \le \ell g$ if $f \le g$. Since $f, -f \le \mathbb{1}$ for every $f \in C(\sigma)$ with $\|f\|_\infty \le 1$, this proves that $\ell \in C(\sigma)'$ and $\|\ell\| = \ell(\mathbb{1})$.

Proof of Theorem A.1. **1.** We first assume that $\sigma = [a,b]$. Let L be a Hahn-Banach extension of ℓ to $L^\infty((a,b),\mathbb{R})$, and define $\omega : [a,b] \to \mathbb{R}$ by

$$\widehat{\omega}(t) := \begin{cases} L(\chi_{[a,t]}), & t > a \\ 0, & t = a. \end{cases}$$

Then, if $a \le s < t \le b$, we have

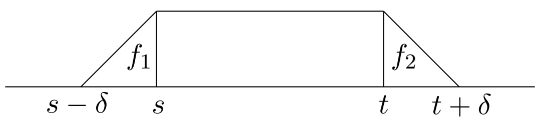
$$\widehat{\omega}(t) - \widehat{\omega}(s) = \begin{cases} L(\chi_{(s,t]}), & s > a, \\ L(\chi_{[s,t]}), & s = a. \end{cases}$$

Let $a = t_0 < t_1 < \ldots < t_n = b$. Then

$$\begin{aligned} \sum_{i=1}^{n} |\widehat{\omega}(t_i) - \widehat{\omega}(t_{i-1})| &= \left|L\left(\chi_{[a,t_1]}\right)\right| + \sum_{i=2}^{n} \left|L(\chi_{(t_{i-1},t_i]})\right| \\ &= L(\pm\chi_1 \pm \ldots \pm \chi_n) \le \|L\| = \|\ell\|. \end{aligned}$$

Hence, $\widehat{\omega} \in BV[a,b]$ and $V_{\widehat{\omega}}(a,b) \le \|\ell\|$. Let us now show that $\widehat{\omega}$ is "essentially" non-decreasing.

Let $\widehat{\omega}$ be continuous at s and t, $a < s < t < b$, and let $\varepsilon > 0$. Let $\delta > 0$ be such that $V_{\widehat{\omega}}(s-\delta, s) < \varepsilon$ and $V_{\widehat{\omega}}(t, t+\delta) < \varepsilon$. Now, let $f : [a,b] \to [0,1]$ be as in the following figure:



Approximate the non-constant parts f_1 and f_2 of f by step functions $g_1 = \sum_i \alpha_i \chi_{\Delta_i}$ and $g_2 = \sum_j \beta_j \chi_{\Gamma_j}$, respectively, such that $\|f_1 - g_1\|_\infty < \varepsilon$ and $\|f_2 - g_2\|_\infty < \varepsilon$, and put

$$g := g_1 + \chi_{(s,t]} + g_2.$$

Then $\|f - g\|_\infty \le \varepsilon$. If $\Delta_i = (a_i, b_i]$ and $\Gamma_j = (\gamma_j, \delta_j]$,

$$\begin{aligned}
0 \le \ell(f) &= L(f-g) + L(g) \le \varepsilon\|L\| + L\left(\sum \alpha_i \chi_{\Delta_i} + \chi_{(s,t]} + \sum \beta_j \chi_{\Gamma_j}\right) \\
&= \varepsilon\|L\| + \sum_i \alpha_i\left(\widehat{\omega}(b_i) - \widehat{\omega}(a_i)\right) + \widehat{\omega}(t) - \widehat{\omega}(s) + \sum_j \beta_j\left(\widehat{\omega}(\delta_j) - \widehat{\omega}(\gamma_j)\right) \\
&\le \varepsilon\|L\| + V_{\widehat{\omega}}(s-\delta, s) + V_{\widehat{\omega}}(t, t+\delta) + \widehat{\omega}(t) - \widehat{\omega}(s) \\
&< \varepsilon(\|L\| + 2) + \widehat{\omega}(t) - \widehat{\omega}(s).
\end{aligned}$$

Letting $\varepsilon \to 0$ implies $\widehat{\omega}(s) \le \widehat{\omega}(t)$. In a similar way (with $s = a$) one shows that $\widehat{\omega}(t) \ge 0$ for all points of continuity $t \in (a,b)$ of $\widehat{\omega}$. In particular, $\widehat{\omega}(a+) \ge 0$.

Now, define $\omega(t) := \widehat{\omega}(t+)$ for $t \in [a,b)$ and $\omega(b) = \widehat{\omega}(b)$. Then ω is right-continuous and non-decreasing. By Theorem A.5, there exists a Borel measure μ on $[a,b]$ with $\mu((s,t]) = \omega(t) - \omega(s)$ for $s < t$ and $\mu(\{a\}) = \widehat{\omega}(a+) = \omega(a)$. Thus,

$$\mu([a,b]) = \omega(b) - \omega(a) + \omega(a) = \widehat{\omega}(b) = L(\mathbb{1}) = \ell(\mathbb{1}) = \|\ell\|.$$

If $f \in C([a,b],\mathbb{R})$ and $\varepsilon > 0$, choose a step function $g = \sum_{i=1}^n \alpha_i \chi_{\Delta_i}$, where $\Delta_1 = [a, b_1]$ and $\Delta_i = (a_i, b_i]$, $i = 2, \ldots, n$, such that $\|f - g\|_\infty \le \varepsilon$. It is no restriction to assume that all a_i and b_i (except $b_n = b$) are points of continuity of $\widehat{\omega}$. Then we have

$$\begin{aligned}
\left|\ell(f) - \int f\,\mathrm{d}\mu\right| &\le |L(f) - L(g)| + \underbrace{\left|\sum_i \alpha_i\left(\widehat{\omega}(b_i) - \widehat{\omega}(a_i)\right) - \int\left(\sum_i \alpha_i \chi_{\Delta_i}\right)\mathrm{d}\mu\right|}_{=0} \\
&\quad + \left|\int g\,\mathrm{d}\mu - \int f\,\mathrm{d}\mu\right| \\
&\le 2\|\ell\| \cdot \|f - g\|_\infty \le 2\|\ell\|\varepsilon.
\end{aligned}$$

This shows $\ell f = \int f\,\mathrm{d}\mu$ and thus completes the proof of the theorem for the case $\sigma = [a,b]$.

2. Let now $\sigma \subset \mathbb{R}$ be an arbitrary compact set, and $[m,M]$ be the convex hull of σ. For $f \in C([m,M],\mathbb{R})$ define $Lf := \ell\left(f|_\sigma\right)$. If $f \ge 0$ then $Lf = \ell\left(f|_\sigma\right) \ge 0$. Hence, $L \ge 0$, and thus $L \in C([m,M],\mathbb{R})'$ with $\|L\| = L(\mathbb{1}) = \ell(\mathbb{1}) = \|\ell\|$. By the first part of this proof, there exists a Borel measure ν on $[m,M]$ with $\nu([m,M]) = \|\ell\|$ and $Lf = \int f\,\mathrm{d}\nu$ for $f \in C([m,M],\mathbb{R})$.

Let us show that $[m,M] \setminus \sigma$ is of measure 0. For this, let $[\alpha,\beta] \subset [m,M] \setminus \sigma$. Then there exists $f \in C([m,M],\mathbb{R})$, $f \ge 0$, such that $f|_{[\alpha,\beta]} = 1$ and $f|_\sigma = 0$. Thus,

$$0 = \ell\left(f|_\sigma\right) = Lf = \int f\,\mathrm{d}\nu \ge \int_{[\alpha,\beta]} f\,\mathrm{d}\nu = \nu([\alpha,\beta]).$$

Continuity of measures from below and σ-additivity give $\nu([m, M] \setminus \sigma) = 0$ and thus $\nu(\sigma) = \nu([m, M]) = \|\ell\|$. We define the measure μ on σ by setting $\mu(\Delta) := \nu(\Delta)$ for Borel sets $\Delta \subset \sigma$. Then $\mu(\sigma) = \|\ell\|$, and for $f \in C(\sigma)$ we have $\ell f = L\widetilde{f} = \int \widetilde{f} \, \mathrm{d}\nu = \int_\sigma \widetilde{f} \, \mathrm{d}\nu = \int_\sigma f \, \mathrm{d}\mu$, where $\widetilde{f} : [m, M] \to \mathbb{R}$ is any continuous extension of f (see Lemma 1.8). $\square$

B. Addendum: Collection from the Exercise Classes

In this section of the appendix, we collect a couple of topics from the exercise classes which were held in addition to the lectures.

B.1. Convolution Theorem and Shannon Sampling Theorem

The following proposition is known as the convolution theorem in Fourier analysis. A similar theorem holds for the Laplace transform.

Proposition B.1. *If $f, g \in L^1(\mathbb{R}^n)$ then $f * g \in L^1(\mathbb{R}^n)$ and $\mathcal{F}(f * g) = (2\pi)^{\frac{n}{2}} \mathcal{F}f \cdot \mathcal{F}g$.*

Proof. It was proved in the lecture that $f * g \in L^1(\mathbb{R}^n)$. For $y \in \mathbb{R}^n$ we have

$$\begin{aligned} \mathcal{F}(f * g)(y) &= c \int_{\mathbb{R}^n} (f * g)(x) e^{-ix\cdot y} \,\mathrm{d}x = c \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f(x-z) g(z) \,\mathrm{d}z \right) e^{-ix\cdot y} \,\mathrm{d}x \\ &= c \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x-z) e^{-i(x-z)\cdot y} g(z) e^{-iz\cdot y} \,\mathrm{d}x \,\mathrm{d}z \\ &= c \int_{\mathbb{R}^n} g(z) e^{-iz\cdot y} \left(\int_{\mathbb{R}^n} f(x) e^{-ix\cdot y} \,\mathrm{d}x \right) \mathrm{d}z = (\mathcal{F}g)(y) \cdot \frac{1}{c} (\mathcal{F}f)(y). \end{aligned}$$

The claim is proved. □

Proposition B.2. *Let $f \in L^1(\mathbb{R}^n)$ be such that $\mathcal{F}f \in L^1(\mathbb{R}^n)$. Then we have a.e.*

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}^n} \mathcal{F}f(y) e^{ixy} \,\mathrm{d}y.$$

The function $\operatorname{sinc} : \mathbb{R} \to \mathbb{R}$ is defined by

$$\operatorname{sinc}(x) := \begin{cases} \frac{\sin(\pi x)}{\pi x}, & \text{for } x \in \mathbb{R} \setminus \{0\}, \\ 1, & \text{for } x = 0. \end{cases}$$

Theorem B.3 (Shannon Sampling Theorem). *Let $f \in C(\mathbb{R}) \cap L^1(\mathbb{R})$, and assume that $\operatorname{supp} \mathcal{F}f \subset \left[-\frac{\pi}{T}, \frac{\pi}{T}\right]$ for some $T > 0$. Then the k-th Fourier coefficient of $\mathcal{F}f|_{[-\frac{\pi}{T}, \frac{\pi}{T}]}$ is given by $\sqrt{T} f(-kT)$, and the sequence of functions (f_n), defined by*

$$f_n(x) := \sum_{k=-n}^{n} f(kT) \operatorname{sinc}\left(\frac{x}{T} - k\right),$$

converges uniformly to f (on $\mathbb{R}$). Indeed, we have $\|f_n - f\|_\infty^2 \leq \sum_{|k|>n} |f(kT)|^2$.

Proof. Let $\langle \cdot, \cdot \rangle$ denote the scalar product on $L^2(-a, a)$, where $a := \frac{\pi}{T}$. For $x \in \mathbb{R}$ define $e_x \in L^2(-a, a)$ by $e_x(y) := \frac{1}{\sqrt{2a}} e^{iTxy}$, $y \in (-a, a)$. The sequence $(e_k)_{k\in\mathbb{Z}}$ is an orthonormal basis of $L^2(-a, a)$. Since $\operatorname{supp} \mathcal{F}f \subset [-a, a]$ we have $\mathcal{F}f \in L^1(\mathbb{R})$. Hence, by Proposition B.2 we get

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} \mathcal{F}f(y) e^{ixy} \,\mathrm{d}y = \frac{1}{\sqrt{T}} \left\langle \mathcal{F}f, e_{-\frac{x}{T}} \right\rangle.$$

Also, $\widetilde{f} := \mathcal{F}f|_{[-a,a]} \in L^2(-a,a)$, and hence by applying the last equation with $x = -kT$,

$$\widetilde{f} = \sum_{k\in\mathbb{Z}} \left\langle \widetilde{f}, e_k \right\rangle e_k = \sum_{k\in\mathbb{Z}} \sqrt{T} f(-kT) e_k = \sqrt{T} \sum_{k\in\mathbb{Z}} f(kT) e_{-k}. \tag{B.1}$$

The convergence is in $L^2(-a,a)$. Now,

$$\begin{aligned}\left\langle e_{-k}, e_{-\frac{x}{T}} \right\rangle &= \frac{1}{2a} \int_{-a}^{a} e^{-iTky} e^{ixy}\,\mathrm{d}y = \frac{1}{2a} \int_{-a}^{a} e^{iy(x-kT)}\,\mathrm{d}y = \frac{1}{2a} \left[\frac{e^{iy(x-kT)}}{i(x-kT)} \right]_{-a}^{a} \\ &= \frac{1}{2a} \frac{e^{ia(x-kT)} - e^{-ia(x-kT)}}{i(x-kT)} = \frac{\sin(a(x-kT))}{a(x-kT)} = \operatorname{sinc}\left(\frac{x}{T} - k \right).\end{aligned}$$

Thus,

$$\begin{aligned} f_n(x) &= \sum_{k=-n}^{n} f(kT) \operatorname{sinc}\left(\frac{x}{T} - k \right) = \sum_{k=-n}^{n} f(kT) \left\langle e_{-k}, e_{-\frac{x}{T}} \right\rangle \\ &= \left\langle \sum_{k=-n}^{n} f(kT) e_{-k}, e_{-\frac{x}{T}} \right\rangle = \frac{1}{\sqrt{T}} \left\langle g_n, e_{-\frac{x}{T}} \right\rangle \end{aligned}$$

where $g_n := \sqrt{T} \sum_{k=-n}^{n} f(kT) e_{-k}$. By (B.1), we know that $\|g_n - \widetilde{f}\|_2 \to 0$ as $n \to \infty$. Now, for each $x \in \mathbb{R}$ we have

$$\begin{aligned} |f_n(x) - f(x)| &= \left| \frac{1}{\sqrt{T}} \left\langle g_n, e_{-\frac{x}{T}} \right\rangle - \frac{1}{\sqrt{T}} \left\langle \widetilde{f}, e_{-\frac{x}{T}} \right\rangle \right| = \frac{1}{\sqrt{T}} \left| \left\langle g_n - \widetilde{f}, e_{-\frac{x}{T}} \right\rangle \right| \\ &\le \frac{1}{\sqrt{T}} \|g_n - \widetilde{f}\|_2 \|e_{-\frac{x}{T}}\|_2 = \frac{1}{\sqrt{T}} \|g_n - \widetilde{f}\|_2 \to 0. \end{aligned}$$

Thus, $f_n \to f$ uniformly on $\mathbb{R}$. Finally, Parseval's identity implies

$$\|g_n - \widetilde{f}\|_2^2 = \sum_{k\in\mathbb{Z}} |\langle g_n - \widetilde{f}, e_k \rangle|^2 = \sum_{|k|>n} \left(\sqrt{T} f(-kT) \right)^2 = T \sum_{|k|>n} |f(kT)|^2.$$

The theorem is proved. □

B.2. Uncertainty Principle

For two linear operators $A : H \supset \operatorname{dom} A \to H$ and $B : H \supset \operatorname{dom} B \to H$ one defines the *commutator* of A and B by

$$[A, B] := AB - BA,$$

i.e., $\operatorname{dom}[A,B] := \operatorname{dom}(AB) \cap \operatorname{dom}(BA)$ and $[A,B]x := ABx - BAx$, $x \in \operatorname{dom}[A,B]$.

Lemma B.4. *Let A, B be symmetric operators in H. Then for all $a, b \in \mathbb{R}$ and all $x \in \operatorname{dom}[A,B]$, we have*

$$\|(A-a)x\| \cdot \|(B-b)x\| \ge \frac{1}{2} |\langle [A,B]x, x \rangle| \tag{B.2}$$

Equality holds in (B.2) *if and only if $(A-a)x = ic(B-b)x$ for some $c \in \mathbb{R}$.*

Proof. Let $x \in \operatorname{dom}[A, B]$. Then

$$\begin{aligned}\langle [A,B]x, x\rangle &= \langle ABx - BAx, x\rangle = \langle (A-a)(B-b)x - (B-b)(A-a)x, x\rangle \\ &= \langle (A-a)(B-b)x, x\rangle - \langle (B-b)(A-a)x, x\rangle \\ &= \langle (B-b)x, (A-a)x\rangle - \langle (A-a)x, (B-b)x\rangle \\ &= \langle (B-b)x, (A-a)x\rangle - \overline{\langle (B-b)x, (A-a)x\rangle} \\ &= 2i \operatorname{Im} \langle (B-b)x, (A-a)x\rangle .\end{aligned}$$

Thus,

$$|\langle [A,B]x, x\rangle| = 2\,|\operatorname{Im}\langle (B-b)x, (A-a)x\rangle| \le 2\,|\langle (B-b)x, (A-a)x\rangle| .$$

Application of Cauchy-Schwartz inequality gives (B.2).

As to the equality claim, we have to secure equality in both $\le$ positions in the last inequality. The first $\le$ holds with equality iff $\langle (B-b)x, (A-a)x\rangle \in i\mathbb{R}$. The second $\le$ is equality iff both arguments in the scalar product are linearly dependent, i.e., iff $(A-a)x = \lambda(B-b)x$ with some $\lambda \in \mathbb{C}$. Together, this shows

$$(A-a)x = ic(B-b)x$$

for some $c \in \mathbb{R}$. □

Define the operators $X : L^2(\mathbb{R}) \supset \operatorname{dom} X \to L^2(\mathbb{R})$ and $P : L^2(\mathbb{R}) \supset \operatorname{dom} P \to L^2(\mathbb{R})$ by

$$\operatorname{dom} X := \left\{ f \in L^2(\mathbb{R}) : \mathrm{id} \cdot f \in L^2(\mathbb{R}) \right\}, \quad Xf := \mathrm{id} \cdot f,$$

and

$$\operatorname{dom} P := H^1(\mathbb{R}), \quad Pf := -if'.$$

Let $f, g \in \operatorname{dom} X$. Then

$$\langle Xf, g\rangle = \int_{\mathbb{R}} x f(x)\overline{g(x)}\,\mathrm{d}x = \int_{\mathbb{R}} f(x)\overline{xg(x)}\,\mathrm{d}x = \langle f, Xg\rangle .$$

If $f, g \in \operatorname{dom} P = H^1(\mathbb{R})$, then there exists $(\varphi_n) \subset C_c^\infty(\mathbb{R})$ such that $\|g - \varphi_n\|_{H^1} \to 0$ as $n \to \infty$. Thus,

$$\langle Pf, g\rangle = \left\langle -if', g\right\rangle = \lim_{n\to\infty} \left\langle -if', \varphi_n\right\rangle = \lim_{n\to\infty} \left\langle if, \varphi_n'\right\rangle = \left\langle if, g'\right\rangle = \left\langle f, -ig'\right\rangle = \langle f, Pg\rangle .$$

Hence, both operators X and P are symmetric. Let us calculate the domains of XP and PX. We have

$$\operatorname{dom}(XP) = \left\{ f \in H^1(\mathbb{R}) : f' \in \operatorname{dom} X \right\} = \left\{ f \in H^1(\mathbb{R}) : \mathrm{id} \cdot f' \in L^2(\mathbb{R}) \right\},$$

and

$$\operatorname{dom}(PX) = \left\{ f \in L^2(\mathbb{R}) : \mathrm{id} \cdot f \in H^1(\mathbb{R}) \right\}.$$

Hence,

$$\begin{aligned}\operatorname{dom}[X,P] &= \left\{ f \in H^1(\mathbb{R}) : \mathrm{id} \cdot f' \in L^2(\mathbb{R}),\ \mathrm{id} \cdot f \in H^1(\mathbb{R}) \right\} \\ &= \left\{ f \in H^1(\mathbb{R}) : \mathrm{id} \cdot f' \in L^2(\mathbb{R}),\ \mathrm{id} \cdot f \in L^2(\mathbb{R}) \right\}.\end{aligned}$$

Theorem B.5. *For $f \in L^2(\mathbb{R})$ and $a, b \in \mathbb{R}$ we have*

$$\left(\int_{\mathbb{R}} (x-a)^2 |f(x)|^2 \,\mathrm{d}x\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}} (y-b)^2 |\mathcal{F}f(y)|^2 \,\mathrm{d}y\right)^{\frac{1}{2}} \geq \frac{1}{2}\|f\|_2^2. \tag{B.3}$$

Proof. Let $f \in L^2(\mathbb{R})$. We do not have to prove anything when either $xf \notin L^2(\mathbb{R})$ or $y\mathcal{F}f \notin L^2(\mathbb{R})$ since in these cases, the left hand side of (B.3) is infinity, so (B.3) is trivially satisfied. Hence, it is no restriction to assume that $xf \in L^2(\mathbb{R})$ and $y\mathcal{F}f \in L^2(\mathbb{R})$. By Theorem 3.30, the latter is equivalent to $f \in H^1(\mathbb{R})$.

First, let us assume in addition that also $xf' \in L^2(\mathbb{R})$. Then $f \in \mathrm{dom}[X, P]$, and we have

$$[X,P]f = XPf - PXf = -i\left(\mathrm{id}\, f' - (\mathrm{id}\, f)'\right) = -i\left(\mathrm{id}\, f' - \left(f + \mathrm{id}\, f'\right)\right) = if.$$

Hence, Lemma B.4 implies

$$\|(X-a)f\|_2 \cdot \|(P-b)f\|_2 \geq \frac{1}{2}\,|\langle if, f\rangle| = \frac{1}{2}\|f\|_2^2.$$

Evidently, $\|(X-a)f\|_2$ coincides with the first factor on the left hand side of (B.3). Moreover,

$$\|(P-b)f\|_2 = \|-if' - bf\|_2 = \|\mathcal{F}(-if' - bf)\|_2 = \|(y-b)\mathcal{F}f\|_2,$$

which is the second factor. Hence, we have established (B.3) for $f \in H^1(\mathbb{R})$ such that $xf', xf \in L^2(\mathbb{R})$.

Let us now drop the assumption that $xf' \in L^2(\mathbb{R})$. Let $(\varphi_n) \subset C_c^\infty(\mathbb{R})$ be a sequence of test functions such that for each $n \in \mathbb{N}$ we have

(a) $0 \leq \varphi_n(x) \leq 1$ for all $x \in \mathbb{R}$,

(b) $\varphi_n|[-n,n] = 1$,

(c) $\mathrm{supp}(\varphi_n) \subset [-n-1, n+1]$,

(d) $0 \leq |\varphi_n'(x)| \leq 2$ for all $x \in \mathbb{R}$.

Put $f_n := \varphi_n f$, $n \in \mathbb{N}$. Then $f_n \in \mathrm{dom}[X, P]$, so that (B.3) holds for each f_n:

$$\left(\int_{\mathbb{R}} (x-a)^2 |f_n(x)|^2 \,dx\right)^{1/2} \left(\int_{\mathbb{R}} (y-b)^2 |\mathcal{F}f_n(y)|^2 \,dy\right)^{1/2} \geq \frac{1}{2}\|f_n\|_2^2. \tag{B.4}$$

An easy application of Lebesgue's theorem (or Beppo-Levi) proves that $\|f_n - f\|_2 \to 0$ and $\|(x-a)f_n - (x-a)f\|_2 \to 0$ as $n \to \infty$. Hence, the first factor on the left hand side of (B.4) and the right hand side of (B.4) converge to their counterparts in (B.3). For the remainder, we observe that

$$\begin{aligned} \|y\mathcal{F}f_n - y\mathcal{F}f\|_2 &= \|y\mathcal{F}(f_n - f)\|_2 = \|\mathcal{F}(f_n' - f')\|_2 = \|f_n' - f'\|_2 \\ &= \|\varphi_n' f + \varphi_n f' - f'\|_2 \leq \|\varphi_n' f\|_2 + \|(1-\varphi_n)f'\|_2. \end{aligned}$$

Now, $\|(1-\varphi_n)f'\|_2 \to 0$ as $n \to \infty$ by Lebesgue, and

$$\|\varphi_n' f\|_2^2 = \int\limits_{n\le|x|\le n+1} (\varphi_n')^2 |f|^2 \,dx \;\le\; \|\varphi_n'\|_\infty^2 \int_{|x|\ge n} |f|^2\,dx \;\le\; 4\int_{|x|\ge n} |f|^2\,dx,$$

which also tends to zero as $n \to \infty$ (again by Lebesgue). This finishes the proof of (B.3) for $f \in L^2(\mathbb{R})$. ☐

B.3. Some More Facts about Sobolev Functions

Sobolev's Lemma 3.27 can be strengthened in the case $n = m = 1$:

Theorem B.6. *Let $I \subset \mathbb{R}$ be an open intervall, $f \in H^1(I)$. Then $f \in C(\overline{I})$ and*

$$f(y) - f(x) = \int_x^y f'(t)\,\mathrm{d}t \quad \textit{for a.e. } x, y \in I. \tag{B.5}$$

Proof. **1.** In a first step, we assume that $I = \mathbb{R}$. Then, since $C_c^\infty(\mathbb{R})$ is dense in $H^1(\mathbb{R})$ we find a sequence $(\varphi_n) \subset C_c^\infty(\mathbb{R})$ such that $\|f - \varphi_n\|_2 \to 0$ and $\|f' - \varphi_n'\|_2 \to 0$. For $R > 0$, choose $\phi \in C_c^\infty(\mathbb{R})$ such that $\phi(x) = 1$ on $[-R, R]$ and $\phi(x) = 0$ for $x \in \mathbb{R}\setminus(-R-1, R+1)$.

Put $F := \phi f$ and $F_n := \phi\varphi_n$. Then $F \in H^1(\mathbb{R})$ and $F_n \in C_c^\infty(\mathbb{R})$ and we have $F' = \phi' f + \phi f' \in L^2(\mathbb{R})$. Moreover,

$$\begin{aligned}\|F_n' - F'\|_2 &= \|\phi'\varphi_n + \phi\varphi_n' - \phi' f - \phi f'\|_2 \le \|\phi'(\varphi_n - f)\|_2 + \|\phi(\varphi_n' - f')\|_2 \\ &\le \|\phi'\|_\infty \|\varphi_n - f\|_2 + \|\phi\|_\infty \|\varphi_n' - f'\|_2 \to 0.\end{aligned}$$

Define

$$\widetilde{F}(x) := \int_{-R-1}^x F'(x)\,\mathrm{d}t.$$

Then

$$\begin{aligned}|F_n(x) - \widetilde{F}(x)| &= \left|\int_{-R-1}^x F_n'(t)\,\mathrm{d}t - \int_{-R-1}^x F'(t)\,\mathrm{d}t\right| = \left|\int_{-R-1}^x (F_n' - F')\,\mathrm{d}t\right| \\ &\le \int_{-R-1}^x |F_n' - F'|\,\mathrm{d}t \le \|F_n' - F'\|_2 \cdot |x + R + 1|^{\frac{1}{2}} \to 0.\end{aligned}$$

This implies that $F_n \to \widetilde{F}$ pointwise. Furthermore,

$$\|F_n - F\|_2 = \|\phi(\varphi_n - f)\|_2 \le \|\phi\|_\infty \cdot \|\varphi_n - f\|_2 \to 0.$$

Hence, by the Riesz-Fischer Theorem, there exists a subsequence (F_{n_k}) converging a.e. (pointwise) to F. In particular, for a.e. $x \in [-R, R]$ we have $\widetilde{F}(x) = F(x)$, and thus

$$f(x) = F(x) = \widetilde{F}(x) = \int_{-R-1}^{-R} F'\,\mathrm{d}t + \int_{-R}^x (\phi' f + \phi f')\,\mathrm{d}t = \int_{-R-1}^{-R} F'\,\mathrm{d}t + \int_{-R}^x f'\,\mathrm{d}t.$$

Hence, for a.e. $x, y \in [-R, R]$ we have $f(y) - f(x) = \int_x^y f'(t)\,\mathrm{d}t$, and (B.5) follows (for a.e. $x, y \in \mathbb{R}$).

2. Let I be an arbitrary open interval. Let $[a,b] \subset I$, $\psi \in C_c^\infty(I)$ such that $\psi = 1$ on $[a,b]$, and put

$$F := \begin{cases} \psi f, & \text{on } I, \\ 0, & \text{on } \mathbb{R} \setminus I, \end{cases} \qquad \text{and} \qquad H := \begin{cases} \psi' f + \psi f', & \text{on } I, \\ 0, & \text{on } \mathbb{R} \setminus I. \end{cases}$$

For $\varphi \in C_c^\infty(\mathbb{R})$, we have

$$\begin{aligned} \int_{\mathbb{R}} H\varphi \,\mathrm{d}x &= \int_I \psi' f\varphi \,\mathrm{d}x + \int_I \psi f'\varphi \,\mathrm{d}x = \int_I \psi' f\varphi \,\mathrm{d}x - \int_I f(\psi\varphi)' \,\mathrm{d}x \\ &= \int_I [\psi' f\varphi - f\varphi'\psi - f\varphi\psi'] \,\mathrm{d}x = -\int_I f\varphi'\psi \,\mathrm{d}x = -\int_I F\varphi' \,\mathrm{d}x = -\int_R F\varphi' \,\mathrm{d}x. \end{aligned}$$

Thus, $H = F'$ in the weak sense and $H \in L^2(\mathbb{R})$, i.e., $F \in H^1(\mathbb{R})$. By what we have shown above, there exits a function $\widetilde{F}_{a,b} \in C(\mathbb{R})$ such that $F = \widetilde{F}_{a,b}$ a.e. on $\mathbb{R}$ and $\widetilde{F}_{a,b}(y) - \widetilde{F}_{a,b}(x) = \int_x^y F'\,\mathrm{d}t$. Hence, $f = \widetilde{F}_{a,b}$ on $[a,b]$ a.e., and for a.e. $x, y \in [a,b]$ we have

$$f(x) - f(y) = \widetilde{F}_{a,b}(y) - \widetilde{F}_{a,b}(x) = \int_x^y F' \,\mathrm{d}t = \int_x^y H \,\mathrm{d}t = \int_x^y f'(t) \,\mathrm{d}t.$$

This establishes (B.5) and $f = \widetilde{f} \in C(I)$ a.e., where $\widetilde{f}(x) = f(y_0) + \int_x^{y_0} f'\,\mathrm{d}t$ for some $y_0 \in I$. Let x_0 be an endpoint of I, and let $(x_n) \subset I$ with $x_n \to x_0$. Then, as for $x, y \in I$ we have

$$|\widetilde{f}(y) - \widetilde{f}(x)| \le \int_x^y |f'| \,\mathrm{d}t \le \left(\int_x^y |f'|^2 \,\mathrm{d}t \right)^{1/2} |x-y|^{1/2} \le \|f'\|_2 \cdot |x-y|^{1/2},$$

$(\widetilde{f}(x_n))$ is a Cauchy sequence, and its limit is not dependent on the choice of the sequence (x_n). Hence, $\widetilde{f}$ admits an extension in $C(\overline{I})$. □

B.4. Metrizable and Normable Locally Convex Spaces

We would like to consider the question, whether there are necessary and/or sufficient conditions for a locally convex space to be metrizable or even normable. Clearly, by metrizable (normable), we mean that the locally convex topology stems from a metric (norm, respectively).

Proposition B.7. *Let $(X, \mathcal{O}_P)$ be a LCS generated by a countable family of seminorms P. Then X is metrizable. More precisely, if $P = \{p_n : n \in \mathbb{N}\}$, then*

$$d(x,y) := \sum_{n=1}^\infty 2^{-n} \frac{p_n(x-y)}{1 + p_n(x-y)}, \qquad x, y \in X,$$

is a metric which defines $\mathcal{O}_P$.

Proof. Define the function $f : [0, \infty) \to [0, 1)$ by $f(t) := \frac{t}{1+t}$, $t \in [0, \infty)$. Then f is strictly increasing, we have $f(s+t) \le f(s)+f(t)$ for $s, t \in [0, \infty)$, $f(at) \le \max\{1, a\}f(t)$ for $a > 0$, $t \in [0, \infty)$, and $f(t) \le \min\{1, t\}$ for all $t \in [0, \infty)$. We have

$$d(x, y) = \sum_{n=1}^{\infty} 2^{-n} f(p_n(x - y)).$$

Let us prove that d is a metric. Obviously, d is symmetric. Further,

$$\begin{aligned} d(x, z) &= \sum_{n=1}^{\infty} 2^{-n} f(p_n(x - z)) \le \sum_{n=1}^{\infty} 2^{-n} f\left(p_n(x - y) + p_n(y - z)\right) \\ &\le \sum_{n=1}^{\infty} 2^{-n} \left[f(p_n(x - y)) + f(p_n(y - z))\right] = d(x, y) + d(y, z). \end{aligned}$$

This proves the triangle inequality. Now, let $x, y \in X$ be such that $d(x, y) = 0$. Then $f(p_n(x - y)) = 0$ for all $n \in \mathbb{N}$ and thus $x = y$, since P is separating. We will now prove the following claims:

(a) $(X, \mathcal{O}_d)$ is a TVS.

(b) $\mathcal{O}_d \subset \mathcal{O}_P$.

(c) $\mathcal{O}_P \subset \mathcal{O}_d$.

(a). We have that $+ : X \times X \to X$ is d-continuous:

$$\begin{aligned} d(u + v, x + y) &= \sum_{n=1}^{\infty} 2^{-n} f(p_n(u + v - x - y)) \\ &\le \sum_{n=1}^{\infty} 2^{-n} \left[f(p_n(u - x)) + f(p_n(v - y))\right] = d(u, x) + d(v, y). \end{aligned}$$

Furthermore,

$$\begin{aligned} d(\lambda x, \lambda_0 x_0) &= \sum_{n=1}^{\infty} 2^{-n} f(p_n(\lambda x - \lambda_0 x_0)) \le \sum_{n=1}^{\infty} 2^{-n} f(p_n(\lambda(x - x_0)) + p_n((\lambda - \lambda_0)x_0)) \\ &\le \sum_{n=1}^{\infty} 2^{-n} \big[f(|\lambda| p_n(x - x_0)) + f(|\lambda - \lambda_0| p_n(x_0))\big] \\ &\le \max\{1, |\lambda|\} \sum_{n=1}^{\infty} 2^{-n} f(p_n(x - x_0)) + \sum_{n=1}^{\infty} 2^{-n} |\lambda - \lambda_0| p_n(x_0) \\ &= \max\{1, |\lambda|\} d(x, x_0) + |\lambda - \lambda_0| d(x_0, 0), \end{aligned}$$

which shows that also $\cdot : \mathbb{K} \times X \to X$ is d-continuous.

(b). Let $\varepsilon > 0$. We prove, that there exists $U \in \mathcal{O}_P$ with $0 \in U$ such that $U \subset B_\varepsilon(0)$. Let $N \in \mathbb{N}$ such that $\sum_{n=N+1}^\infty 2^{-n} < \frac{\varepsilon}{2}$, and put

$$U := \bigcap_{n=1}^N V\left(p_n, \frac{\varepsilon}{2N-\varepsilon}\right) \in \mathcal{O}_P.$$

We have $0 \in U$, and for $x \in U$,

$$\begin{aligned} d(x,0) &= \sum_{n=1}^N 2^{-n} f(p_n(x)) + \sum_{n=N+1}^\infty 2^{-n} f(p_n(x)) < \sum_{n=1}^N f(p_n(x)) + \frac{\varepsilon}{2} \\ &< \sum_{n=1}^N f\left(\frac{\varepsilon}{2N-\varepsilon}\right) + \frac{\varepsilon}{2} = N\frac{\frac{\varepsilon}{2N-\varepsilon}}{1+\frac{\varepsilon}{2N-\varepsilon}} + \frac{\varepsilon}{2} = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Thus, $x \in B_\varepsilon(0)$, and hence $U \subset B_\varepsilon(0)$.

(c). Let $\varepsilon > 0$ and $n \in \mathbb{N}$. Then with $\delta := 2^{-n} f(\varepsilon)$, we have $B_\delta(0) \subset V(p_n, \varepsilon)$. Indeed, if $x \in B_\delta(0)$, i.e.,

$$\sum_{k=1}^\infty 2^{-k} f(p_k(x)) < 2^{-n} f(\varepsilon),$$

then $2^{-n} f(p_n(x)) < 2^{-n} f(\varepsilon)$, i.e., $p_n(x) < \varepsilon$. □

Theorem B.8. *A LCS is metrizable if and only if its topology is generated by a countable family of seminorms.*

Proof. Let $(X, \mathcal{O}_P)$ be a LCS. If P is countable, then X is metrizable by Proposition B.7. Let X be metrizable with metric D. For all $n \in \mathbb{N}$ we have $B_{1/n}(0) \in \mathcal{O}_D = \mathcal{O}_P$. Hence, for all $n \in \mathbb{N}$ there exists a finite subset $F_n \subset P$ and $\varepsilon_n > 0$ such that

$$\bigcap_{p \in F_n} V(p, \varepsilon_n) \subset B_{\frac{1}{n}}(0).$$

Put $F := \bigcup_{n=1}^\infty F_n \subset P$. Then F is countable and $\mathcal{O}_D \subset \mathcal{O}_F \subset \mathcal{O}_P = \mathcal{O}_D$, implying $\mathcal{O}_F = \mathcal{O}_P$. □

Definition B.9. Let $(X, \mathcal{O})$ be a TVS. A set $B \subset X$ is called **bounded** if for all $U \in \mathcal{O}$ with $0 \in U$ there exists $t > 0$ such that $B \subset tU$.

Theorem B.10. *A LCS is normalizable if and only if it contains a bounded neighborhood of zero.*

Proof. Let $(X, \mathcal{O}_P)$ be a LCS. If X is normalizable, the claim is clear. Let $U \in \mathcal{O}_P$, $0 \in U$, be bounded. There exist a finite set $F \subset P$ and $\varepsilon > 0$ such that $\bigcap_{p \in F} V(p, \varepsilon) \subset U$. Put $q(x) := \varepsilon^{-1} \max_{p \in F} p(x)$, $x \in X$. Then q is a continuous seminorm on X and satisfies $V(q, 1) \subset U$. Let us show that q is a norm.

Let $x \in X$ with $x \neq 0$. Then there exists $W \in \mathcal{O}_P$ with $0 \in W$ and $x \notin W$. Since U is bounded, we find some $t > 0$ such that $V(q,1) \subset U \subset tW$. As $tx \notin tW$, also $tx \notin V(q,1)$, and thus $q(x) = t^{-1}q(tx) \geq t^{-1} > 0$. Thus, q is indeed a norm.

Now, let $p \in P$, $\varepsilon > 0$. Then there exists $t > 0$ such that $V(q,1) \subset U \subset tV(p,\varepsilon)$. Hence,

$$V(q,t^{-1}) = t^{-1}V(q,1) \subset V(p,\varepsilon).$$

This shows $\mathcal{O}_P \subset \mathcal{O}_q$. And as $\mathcal{O}_q \subset \mathcal{O}_P$ is evident, the theorem is proved. □

B.5. The Stone-Weierstraß Theorem

We begin this section with a couple of measure-theoretical facts. Let T be a compact topological space. As usual, we denote the Borel sigma algebra on T by Σ. Let $\mu : \Sigma \to \mathbb{C}$ be a complex-valued measure (i.e., $\mu(\varnothing) = 0$, and for mutually disjoint sets $A_n \in \Sigma$, we have $\mu\left(\bigcup_n A_n\right) = \sum_n \mu(A_n)$). We define the *variation* $|\mu|$ of μ by

$$|\mu|(A) := \sup\left\{\sum_{i=1}^n |\mu(E_i)| : A = \bigcup_{i=1}^n E_i,\, E_i \in \Sigma,\, E_i \cap E_j = \varnothing \text{ if } i \neq j\right\}, \quad A \in \Sigma.$$

It can be shown that $|\mu|$ is a finite positive measure on T. The complex-valued measure μ is called *regular* if for all $A \in \Sigma$ we have

$$|\mu|(A) = \inf\{|\mu|(U) : A \subset U,\, U \text{ open}\} = \sup\{|\mu|(C) : C \subset A \text{ compact}\}.$$

By $M(T)$ we denote the set of all regular complex valued Borel measures on T. Obviously, $M(T)$ is a linear space. For $\mu \in M(T)$ we define $\|\mu\| := |\mu|(T)$. It can be shown that $\|\cdot\|$ is a norm on $M(T)$ and that $(M(T), \|\cdot\|)$ is a Banach space. The norm $\|\cdot\|$ is called the *variation norm* of μ.

Each $\mu \in M(T)$ admits a decomposition $\mu = \mu_+ - \mu_- + i(\nu_+ - \nu_-)$ with positive measures $\mu_\pm$ and $\nu_\pm$ (this is a non-trivial fact!). For $f \in L^1(\mu) := L^1(\mu_+ + \mu_- + \nu_+ + \nu_-)$, one puts

$$\int f\,\mathrm{d}\mu := \int f\,\mathrm{d}\mu_+ - \int f\,\mathrm{d}\mu_- + i\left(\int f\,\mathrm{d}\nu_+ - \int f\,\mathrm{d}\nu_-\right).$$

For $\mu \in M(T)$ and $f \in L^1(\mu)$, we define a complex-valued measure μ_f by $\mu_f(A) := \int_A f\,\mathrm{d}\mu$, $A \in \Sigma$. One can show that

$$\|\mu_f\| = \int |f|\,\mathrm{d}|\mu|.$$

The following version of the Riesz Representation Theorem is probably the most general one.

Theorem B.11 (Riesz Representation Theorem). *$C(T)' \cong M(T)$ via the mapping $\phi : M(T) \to C(T)'$, defined by*

$$(\phi\mu)f := \int f\,\mathrm{d}\mu, \quad \mu \in M(T),\, f \in C(T).$$

Here, we finish our preliminaries from measure theory. For proofs and details, the reader is refered to [8].

It is well known that $C(T) := \{f : T \to \mathbb{K} \,|\, f \text{ continuous}\}$ is a Banach space with the norm $\|\cdot\|_\infty$. An *algebra* in $C(T)$ is a (not necessarily closed) subspace $\mathcal{A} \subset C(T)$ such that $fg \in \mathcal{A}$ for all $f, g \in \mathcal{A}$. An algebra $\mathcal{A}$ in $C(T)$ is called *unital* if $\mathbb{1} \in \mathcal{A}$. The following Stone-Weierstraß Theorem yields a condition on a unital algebra to be dense in $C(T)$.

Theorem B.12 (Stone-Weierstraß). *A unital algebra $\mathcal{A} \subset C(T)$ is dense in $C(T)$ if it satisfies the following conditions:*

(i) *$\mathcal{A}$ separates points (i.e., $\forall s, t \in T$, $s \neq t$, $\exists f \in \mathcal{A} : f(s) \neq f(t)$).*

(ii) *$\mathcal{A}$ is selfadjoint, i.e., $f \in \mathcal{A} \implies \overline{f} \in \mathcal{A}$.*

Proof. Recall that the annihilator of $\mathcal{A}$ is defined by

$$\mathcal{A}^\perp = \big\{F \in C(T)' : F(f) = 0 \;\forall f \in \mathcal{A}\big\} = \left\{\mu \in M(T) : \int f \,\mathrm{d}\mu = 0 \;\forall f \in \mathcal{A}\right\}.$$

By a theorem from Functional Analysis I (see [5]), it suffices to show that $\mathcal{A}^\perp = \{0\}$.

Suppose $\mathcal{A}^\perp \neq \{0\}$. The weak* topology on $C(T)'$ is generated by the seminorms $p_f : C(T)' \to [0, \infty)$, $f \in C(T)$, with $p_f(F) := |Ff|$, $F \in C(T)'$. Hence, $A^\perp = \bigcap_{f \in A} p_f^{-1}(\{0\})$ is weak*-closed[10]. Hence, $K := \{\mu \in \mathcal{A}^\perp : \|\mu\| \leq 1\}$ is weak*-compact by Alaoglu's Theorem. By the Krein-Milman Theorem, $\mathrm{ex}(K) \neq \varnothing$. Let $\mu \in \mathrm{ex}(K)$. By assumption, $\mathcal{A}^\perp \neq \{0\}$. Hence, $K \neq \{0\}$. Choose $\nu \in K \setminus \{0\}$. Then $0 = \frac{1}{2}\nu + \frac{1}{2}(-\nu)$ implies that $\mu \neq 0$. Define

$$S := T \setminus \bigcup \{U \subset T \text{ open} : |\mu|(U) = 0\} = \bigcap \{A \subset T \text{ closed} : |\mu|(A^c) = 0\},$$

the support of μ. Then $|\mu|(S^c) = 0$ (impliying $S \neq \varnothing$), and thus for all $f \in C(T)$ we have $\int f \,\mathrm{d}\mu = \int_S f \,\mathrm{d}\mu$, since

$$\left|\int_{S^c} f \,\mathrm{d}\mu\right| \leq \int_{S^c} |f| \,\mathrm{d}|\mu| = 0.$$

We have $\|\mu\| = 1$, since otherwise $\mu = \|\mu\| \frac{\mu}{\|\mu\|} + (1 - \|\mu\|)0$ is in the relative interior of a segment in K.

As $S \neq \varnothing$, we can fix $t_0 \in S$. Let us show that $S = \{t_0\}$. Suppose there is $t \in S$ with $t \neq t_0$. By (i), there exists $f_1 \in \mathcal{A}$ such that $f_1(t_0) \neq f_1(t) =: \beta$. Moreover, $\beta\mathbb{1} \in \mathcal{A}$ as $\mathcal{A}$ is unital. Thus, $f_2 := f_1 - \beta \in \mathcal{A}$ and $f_2(t_0) \neq f_2(t) = 0$. By (ii), we have $f_3 := |f_2|^2 = f_2\overline{f_2} \in \mathcal{A}$, $f_3 \geq 0$, and $f_3(t_0) > 0 = f_3(t)$. Set

$$f := \frac{f_3}{1 + \|f_3\|_\infty}.$$

[10] Alternatively, $\mathcal{A}^\perp = \mathcal{A}^\circ$ w.r.t. the dual pair $(C(T), C(T)')$ and is thus weak*-closed by Lemma 5.40.

Then $f \in \mathcal{A}$, $0 \le f < 1$, and $f(t_0) > 0 = f(t)$. As $\mu \in \mathcal{A}^\perp$,

$$0 = \int gf \, \mathrm{d}\mu = \int g \, \mathrm{d}\mu_f \quad \forall g \in \mathcal{A}.$$

This implies that $\mu_f \in \mathcal{A}^\perp$. Similarly $\mu_{1-f} \in \mathcal{A}^\perp$. We have $\|\mu_f\| = \int f \, \mathrm{d}|\mu| =: \alpha$.

Since $f(t_0) > 0$ there exist $\varepsilon > 0$ and an open set $U \subset T$ with $t_0 \in U$ such that $f(s) > \varepsilon$ for all $s \in U$. Thus,

$$\alpha = \int f \, \mathrm{d}|\mu| \ge \int_U f \, \mathrm{d}|\mu| \ge \varepsilon|\mu|(U) > 0.$$

(If $|\mu|(U) = 0$, then $U \cap S = \varnothing$, but at the same time we have $t_0 \in U \cap S$.) Furthermore,

$$\alpha = \int f \, \mathrm{d}|\mu| \le \int \mathrm{d}|\mu| = \|\mu\| = 1.$$

If $\alpha = 1$ we would have $f = 1$ $|\mu|$-a.e., which is a contradiction to $f < 1$. Thus $0 < \alpha < 1$. Also,

$$1 - \alpha = \int (1 - f) \, \mathrm{d}|\mu| = \|\mu_{1-f}\|.$$

For all $A \in \Sigma$, we have

$$\mu(A) = \mu_f(A) + \mu_{1-f}(A) = \Bigg(\alpha \underbrace{\frac{\mu_f}{\|\mu_f\|}}_{\in K} + (1 - \alpha) \underbrace{\frac{\mu_{1-f}}{\|\mu_{1-f}\|}}_{\in K}\Bigg)(A).$$

Since $\mu \in \mathrm{ex}(K)$, it follows that $\mu = \alpha^{-1}\mu_f$. Hence $\alpha\mu = \mu_f$. Thus,

$$\int_A (f - \alpha) \, \mathrm{d}\mu = \mu_f(A) - \alpha\mu(A) = 0$$

for all $A \in \Sigma$. Therefore, $\mu_{f-\alpha} = 0$ and hence $0 = \|\mu_{f-\alpha}\| = \int_S |f - \alpha| \, \mathrm{d}|\mu|$, i.e. $f = \alpha$ $|\mu|$-a.e. on S. As $U := \{s \in T : f(s) \neq \alpha\}$ is open and $|\mu|(U) = 0$, we have $S \cap U = \varnothing$ and thus $f \equiv \alpha$ on S. Since $f(t) = 0$, we obtain $t \notin S$, contradicting the initial assumption that $t \in S$. Hence, $S = \{t_0\}$. This shows, that for all $A \in \Sigma$ we have $\mu(A) = \mu(A \cap \{t_0\})$ and thus (note that $\mathbb{1} \in \mathcal{A}$ and $\mu \in \mathcal{A}^\perp$)

$$0 = \int \mathbb{1} \, \mathrm{d}\mu = \mu(T) = \mu(\{t_0\}).$$

Therefore, $\mu = 0$, which contradicts $\|\mu\| = 1$. This shows $\mathcal{A}^\perp = \{0\}$, which completes the proof. □

B.6. Bishop's Theorem

Definition B.13. Let T be a compact Hausdorff space and $\mathcal{A} \subset C(T)$. A subset $E \subset T$ is called $\mathcal{A}$-antisymmetric if for each $f \in \mathcal{A}$ for which $f|_E$ is real-valued we have $f|_E = \text{const}$.

Example B.14. Let $S \subset \mathbb{C}$ be compact and $\mathcal{A} = \{f \in C(S) : f|_{\operatorname{int} S} \text{ is holomorphic}\}$. Then every connected component of $\operatorname{int} S$ is $\mathcal{A}$-antisymmetric.

Theorem B.15 (Bishop). *Let T be a compact Hausdorff space and $\mathcal{A} \subset C(T)$ a closed unital algebra. Then, whenever $g \in C(T)$ is such that for each $\mathcal{A}$-antisymmetric set $E \subset T$ there exists $f \in \mathcal{A}$ satisfying $f|_E = g|_E$, it follows that $g \in \mathcal{A}$.*

Remark B.16. Let us derive the Stone-Weierstraß Theorem from Bishop's Theorem. For this, let $\mathcal{A} \subset C(T)$ be a unital algebra which is separating points and selfadjoint. Then $\{f \in \mathcal{A} : f \text{ real-valued}\}$ is separating points. Indeed, if $s, t \in T$ with $s \neq t$, then there exists $f \in \mathcal{A}$ such that $f(s) \neq f(t)$ and thus $\operatorname{Re} f(s) \neq \operatorname{Re} f(t)$ or $\operatorname{Im} f(s) \neq \operatorname{Im} f(t)$. We have $\operatorname{Re} f = \frac{1}{2}(f + \overline{f}) \in \mathcal{A}$ and $\operatorname{Im} f = \frac{1}{2}(f - \overline{f}) \in \mathcal{A}$.

Note that also $\overline{\mathcal{A}}$ is a unital algebra. Now, let $E \subset T$ be $\overline{\mathcal{A}}$-antisymmetric. Suppose there exist $s, t \in E$ with $s \neq t$. Then there is a real-valued $f \in \mathcal{A} \subset \overline{\mathcal{A}}$ such that $f(s) \neq f(t)$. This contradicts the assumption that E is $\overline{\mathcal{A}}$-antisymmetric. Thus, the $\overline{\mathcal{A}}$-antisymmetric sets are the singleton sets.

Now, let $g \in C(T)$. Then, for each $t \in T$ we have that $f|_{\{t\}} = g|_{\{t\}}$, where $f := g(t)\mathbb{1} \in \mathcal{A} \subset \overline{\mathcal{A}}$. By Bishop's Theorem, we conclude $g \in \overline{\mathcal{A}}$. Thus, $\overline{\mathcal{A}} = C(T)$.

Proof of Bishop's Theorem. Again, we make use of the fact that $C(T)' \cong M(T)$ via $\phi : M(T) \to C(T)'$, $\phi(\mu)f = \int f \,\mathrm{d}\mu$. Again,

$$\mathcal{A}^\perp = \{F \in C(T)' : F(f) = 0 \;\forall f \in \mathcal{A}\} \cong \{\mu \in M(T) : \int f \,\mathrm{d}\mu = 0 \;\forall f \in \mathcal{A}\}.$$

We define $K := \{\mu \in \mathcal{A}^\perp : \|\mu\| \leq 1\}$ (which is a convex and weak*-compact set (see the proof of Theorem B.12)). If $K = \{0\}$ then $\mathcal{A}^\perp = \{0\}$, and thus $\mathcal{A} = C(T)$. Therefore, assume $K \neq \{0\}$. By the Krein-Milman Theorem 5.25, $\operatorname{ex}(K) \neq \varnothing$. Let $\mu \in \operatorname{ex}(K)$ and $S := \operatorname{supp}(\mu)$, the smallest compact set, for which $|\mu|(S) = |\mu|(T) = \|\mu\|$.

Let us show that S is $\mathcal{A}$-antisymmetric. For this, let $f \in \mathcal{A}$ such that $f|_S$ is real-valued. Then, for $g \in \mathcal{A}$, defined by

$$g(t) := \frac{f(t) - \inf_{s \in S} f(s)}{\|f - \inf_{s \in S} f(s)\|_\infty}, \quad t \in T,$$

we have $0 \leq g|_S \leq 1$. Thus, we may assume w.l.o.g. that $0 \leq f|_S \leq 1$. Then we can proceed as in the proof of Theorem B.12 and obtain $f|_S = \|\mu_f\|$. Hence, $S = \operatorname{supp}(\mu)$ is indeed $\mathcal{A}$-antisymmetric for each $\mu \in \operatorname{ex}(K)$.

To prove Bishop's Theorem, let $g \in C(T)$ be such that for all $\mathcal{A}$-antisymmetric sets $E \subset T$ there exists $f \in \mathcal{A}$ such that $f|_E = g|_E$. Then for each $\mu \in \operatorname{ex}(K)$ there exists $f \in \mathcal{A}$ such that $g|_{\operatorname{supp}(\mu)} = f|_{\operatorname{supp}(\mu)}$, and thus

$$\int g \,\mathrm{d}\mu = \int f \,\mathrm{d}\mu = 0,$$

since $\mu \in \mathcal{A}^\perp$. Thus, $\int g \, d\mu = 0$ for all $\mu \in \operatorname{co}(\operatorname{ex}(K)) = K$. Therefore, even $\int g \, d\mu = 0$ for all $\mu \in \mathcal{A}^\perp$.

Suppose, $g \notin \mathcal{A} = \overline{\mathcal{A}}$. By Hahn-Banach, there exists $F \in C(T)'$ such that $F(g) = 1$ and $F|_{\mathcal{A}} = 0$. Then there exists $\nu \in M(T)$ such that $\int g \, d\nu = 1$ and $\int f \, d\nu = 0$ for all $f \in \mathcal{A}$. Thus, $\nu \in \mathcal{A}^\perp$, and hence $\int g \, d\nu = 0$. A contradiction. □

B.7. Additional Remarks to the Introduction of Distributions

Lemma B.17. *The family* $P_K := \{p_m|_{\mathcal{D}_K(\Omega)} : m \in \mathbb{N}_0\}$ *is separating.*

Proof. Each p_m is a norm, and a norm on a linear space is separating. □

Lemma B.18. *For* $C \in \operatorname{Comp}(\Omega)$*, define* $p(\varphi) := \sup_{|\alpha| \le m} \sup_{x \in C} |(\partial^\alpha \varphi)(x)|$, $\varphi \in \mathcal{D}(\Omega)$*. Then* p *is a seminorm on* $\mathcal{D}(\Omega)$ *and* $p \in P$*.*

Proof. It is clear that p is a seminorm. Let $K \subset \Omega$ be compact. Then for $\varphi \in \mathcal{D}_K(\Omega)$ we have

$$p(\varphi) = \sup_{|\alpha| \le m} \sup_{x \in C \cap K} |(\partial^\alpha \varphi)(x)| \le \sup_{|\alpha| \le m} \sup_{x \in K} |(\partial^\alpha \varphi)(x)| = p_m(\varphi).$$

Hence, $p \in P$. □

Remark B.19. Clearly, $P' = \{p_m : m \in \mathbb{N}_0\}$ is a separating family of seminorms on $\mathcal{D}(\Omega)$. Since $p_m|_{\mathcal{D}_K(\Omega)}$ is $\mathcal{O}_{P_K}$-continuous for each $K \in \operatorname{Comp}(\Omega)$, we have $p_m \in P$ for each $m \in \mathbb{N}_0$, and thus $P' \subset P$. Therefore, $\mathcal{O}_{P'} \subset \mathcal{O}_P$.

It is now obvious to ask whether $\mathcal{O}_P = \mathcal{O}_{P'}$ holds. This can be negated as the following counter example shows: Let $\Omega = \mathbb{R}$, $K_j := [-j, j]$, and $x_j := j - \frac{1}{2} \in K_j$ for all $j \in \mathbb{N}$. Define $q_j(\varphi) := |\varphi(x_j)|$, $\varphi \in \mathcal{D}(\mathbb{R})$. From the lecture we know that $q_j \in P$ for each $j \in \mathbb{N}$. Put $p := \sum_{j \in \mathbb{N}} q_j$. This is a well defined seminorm on $\mathcal{D}(\mathbb{R})$. If $K \subset \mathbb{R}$ is compact, there exists $N \in \mathbb{N}$ such that $K \subset K_N$ and $K \not\subseteq K_{N-1}$ (where $K_0 := \varnothing$), and thus

$$p(\varphi) = \sum_{j=1}^{N-1} |\varphi(x_j)| \quad \text{for } \varphi \in \mathcal{D}_K(\Omega).$$

Hence, $p|_{\mathcal{D}_K(\mathbb{R})} = \sum_{j=1}^{N-1} q_j|_{\mathcal{D}_K(\mathbb{R})}$ is $\mathcal{O}_{P_K}$-continuous, and thus $p \in P$.

For given $m \in \mathbb{N}_0$ and $\varepsilon > 0$ choose $\varphi \in \mathcal{D}(\mathbb{R})$ such that[11] $\varphi \equiv \frac{1}{N}$ on K_N and $p_m(\varphi) < \varepsilon$. Then $p(\varphi) \ge \sum_{j=1}^{N} |\varphi(x_j)| = N \cdot \frac{1}{N} = 1$. Thus

$$\{\varphi \in \mathcal{D}(\mathbb{R}) : p_m(\varphi) < \varepsilon\} \not\subseteq \{\varphi \in \mathcal{D}(\mathbb{R}) : p(\varphi) < 1\}$$

for each $m \in \mathbb{N}_0$ and $\varepsilon > 0$. Thus, $\{\varphi \in \mathcal{D}(\mathbb{R}) : p(\varphi) < 1\} \in \mathcal{O}_P \setminus \mathcal{O}_{P'}$.

[11]To construct such a φ, choose any $\varphi_0 \in \mathcal{D}(\mathbb{R})$ with $\varphi_0|_{K_N} = 1$ and set $\varphi(x) := N^{-1}\varphi_0(p_m(\varphi_0)^{-1}x)$.

B.8. Tempered Distributions

We will define derivatives of distributions $T \in \mathcal{D}(\Omega)'$ in the following way:

$$(\partial^\alpha T)\varphi := (-1)^{|\alpha|} T(\partial^\alpha \varphi).$$

For regular distributions T_f with $f \in C^\infty(\Omega)$, we have

$$(\partial^\alpha T_f)\,\varphi = (-1)^{|\alpha|} \int (\partial^\alpha \varphi) f \,\mathrm{d}x = \int \varphi(\partial^\alpha f)\,\mathrm{d}x = T_{\partial^\alpha f}\varphi,$$

i.e., $\partial^\alpha T_f = T_{\partial^\alpha f}$.

The aim is to define a Fourier transform $\mathcal{F}T$. An analogous approach to the definition of derivatives is setting $(\mathcal{F}T)\varphi := T(\mathcal{F}\varphi)$. But $\mathcal{F}\varphi \notin \mathcal{D}(\mathbb{R}^n)$ for all $\varphi \in \mathcal{D}(\mathbb{R}^n) \setminus \{0\}$! The solution is to consider $\varphi \in \mathcal{S}(\mathbb{R}^n) \supset \mathcal{D}(\mathbb{R}^n)$. Thus, we have to define a topology on $\mathcal{S}(\mathbb{R}^n)$. For this, one defines the seminorms

$$p_{\alpha,m}(\varphi) := \sup_{x \in \mathbb{R}^n} (1 + |x|^m)\,|(\partial^\alpha \varphi)(x)|\,, \quad \varphi \in \mathcal{S}(\mathbb{R}^n),$$

where $m \in \mathbb{N}_0$ and $\alpha \in \mathbb{N}_0^n$. As $p_{0,0} = 2\|\cdot\|_\infty$, this set of seminorms is separating on $\mathcal{S}(\mathbb{R}^n)$. An equivalent set of seminorms is given by

$$p_{\alpha,Q}(\varphi) := \sup_{x \in \mathbb{R}^n} |Q(x)(\partial^\alpha \varphi)(x)|\,, \quad \varphi \in \mathcal{S}(\mathbb{R}^n),$$

where the Q's are (multivariate) polynomials. The functionals in $\mathcal{S}'(\mathbb{R}^n) := \mathcal{S}(\mathbb{R}^n)'$ are called *tempered distributions.*

Lemma B.20. *The following statements hold.*

(i) *For each $\alpha \in \mathbb{N}_0^n$, the map $\varphi \mapsto \partial^\alpha \varphi$ is continuous on $S(\mathbb{R}^n)$.*

(ii) *For each $\alpha \in \mathbb{N}_0^n$, the map $\varphi \mapsto x^\alpha \varphi$ is continuous on $S(\mathbb{R}^n)$.*

(iii) *There exists $C > 0$ such that for all $\varphi \in \mathcal{S}(\mathbb{R}^n)$ and all $y \in \mathbb{R}^n$ we have*

$$|(\mathcal{F}\varphi)(y)| \le C p_{0,n+1}(\varphi).$$

Proof. (i). Define $T : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ by $T\varphi := \partial^\alpha \varphi$, $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Then, for each $\beta \in \mathbb{N}_0^n$ and each $m \in \mathbb{N}_0$, we have

$$(p_{\beta,m} \circ T)(\varphi) = \sup_{x \in \mathbb{R}^n} (1 + |x|^m) \left|\partial^\beta \partial^\alpha \varphi(x)\right| = p_{\alpha+\beta,m}(\varphi), \quad \varphi \in \mathcal{S}(\mathbb{R}^n).$$

Hence, T is continuous by Lemma 5.14.

(ii). Define $T : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ by $T\varphi := x^\alpha \varphi$, $\varphi \in \mathcal{S}(\mathbb{R}^n)$. For each $\beta \in \mathbb{N}_0^n$, we have

$$\partial^\beta (x^\alpha \varphi)(x) = \sum_{\gamma \le \min\{\alpha,\beta\}} \binom{\beta}{\gamma} \binom{\alpha}{\gamma} \gamma!\, x^{\alpha-\gamma} (\partial^{\beta-\gamma} \varphi)(x), \quad \varphi \in \mathcal{S}(\mathbb{R}^n),\ x \in \mathbb{R}^n.$$

Thus, if $m \in \mathbb{N}_0$, then

$$\begin{aligned}(p_{\beta,m} \circ T)(\varphi) &= \sup_{x\in\mathbb{R}^n} (1+|x|^m) \left| \sum_{\gamma\le\min\{\alpha,\beta\}} \binom{\beta}{\gamma}\binom{\alpha}{\gamma} \gamma!\, x^{\alpha-\gamma}(\partial^{\beta-\gamma}\varphi)(x) \right| \\ &\le \sum_{\gamma\le\min\{\alpha,\beta\}} c_\gamma \sup_{x\in\mathbb{R}^n} (1+|x|^m)\, |x|^{m_\gamma} \left|(\partial^{\beta-\gamma}\varphi)(x)\right| \\ &\le 3 \sum_{\gamma\le\min\{\alpha,\beta\}} c_\gamma \sup_{x\in\mathbb{R}^n} (1+|x|^{mm_\gamma}) \left|(\partial^{\beta-\gamma}\varphi)(x)\right| \\ &= 3 \sum_{\gamma\le\min\{\alpha,\beta\}} c_\gamma\, p_{\beta-\gamma,mm_\gamma}(\varphi),\end{aligned}$$

where $c_\gamma = \binom{\beta}{\gamma}\binom{\alpha}{\gamma}\gamma!$ and $m_\gamma = |\alpha| - |\gamma|$. This proves the claim.

(iii). We observe

$$\begin{aligned}c^{-1}|(\mathcal{F}\varphi)(y)| &= \left|\int \varphi(x) e^{-ixy}\,\mathrm{d}x\right| \le \int |\varphi(x)|\,\mathrm{d}x \le \int \frac{\sup_{y\in\mathbb{R}^n}(1+|y|^{n+1})|\varphi(y)|}{1+|x|^{n+1}}\,\mathrm{d}x \\ &= \left(\int \frac{\mathrm{d}x}{1+|x|^{n+1}}\right) p_{0,n+1}(\varphi),\end{aligned}$$

where the integral $\int \frac{\mathrm{d}x}{1+|x|^{n+1}}$ converges. □

Theorem B.21. *$\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is a homeomorphism.*

Proof. Let $\alpha \in \mathbb{N}_0^n$, Q a polynomial, $Q(x) = \sum_\gamma a_\gamma x^\gamma$, and $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Define the differential operator $D := \sum_\gamma i^{-|\gamma|-|\alpha|} a_\gamma \partial^\gamma$, and recall that $\mathcal{F}(x^\alpha\varphi) = i^{|\alpha|}\partial^\alpha\mathcal{F}\varphi$ as well as $\mathcal{F}(\partial^\alpha\varphi) = i^{|\alpha|} y^\alpha \mathcal{F}\varphi$. Thus, we have

$$\begin{aligned}\mathcal{F}\left(D(x^\alpha\varphi)\right) &= \mathcal{F}\left(\sum_\gamma i^{-|\gamma|-|\alpha|} a_\gamma \partial^\gamma(x^\alpha\varphi)\right) = \sum_\gamma i^{-|\gamma|-|\alpha|} a_\gamma \mathcal{F}(\partial^\gamma x^\alpha \varphi) \\ &= \sum_\gamma a_\gamma i^{-|\alpha|} y^\gamma \mathcal{F}(x^\alpha\varphi) = \sum_\gamma a_\gamma y^\gamma \partial^\alpha \mathcal{F}\varphi = Q(y)\partial^\alpha\mathcal{F}\varphi.\end{aligned}$$

This implies

$$\begin{aligned}|Q(y)(\partial^\alpha\mathcal{F}\varphi)(y)| &= |\mathcal{F}(D(x^\alpha\varphi))(y)| \le C p_{0,n+1}(D(x^\alpha\varphi)) \\ &= C p_{0,n+1}\left(\sum_\gamma i^{-|\gamma|-|\alpha|} a_\gamma \partial^\gamma(x^\alpha\varphi)\right) \le C \cdot \sum_\gamma |a_\gamma| p_{0,n+1}(\partial^\gamma(x^\alpha\varphi)).\end{aligned}$$

Hence, we have proved that $p_{\alpha,Q} \circ \mathcal{F} \le p$, where p is a continuous seminorm on $\mathcal{S}(\mathbb{R}^n)$. Lemma 5.13 now shows that $p_{\alpha,Q} \circ \mathcal{F}$ is a continuous seminorm. Thus, Lemma 5.14 implies that $\mathcal{F}$ is continuous. Hence, also $\mathcal{F}^{-1} = \mathcal{F}^3$ is continuous. □

For a tempered distribution $T \in \mathcal{S}'(\mathbb{R}^n)$, we now define its Fourier transform by

$$(\mathcal{F}T)(\varphi) := T(\mathcal{F}\varphi), \quad \varphi \in \mathcal{S}(\mathbb{R}^n).$$

Alternatively, $\mathcal{F}T := T \circ \mathcal{F}$. Then $\mathcal{F}T \in \mathcal{S}'(\mathbb{R}^n)$.

References

[1] H.W. Alt, Lineare Funktionalanalysis. Eine anwendungsorientierte Einführung, 6th edition, Springer, 2012.

[2] J.B. Conway, A Course in Functional Analysis, 2nd edition, Springer, 1990.

[3] H. Heuser, Funktionalanalysis. Theorie und Anwendung, 4th edition, Teubner, 2006.

[4] F. Hirzebruch and W. Scharlau, Einführung in die Funktionalanalysis, 1st edition, Spektrum Akademischer Verlag, 1971.

[5] G. Kutyniok, Functional Analysis I, Lecture Notes, TU Berlin, 2012.

[6] F. Riesz and B. Nagy, Functional Analysis, Blackie & Son Ltd., 1955.

[7] W. Rudin, Functional Analysis, McGraw-Hill, 1991.

[8] W. Rudin, Real and Complex Analysis, 3rd edition, McGraw-Hill, 1986.

[9] J. Weidmann, Linear Operators in Hilbert Spaces, 1st edition, Springer, 1980.

[10] D. Werner, Funktionalanalysis, 7th edition, Springer, 2011.

[11] H. Winkler, Funktionalanalysis II, Lecture Notes, TU Berlin, 2007.