Lecture Notes
Spectral Theory

ROLAND SCHNAUBELT

These lecture notes are based on my course from the summer semester 2010. I kept the numbering and the contents of the results presented in the lectures (except for some minor corrections). Typically, the proofs and calculations in the notes are a bit shorter than those given in the lecture. Moreover, the drawings and many additional, mostly oral remarks from the lectures are omitted here. On the other hand, I have added a few lengthy proofs not shown in the lectures.

I like to thank Bernhard Konrad for his support in preparing these notes and Martin Meyries for pointing out a few mistakes and gaps in Section 4.3.

Karlsruhe, 6. August 2012

Roland Schnaubelt
Contents

1 General spectral theory .............................................. 3
   1.1 Closed operators ............................................ 3
   1.2 The spectrum ................................................. 7

2 Spectral theory of compact operators ......................... 17
   2.1 Compact operators ........................................... 17
   2.2 The Fredholm alternative .................................. 20

3 Sobolev spaces and weak derivatives ......................... 29
   3.1 Basic properties ............................................. 29
   3.2 Density and embedding theorems .......................... 39

4 Selfadjoint operators .............................................. 53
   4.1 Closed operators and their spectra, revisited .......... 53
   4.2 Selfadjoint operators and their spectra ............... 60
   4.3 The spectral theorems ..................................... 66

5 Holomorphic functional calculus ................................ 76
   5.1 The bounded case ............................................ 76
   5.2 Sectorial operators ......................................... 81

Bibliography .......................................................... 89
Chapter 1

General spectral theory

**General notation.** $X \neq \{0\}$, $Y \neq \{0\}$ and $Z \neq \{0\}$ are Banach spaces over $\mathbb{C}$ with norms $\|\cdot\|$ (or $\|\cdot\|_X$ etc.). The space

$$\mathcal{L}(X, Y) = \{ T : X \to Y : T \text{ is linear and continuous} \}$$

is endowed with the operator norm $\|T\| = \sup_{\|x\| \leq 1} \|Tx\|$, and we abbreviate $\mathcal{L}(X) := \mathcal{L}(X, X)$.

Let $D(A)$ be a linear subspace of $X$ and $A : D(A) \to Y$ be linear. Then $A$, or $(A, D(A))$, is called **linear operator from** $X$ to $Y$ (and on $X$ if $X = Y$) with domain $D(A)$. We denote by

$$N(A) = \{ x \in D(A) : Ax = 0 \},$$

and $R(A) = \{ y \in Y : \text{there exists some } x \in D(A) \text{ with } y = Ax \}$

the **kernel** and **range** of $A$.

### 1.1 Closed operators

**Example 1.1.** Let $X = C([0, 1])$ be endowed with the supremum norm and let $Af = f'$ with $D(A) = C^1([0, 1])$. Then $A$ is linear, but not bounded. Indeed, consider the functions $u_n \in D(A)$ given by $u_n(x) = (1/\sqrt{n}) \sin(nx)$ for $n \in \mathbb{N}$, which satisfy $\|u_n\|_\infty \to 0$ and

$$\|Au_n\|_\infty \geq |u_n'(0)| = \sqrt{n} \to \infty \quad \text{as } n \to \infty.$$

However, if $f_n \in D(A)$ satisfy $f_n \to f$ and $Af_n \to g$ in $C([0,1])$ as $n \to \infty$, then $f \in D(A) = C^1([0,1])$ and $Af = g$ (see Analysis 1 & 2).

**Definition 1.2.** Let $A$ be a linear operator from $X$ to $Y$. The operator $A$ is called **closed** if for all $x_n \in D(A)$, $n \in \mathbb{N}$, such that there exists $x = \lim_{n \to \infty} x_n$ in $X$ and $y = \lim_{n \to \infty} Ax_n$ in $Y$ it holds that $x \in D(A)$ and $Ax = y$.

*Hence, $\lim_{n \to \infty} (Ax_n) = A(\lim_{n \to \infty} x_n)$ if both $(x_n)$ and $(Ax_n)$ converge.*
Remark. It is clear that every operator \( A \in \mathcal{L}(X,Y) \) is closed (with \( D(A) = X \)). The operator \( A \) from Example 1.1 is closed.

**Example 1.3.** a) Let \( X = C([0,1]) \) and \( Af = f' \) with
\[
D(A) = \{ f \in C^1([0,1]) : f(0) = 0 \}.
\]
Let \( f_n \in D(A) \) and \( f,g \in X \) be such that \( f_n \to f \) and \( Af_n = f'_n \to g \) in \( X \) as \( n \to \infty \). Again by Analysis 1 & 2, there exists \( f \in C^1([0,1]) \) such that \( f' = g \). Since \( 0 = f_n(0) \to f(0) \) as \( n \to \infty \), we obtain \( f \in D(A) \). This means that \( A \) is closed on \( X \). In the same way we see that \( A_1 f = f' \) with
\[
D(A_1) = \{ f \in C^1([0,1]) : f(0) = f'(0) = 0 \}
\]
is closed. There are many more variants.

b) Let \( X = C([0,1]) \) and \( Af = f' \) with
\[
D(A) = C^1_c((0,1]) = \{ f \in C^1([0,1]) : \text{supp} \, f \subset (0,1] \},
\]
where the support \( \text{supp} \, f \) of \( f \) is the closure of \( \{ t \in [0,1] : f(t) \neq 0 \} \) in \( \mathbb{R} \).

This operator is not closed. In fact, consider the functions \( f_n \in D(A) \) given by
\[
f_n(t) = \begin{cases} 0, & 0 \leq t \leq \frac{1}{n}, \\ (t-\frac{1}{n})^2, & \frac{1}{n} \leq t \leq 1, \end{cases}
\]
for every \( n \in \mathbb{N} \). Then, \( f_n \to f \) and \( f'_n \to f' \) in \( X \) as \( n \to \infty \), where \( f(t) = t^2 \).

However, \( \text{supp} \, f = [0,1] \) and so \( f \notin D(A) \).

c) Let \( X = L^p(\mathbb{R}^d), 1 \leq p < \infty \), and \( m : \mathbb{R}^d \to \mathbb{C} \) be measurable. Define \( Af = mf \) with
\[
D(A) = \{ f \in X : mf \in X \}.
\]
This is the maximal domain. Then \( A \) is closed. Indeed, let \( f_n \to f \) and \( Af_n = mf_n \to g \) in \( X \) as \( n \to \infty \). Then there is a subsequence such that \( f_{n_j}(x) \to f(x) \) and \( m(x)f_{n_j}(x) \to g(x) \) for a.e. \( x \in \mathbb{R}^d \), as \( j \to \infty \). Hence, \( mf = g \) in \( L^p(\mathbb{R}^d) \) and we thus obtain \( f \in D(A) \) and \( Af = g \).

d) Let \( X = L^1([0,1]), Y = \mathbb{C} \), and \( Af = f(0) \) with \( D(A) = C([0,1]) \). Then \( A \) is not closed. In fact, consider the functions \( f_n \in D(A) \) given by
\[
f_n(t) = \begin{cases} 1 - nt, & 0 \leq t \leq \frac{1}{n}, \\ 0, & \frac{1}{n} \leq t \leq 1, \end{cases}
\]
for every \( n \in \mathbb{N} \). Then \( \|f_n\|_1 = 1/(2n) \to 0 \) as \( n \to \infty \), but \( Af_n = f_n(0) = 1 \).

**Definition 1.4.** Let \( A \) be a linear operator from \( X \) to \( Y \). The graph of \( A \) is given by
\[
\text{gr}(A) = \{ (x,Ax) \in X \times Y : x \in D(A) \}.
\]
The graph norm of \( A \) is defined by \( \|x\|_A = \|x\|_X + \|Ax\|_Y \). We write \([D(A)]\) if we equip \( D(A) \) with \( \|\cdot\|_A \).

4
1.1. CLOSED OPERATORS

Of course, $\|\cdot\|_A$ is equivalent to $\|\cdot\|_X$ if $A \in \mathcal{L}(X,Y)$. We endow $X \times Y$ with the norm $\|(x,y)\|_{X \times Y} = \|x\|_X + \|y\|_Y$.

**Lemma 1.5.** For a linear operator $A$ from $X$ to $Y$ the following assertions hold.

1. $\text{gr}(A) \subset X \times Y$ is a linear subspace.

2. $[D(A)]$ is a normed vector space and $A \in \mathcal{L}([D(A)], Y)$.

3. $A$ is closed if and only if $\text{gr}(A)$ is closed in $X \times Y$ if and only if $[D(A)]$ is a Banach space.

4. Let $A$ be injective and put $D(A^{-1}) := R(A)$. Then, $A$ is closed from $X$ to $Y$ if and only if $A^{-1}$ is closed from $Y$ to $X$.

**Proof.** Assertions 1) and 2) are straightforward to check.

3) The operator $A$ is closed if and only if for all $x_n \in D(A)$, $n \in \mathbb{N}$, and $(x,y) \in X \times Y$ with $(x_n, Ax_n) \to (x,y)$ in $X \times Y$ as $n \to \infty$, we have $(x,y) \in \text{gr}(A)$. This property is equivalent to the closedness of $\text{gr}(A)$. Since $\|(x, Ax)\|_{X \times Y} = \|x\|_X + \|Ax\|_Y$, a Cauchy sequence or a converging sequence in $\text{gr}(A)$ corresponds to a Cauchy or a converging sequence in $[D(A)]$, respectively. Hence, $[D(A)]$ is complete if and only if $\text{gr}(A)$ is complete if and only if $\text{gr}(A) \subset X \times Y$ is closed.

4) Assertion 4) follows from 3) since

$$\text{gr}(A^{-1}) = \{(y, A^{-1}y) : y \in R(A)\} = \{(Ax, x) : x \in D(A)\}$$

is closed in $Y \times X$ if and only if $\text{gr}(A)$ is closed in $X \times Y$. \hfill \Box

**Theorem 1.6** (Closed Graph Theorem). Let $X$ and $Y$ be Banach spaces and $A$ be a closed operator from $X$ to $Y$. Then $A$ is bounded (i.e., $\|Ax\| \leq c\|x\|$ for some $c \geq 0$ and all $x \in D(A)$) if and only if $D(A)$ is closed in $X$.

In particular, a closed operator with $D(A) = X$ already belongs to $\mathcal{L}(X, Y)$.

**Proof.** “$\Leftarrow$”: Let $D(A)$ be closed in $X$. Then $D(A)$ is a Banach space for $\|\cdot\|_X$ and $\|\cdot\|_A$. Since $\|x\|_X \leq \|x\|_A$ for all $x \in D(A)$, a corollary to the open mapping theorem (see e.g. Corollary 3.17 in [FA]) shows that there is some $c > 0$ such that $\|Ax\|_Y \leq \|x\|_A \leq c\|x\|_X$ for all $x \in D(A)$.

“$\Rightarrow$”: Let $A$ be bounded and let $x_n \in D(A)$ converge to $x \in X$ with respect to $\|\cdot\|_X$. Then $\|Ax_n - Ax_m\|_Y \leq c\|x_n - x_m\|_X$, and so the sequence $(Ax_n)_n$ is Cauchy in $Y$. Thus there exists $y := \lim_{n \to \infty} Ax_n$ in $Y$. The closedness of $A$ shows that $x \in D(A)$; i.e., $D(A)$ is closed in $X$. \hfill \Box

**Remark.** a) Theorem 1.6 is wrong without completeness. Consider for instance the operator $T$ given by $(Tf)(t) = tf(t)$, $t \in \mathbb{R}$, on $C_c(\mathbb{R})$ with supremum norm. This linear operator is everywhere defined, unbounded and closed: Take $f_n, f, g \in C_c(\mathbb{R})$ such that $f_n(t) \to f(t)$ and $(Tf_n)(t) = tf_n(t) \to g(t)$ uniformly for $t \in \mathbb{R}$ as $n \to \infty$. Then $g(t) = tf(t)$ for all $t \in \mathbb{R}$ and so $g = Tf$. 

5
b) Let \( X \) be an infinite dimensional Banach space and let \( B \) be an algebraic basis of \( X \) (i.e., for all \( x \in X \) there are unique \( \alpha_1, \ldots, \alpha_n \in \mathbb{C} \), \( b_1, \ldots, b_n \in B \) and \( n \in \mathbb{N} \) such that \( x = \alpha_1 b_1 + \cdots + \alpha_n b_n \)). We may assume that \( \|b\| = 1 \) for all \( b \in B \). Choose a countable subset \( B_0 = \{ b_k : k \in \mathbb{N} \} \) of \( B \) and set
\[
Tb_k = kb_k \quad \text{for each } \; b_k \in B_0, \quad \text{and } \; Tb = 0 \quad \text{for each } \; b \in B\setminus B_0.
\]
Then \( T \) can be extended to a linear operator on \( X \) which is unbounded, since \( \|Tb_k\| = k \) and \( \|b_k\| = 1 \). Thus \( T \) is not closed.

**Proposition 1.7.** Let \( A \) be closed from \( X \) to \( Y \), \( T \in \mathcal{L}(X,Y) \), and \( S \in \mathcal{L}(Z,X) \). Then the following operators are closed.

1. \( B = A + T \) with \( D(B) = D(A) \),
2. \( C = AS \) with \( D(C) = \{ z \in Z : Sz \in D(A) \} \).

**Proof.** a) Let \( x_n \in D(B), \; n \in \mathbb{N}, \) and \( x \in X, \; y \in Y \) such that \( x_n \to x \) in \( X \) and \( Bx_n = Ax_n + Tx_n \to y \) in \( Y \) as \( n \to \infty \). Since \( T \) is bounded, there exists \( Tx = \lim_{n \to \infty} Tx_n \) and so \( Ax_n \to y - Tx \) as \( n \to \infty \). Since \( A \) is closed, we deduce that \( x \in D(A) = D(B) \) and \( Ax = y - Tx \); i.e., \( Bx = Ax + Tx = y \).

b) Let \( z_n \in D(C), \; n \in \mathbb{N}, \) and \( z \in Z, \; y \in Y \) such that \( z_n \to z \) in \( Z \) and \( ASz_n \to y \) in \( Y \) as \( n \to \infty \). Since \( S \) is bounded, \( x_n := Sz_n \) converges to \( Sz \). Since \( A x_n \to y \) and \( A \) is closed, we obtain \( Sz \in D(A) \) and \( ASz = y \); i.e., \( z \in D(C) \) and \( Cz = y \). \( \square \)

**Corollary 1.8.** Let \( A \) be linear on \( X \) and \( \lambda \in \mathbb{C} \). Then the following assertions hold.

1. If \( \lambda I - A \) (or \( \lambda I + A \)) is closed, then \( A \) is closed.
2. If \( \lambda I - A \) is bijective with \( (\lambda I - A)^{-1} \in \mathcal{L}(X) \), then \( A \) is closed.

**Proof.** Assertion 1) is a consequence of Proposition 1.7 since \( A = \pm((\lambda I \pm A) - \lambda I) \). For 2), Lemma 1.5 shows that \( \lambda I - A \) is closed, and then assertion 1) yields 2). \( \square \)

The following examples show that closedness can be lost when taking sums or products of closed operators. See the exercises for further related results.

**Example 1.9.** a) Let \( E = C_b(\mathbb{R}^2) \) and \( A_k = \partial_k \) with
\[
D(A_k) = \{ f \in E : \; \text{the partial derivative } \partial_k f \text{ exists and belongs to } E \},
\]
for \( k = 1, 2 \). Set \( B = \partial_1 + \partial_2 \) on
\[
D(B) := D(A_1) \cap D(A_2) = C^1_b(\mathbb{R}^2) = \{ f \in C^1(\mathbb{R}^2) : \; f, \partial_1 f, \partial_2 f \in E \}.
\]
By an exercise, \( A_1 \) and \( A_2 \) are closed.

However, \( B \) is not closed: Take \( \phi_n \in C^1_b(\mathbb{R}) \) converging uniformly to some \( \phi \in C_b(\mathbb{R}) \setminus C^1(\mathbb{R}) \). Set \( f_n(x,y) = \phi_n(x-y) \) and \( f(x,y) = \phi(x-y) \) for \( (x,y) \in \mathbb{R}^2 \) and
1.2. THE SPECTRUM

Definition 1.10. Let $A$ be a closed operator on $X$. The resolvent set of $A$ is given by

$$\rho(A) = \{ \lambda \in \mathbb{C}; \lambda I - A : D(A) \to X \text{ is bijective} \},$$

and its spectrum by

$$\sigma(A) = \mathbb{C} \setminus \rho(A).$$

We further define the point spectrum of $A$ by

$$\sigma_p(A) = \{ \lambda \in \mathbb{C} : \text{there exists some } v \in D(A) \setminus \{ 0 \} \text{ with } \lambda v = Av \} \subset \sigma(A),$$

where we call $\lambda \in \sigma_p(A)$ an eigenvalue of $A$ and the corresponding $v$ an eigenvector or eigenfunction of $A$. For $\lambda \in \rho(A)$ the operator

$$R(\lambda, A) := (\lambda I - A)^{-1} : X \to X$$

and the set $\{ R(\lambda, A) : \lambda \in \rho(A) \}$ are called the resolvent.

Remark 1.11. a) Let $A$ be closed in $X$ and $\lambda \in \rho(A)$. The resolvent $R(\lambda, A)$ has the range $D(A)$. Lemma 1.5 and Corollary 1.8 further show that $R(\lambda, A)$ is closed and thus it belongs to $\mathcal{L}(X)$ by Theorem 1.6.

b) Let $A$ be a linear operator such that $\lambda I - A : D(A) \to X$ has a bounded inverse for some $\lambda \in \mathbb{C}$. Then $A$ is closed by Corollary 1.8. In this case, the closedness assumption in Definition 1.10 is redundant.

We set $e^{\lambda t} = e^{\lambda t}$ for $\lambda \in \mathbb{C}$, $t \in J$, and any interval $J \subset \mathbb{R}$.

Example 1.12. a) Let $X = \mathbb{C}^d$ and $T \in \mathcal{L}(X)$. Then $\sigma(T)$ only consists of the eigenvalues $\lambda_1, \ldots, \lambda_m$ of $T$, where $m \leq d$.

b) Let $X = C([0, 1])$, and $Au = u'$ with $D(A) = C^1([0, 1])$. Then $e^{\lambda t} \in D(A)$ and $Ae^{\lambda t} = \lambda e^{\lambda t}$ for each $\lambda \in \mathbb{C}$. Hence, $\lambda \in \sigma_p(A)$ and so $\sigma(A) = \sigma_p(A) = \mathbb{C}$.

c) Let $X = C([0, 1])$, and $Au = u'$ with $D(A) = \{ u \in C^1([0, 1]) : u(0) = 0 \}$. Let $\lambda \in \mathbb{C}$ and $f \in X$. We then have $u \in D(A)$ and $(\lambda I - A)u = f$ if and only if $u \in C^1([0, 1])$, $u'(t) = \lambda u(t) - f(t)$, $t \in [0, 1]$, and $u(0) = 0$, which is equivalent to

$$u(t) = - \int_0^t e^{\lambda(t-s)} f(s) ds.$$
CHAPTER 1. GENERAL SPECTRAL THEORY

for all $0 \leq t \leq 1$. Hence, $\sigma(A) = \emptyset$ and the resolvent is given by

$$R(\lambda, A)f(t) = -\int_0^t e^{\lambda(t-s)} f(s)ds,$$

for all $0 \leq t \leq 1$, $f \in X$, and $\lambda \in \mathbb{C}$.

Let $U \subset \mathbb{C}$ be open. The derivative of a function $f : U \to Y$ at $\lambda \in U$ is given by

$$f'(\lambda) = \lim_{\mu \to \lambda} \frac{1}{\mu - \lambda} (f(\mu) - f(\lambda)) \in Y,$$

if the limit exists in $Y$.

**Theorem 1.13.** Let $A$ be a closed operator on $X$ and let $\lambda \in \rho(A)$. Then the following assertions hold.

1. $AR(\lambda, A) = \lambda R(\lambda, A) - I$, $AR(\lambda, A)x = R(\lambda, A)Ax$ for all $x \in D(A)$, and

$$\frac{1}{\mu - \lambda} (R(\lambda, A) - R(\mu, A)) = R(\lambda, A)R(\mu, A) = R(\mu, A)R(\lambda, A)$$

if $\mu \in \rho(A) \setminus \{\lambda\}$. The latter identity is called the resolvent equation.

2. The spectrum $\sigma(A)$ is closed, where $B(\lambda, 1/\|R(\lambda, A)\|) \subset \rho(A)$ and

$$R(\mu, A) = \sum_{n=0}^{\infty} (\lambda - \mu)^n R(\lambda, A)^{n+1} =: R_\mu,$$

if $|\lambda - \mu| < 1/\|R(\lambda, A)\|$. The series converges absolutely in $\mathcal{L}(X, [D(A)])$, uniformly on $B(\lambda, 6/\|R(\lambda, A)\|)$ for each $\delta \in (0, 1)$. Moreover,

$$\|R(\mu, A)\|_{\mathcal{L}(X, [D(A)])} \leq \frac{c(\lambda)}{1 - \delta}$$

for all $\mu \in B(\lambda, 6/\|R(\lambda, A)\|)$ and a constant $c(\lambda)$ depending only on $\lambda$.

3. The function $\rho(A) \to \mathcal{L}(X, [D(A)])$, $\lambda \mapsto R(\lambda, A)$, is infinitely often differentiable with

$$\left(\frac{d}{d\lambda}\right)^n R(\lambda, A) = (-1)^n n! R(\lambda, A)^{n+1} \quad \text{for every } n \in \mathbb{N}.$$

4. $$\|R(\lambda, A)\| \geq \frac{1}{d(\lambda, \sigma(A))}.$$
1.2. THE SPECTRUM

Proof. 1) The first assertions follow from $x = (\lambda I - A)R(\lambda, A)x = R(\lambda, A)(\lambda I - A)x$, where $x \in X$ in the first equality and $x \in D(A)$ in the second one. For $\mu \in \rho(A)$ we further have

\[
(\lambda R(\lambda, A) - AR(\lambda, A))R(\mu, A) = R(\mu, A), \\
R(\lambda, A)(\mu R(\mu, A) - AR(\mu, A)) = R(\lambda, A).
\]

The resolvent equation then follows by subtraction and interchanging $\lambda$ and $\mu$.

2) Let $|\mu - \lambda| \leq \delta/\|R(\lambda, A)\|$ for some $\delta \in (0, 1)$ and $x \in X$ with $\|x\| \leq 1$. We then have

\[
\|(\lambda - \mu)^n R(\lambda, A)^{n+1}x\|_A \leq \frac{\delta^n}{\|R(\lambda, A)\|_A} \left(\|AR(\lambda, A)R(\lambda, A)^{n}x\| + \|R(\lambda, A)^{n+1}x\|\right) \\
\leq \delta^n \left(\|AR(\lambda, A)\| + 1 + \|R(\lambda, A)\|\right) =: \delta^n c(\lambda),
\]

where we used 1). So the series in 2) converges absolutely in $L(\lambda, [D(A)])$ uniformly on $B(\lambda, \delta/\|R(\lambda, A)\|)$ and can be estimated in norm by $c(\lambda)(1 - \delta)^{-1}$ (cf. Proposition 3.12 in [FA]). Using also $(\mu I - A)R(\lambda, A) = (\mu - \lambda)R(\lambda, A) + I$, we obtain

\[
(\mu I - A)R_\mu = -\sum_{n=0}^{\infty} (\lambda - \mu)^n R(\lambda, A)^{n+1} + \sum_{n=0}^{\infty} (\lambda - \mu)^n R(\lambda, A)^n = I,
\]

and similarly $R_\mu(\mu I - A)x = x$ for all $x \in D(A)$. Hence, $\mu \in \rho(A)$ and $R_\mu = R(\mu, A)$.

3) Since $\lambda \mapsto R(\lambda, A) \in \mathcal{L}(X, [D(A)])$ is locally bounded, due to the estimate in 2), the resolvent equation implies that the map $\lambda \mapsto R(\lambda, A) \in \mathcal{L}(X, [D(A)])$ is continuous. The resolvent equation then also yields assertion 3) for $n = 1$. Assume that 3) holds for some $n \in \mathbb{N}$. We then obtain

\[
\left(\frac{d}{d\lambda}\right)^{n+1} R(\lambda, A) = \frac{d}{d\lambda} \left((-1)^n n! R(\lambda, A)^{n+1}\right).
\]

Using the formula

\[
R(\mu, A)^{n+1} - R(\lambda, A)^{n+1} = \sum_{j=0}^{n} R(\mu, A)^{n-j}(R(\mu, A) - R(\lambda, A))R(\lambda, A)^j,
\]

the continuity of $R(\cdot, A)$ and the assertion for $n = 1$, we then conclude that 3) holds for $n + 1$.

4) Assertion 4) follows from 2). \qed

Proposition 1.14. Let $\Omega \subset \mathbb{R}^d$, $m \in C(\Omega)$, $X = C_b(\Omega)$, and $Af = mf$ with $D(A) = \{f \in X : mf \in X\}$. Then $A$ is closed,

\[
\sigma(A) = \overline{m(\Omega)},
\]

and $R(\lambda, A)f = \frac{1}{\lambda - m}f$ for all $\lambda \in \rho(A)$ and $f \in X$.

In particular, for every closed (compact) subset $S \subset \mathbb{C}$ there is a closed (bounded) operator $B$ on a Banach space with $\sigma(B) = S$. 

9
Chapter 1. General Spectral Theory

Proof. The closedness of $A$ can be shown as in the remark on page 5. Let $\lambda \notin \overline{m(\Omega)}$ and $g \in C_0(\Omega)$. The function $f := \frac{1}{\lambda - m} g$ then belongs to $C_0(\Omega)$ and $\lambda f - m f = g$ so that $m f = \lambda f - g \in C_0(\Omega)$. As a result, $f \in D(A)$ and $f$ is the unique solution in $D(A)$ of the equation $\lambda f - Af = g$. This means that $\lambda \in \rho(A)$ and $R(\lambda, A) g = \frac{1}{\lambda - m} g$.

In the case that $\lambda = m(z)$ for some $z \in \Omega$, we obtain

$$((\lambda I - A)f)(z) = \lambda f(z) - m(z)f(z) = 0$$

for every $f \in D(A)$. Consequently, $\lambda I - A$ is not surjective and so $\lambda \in \sigma(A)$. We now conclude that $\sigma(A) = \overline{m(\Omega)}$ since the spectrum is closed.

The final assertion follows from Example 1.12c) if $S = \emptyset$. Otherwise, consider $\Omega = S$. Define $A$ and $X$ as above. Then, $\sigma(A) = S$ and $A$ is bounded if $S$ is compact (where $C_b(S) = C(S)$).

A similar result holds on $L^p$-spaces, see e.g. Example IX.2.6 in [Con90].

Example 1.15. Let $X = C_0(\mathbb{R}_+) = \{ f \in C(\mathbb{R}_+) : \lim_{t \to \infty} f(t) = 0 \}$ with the supremum norm and $Af = f'$ with

$$D(A) = C_0^1(\mathbb{R}_+) = \{ f \in C^1(\mathbb{R}_+) : f, f' \in X \}.$$ 

As in Example 1.3 one sees that $A$ is closed. Let $\lambda \in \mathbb{C}$ with $\text{Re} \lambda > 0$ and $f \in X$. We then have $u \in D(A)$ and $\lambda u - Au = f$ if and only if $u \in X \cap C^1(\mathbb{R}_+)$ and $u'(t) = \lambda u(t) - f(t)$ for all $t \geq 0$. This equation is uniquely solved by

$$u(t) = \int_t^\infty e^{\lambda(t-s)} f(s) ds =: (R_\lambda f)(t),$$

for $t \geq 0$.

We have to show that $R_\lambda f \in X$. Let $\varepsilon > 0$. Then there is an $t_\varepsilon \geq 0$ such that $|f(s)| \leq \varepsilon$ for all $s \geq t_\varepsilon$. Hence

$$|R_\lambda f(t)| \leq \int_t^\infty e^{(\text{Re} \lambda)(t-s)} |f(s)| ds \leq \varepsilon \int_0^\infty e^{-s \text{Re} \lambda} dr = \frac{\varepsilon}{\text{Re} \lambda},$$

for all $t \geq t_\varepsilon$, where we substituted $r = s - t$. As a result, $R_\lambda f \in C_0(\mathbb{R}_+)$ and so $\lambda \in \rho(A)$ with $R_\lambda = R(\lambda, A)$. If $\text{Re} \lambda < 0$, then $e_\lambda \in X$ and $e'_\lambda = \lambda e_\lambda \in X$. Hence, $e_\lambda$ is an eigenfunction for the eigenvalue $\lambda$ and $\{ \lambda \in \mathbb{C} : \text{Re} \lambda < 0 \} \subset \sigma(A)$. Since $\sigma(A)$ is closed, we deduce that

$$\{ \lambda \in \mathbb{C} : \text{Re} \lambda \leq 0 \} = \sigma(A).$$

Theorem 1.16. Let $T \in \mathcal{L}(X)$. Then $\sigma(T)$ is a non-empty compact set. The spectral radius $r(T) := \max \{ |\lambda| : \lambda \in \sigma(T) \}$ is given by

$$r(T) = \lim_{n \to \infty} \|T^n\|^{1/n} = \inf_{n \in \mathbb{N}} \|T^n\|^{1/n} \leq \|T\|,$$

and for $\lambda \in \mathbb{C}$ with $|\lambda| > r(T)$ we have

$$R(\lambda, T) = \sum_{n=0}^\infty \lambda^{-n-1} T^n =: R_\lambda.$$
1.2. THE SPECTRUM

Proof. 1) Since \( \|T^{n+m}\| \leq \|T^n\| \|T^m\| \) for all \( n, m \in \mathbb{N} \), an elementary lemma (see Lemma VI.1.4 in [Wer05]) yields that there exists

\[
\lim_{n \to \infty} \|T^n\|^{1/n} = \inf_{n \in \mathbb{N}} \|T^n\|^{1/n} =: r \leq \|T\|.
\]

If \( |\lambda| > r \), then

\[
\limsup_{n \to \infty} |\lambda^{-n}T^n|^{1/n} = \frac{1}{|\lambda|} \lim_{n \to \infty} \|T^n\|^{1/n} = \frac{r}{|\lambda|} < 1.
\]

Therefore the series \( R_\lambda \) converges absolutely in \( \mathcal{L}(X) \), and uniformly for \( \lambda \) in all compact subsets of \( \mathbb{C} \setminus B(0, r) \) (proof as in Analysis 1). Moreover,

\[
(\lambda I - T)R_\lambda = \sum_{n=0}^{\infty} \lambda^{-n}T^n - \sum_{n=0}^{\infty} \lambda^{-n-1}T^{n+1} = I,
\]

and similarly \( R_\lambda(\lambda I - T) = I \). Hence, \( \lambda \in \rho(T) \) and \( R_\lambda = R(\lambda, T) \). Due to its closedness, the spectrum \( \sigma(T) \subset \overline{B}(0, r) \) is compact. Therefore, \( r(T) \) exists as the maximum of a compact subset of \( \mathbb{R} \), and \( r(T) \leq r \).

2) Take \( s > r(T) \), \( \Phi \in \mathcal{L}(X)^* \), and \( m \in \mathbb{N} \). We define \( f_\Phi(\lambda) = \Phi(R(\lambda, T)) \) for \( \lambda \in \rho(T) \). Note that \( f_\Phi : \rho(T) \to \mathbb{C} \) is complex differentiable. We set

\[
C_m(\Phi) = \frac{1}{2\pi i} \int_{|\lambda|=s} \lambda^m \Phi(R(\lambda, T)) d\lambda.
\]

Since \( f_\Phi \) is holomorphic, this integral does not depend on \( s > r(T) \) due to complex analysis. So we may choose for a moment \( s > r \) and use the uniformly convergent series of step 1) to deduce

\[
C_m(\Phi) = \sum_{n=0}^{\infty} \frac{1}{2\pi i} \int_{|\lambda|=s} \lambda^{m-n-1} d\lambda \Phi(T^n)
= \sum_{n=0}^{\infty} \frac{1}{2\pi i} \int_0^{2\pi} (se^{it})^{m-n-1} ise^{it} dt \Phi(T^n) = \Phi(T^m).
\]

Employing the Hahn-Banach theorem, we now choose a functional \( \Phi_m \in \mathcal{L}(X)^* \) with \( \|\Phi_m\| = 1 \) and \( \Phi_m(T^m) = \|T^m\| \) (see e.g. Corollary 4.9 in [FA]). Again for any \( s > r(T) \), we can then estimate

\[
\|T^m\| = C_m(\Phi_m) \leq \frac{1}{2\pi} \int_0^{2\pi} |se^{it}| m \|\Phi_m\| \|R(se^{it}, T)\| |se^{it}| dt
\leq s^m s \max_{|\lambda|=s} \|R(\lambda, T)\| =: c(s)s^m.
\]

Thus, \( \|T^m\|^{1/m} \leq sc(s)^{1/m} \) and so \( r \leq s \). This means that \( r(T) = r \).

Finally, suppose that \( \sigma(T) = \emptyset \). Then the functions \( f_\Phi \) are entire for every \( \Phi \in \mathcal{L}(X)^* \). Moreover, step 1) yields

\[
|f_\Phi(\lambda)| \leq \|\Phi\| |\lambda|^{-1} \sum_{n=0}^{\infty} \frac{\|T^n\|}{|\lambda|^n} \leq \frac{2\|\Phi\|}{|\lambda|},
\]

11
for all \( \lambda \in \mathbb{C} \) with \( |\lambda| \geq 2\|T\| \). Hence, \( f_\lambda \) is bounded and thus constant by Liouville’s theorem from complex analysis. The above estimate then shows that \( \Phi(R(\lambda, T)) = 0 \) for all \( \lambda \in \mathbb{C} \) and \( \Phi \in \mathcal{L}(X^*) \). Using again the Hahn-Banach theorem (see e.g. Corollary 4.9 in [FA]), we obtain \( R(\lambda, T) = 0 \), which is impossible since \( R(\lambda, T) \) is injective and \( X \neq \{0\} \).

Example 1.17. a) Let \( X = C([0, 1]) \) and define the Volterra operator \( V \) on \( X \) by

\[
Vf(t) = \int_0^t f(s)\,ds
\]

for \( t \in [0, 1] \) and \( f \in X \). Then \( V \in \mathcal{L}(X) \) and

\[
|V^n f(t)| \leq \int_0^t \int_0^{s_1} \cdots \int_0^{s_{n-1}} \|f\|_\infty ds_1 \cdots ds_n \leq \frac{1}{n!}\|f\|_\infty,
\]

for all \( n \in \mathbb{N} \), \( t \in [0, 1] \), and \( f \in X \). Hence, \( \|V^n\| \leq 1/n! \). For \( f = 1 \) we obtain \( \|V^n\| \geq \|V^n 1\|_\infty = 1/n! \) and so \( \|V^n\| = 1/n! \). This gives

\[
r(V) = \lim_{n \to \infty} \left( \frac{1}{n!} \right)^{1/n} = 0 \quad \text{and} \quad \sigma(V) = \{0\}.
\]

Observe that \( \sigma_p(V) = \emptyset \) since \( Vf = 0 \) implies that \( f = (Vf)' = 0 \). Moreover, \( \|V\| = 1 > r(V) = 0 \).

b) Let left shift \( L \) given by \( Lx = (x_{n+1}) \) on \( X \in \{c_0, \ell^p : 1 \leq p \leq \infty\} \) has the spectrum

\[
\sigma(L) = \overline{B}(0, 1).
\]

In fact, \( L \in \mathcal{L}(X) \) has norm 1 (see e.g. Example 1.57 in [FA]), and so \( \sigma(L) \subset \overline{B}(0, 1) \). Moreover, \( L(1,0,\cdots) = 0 \), and for \( |\lambda| < 1 \) the sequence \( v = (\lambda^n)_{n \in \mathbb{N}} \) belongs to \( X \) and satisfies \( \lambda v = Lv \) so that \( B(0,1) \subset \sigma_p(L) \subset \sigma(L) \). The closedness of \( \sigma(L) \) then yields \( \sigma(L) = \overline{B}(0, 1) \). Note that \( \sigma_p(L) = \overline{B}(0,1) \) if \( X = \ell^\infty \), but \( \sigma_p(L) = B(0,1) \) in the other cases.

Definition 1.18. Let \( A \) be a closed operator on \( X \). Then we call

\[
\sigma_{ap}(A) = \{\lambda \in \mathbb{C} : \text{there exist } x_n \in D(A) \text{ with } \|x_n\| = 1 \text{ for all } n \in \mathbb{N}, \text{ and } \lambda x_n - Ax_n \to 0 \text{ as } n \to \infty\}
\]

the approximate point spectrum of \( A \) and

\[
\sigma_r(A) = \{\lambda \in \mathbb{C} : (\lambda I - A)D(A) \text{ is not dense in } X\}
\]

the residual spectrum of \( A \).

Proposition 1.19. For a closed operator \( A \) on \( X \) the following assertions hold.

1. \( \sigma_{ap}(A) = \sigma_p(A) \cup \{\lambda \in \mathbb{C} : (\lambda I - A)D(A) \text{ is not closed in } X\} \).

2. \( \sigma(A) = \sigma_{ap}(A) \cup \sigma_r(A) \).
1.2. THE SPECTRUM

3. \( \partial \sigma(A) \subset \sigma_{ap}(A) \).

(Note that the unions need not be disjoint.)

Proof. 1) We have \( \lambda \notin \sigma_{ap}(A) \) if and only if there is a \( c > 0 \) such that \( \| (\lambda I - A)x \| \geq c \| x \| \) for all \( x \in D(A) \). This lower estimate implies that \( \lambda \notin \sigma_p(A) \). Moreover, if \( y_n = \lambda x_n - Ax_n \to y \) in \( X \) as \( n \to \infty \) for some \( x_n \in D(A) \), then the lower estimate shows that \( (x_n) \) is Cauchy in \( X \), so that \( x_n \) converges to some \( x \in X \). Hence, \( Ax_n = \lambda x_n - y_n \to \lambda x - y \) and the closedness of \( A \) yields \( x \in D(A) \) and \( \lambda x - Ax = y \). Consequently, \( (\lambda I - A)D(A) \) is closed.

Conversely, if \( (\lambda I - A)D(A) \) is closed and \( \lambda \notin \sigma_p(A) \), then the inverse \( (\lambda I - A)^{-1} \) exists and is closed on its closed domain \( (\lambda I - A)D(A) \). The closed graph theorem 1.6 then yields the boundedness of \( (\lambda I - A)^{-1} \). Thus,

\[
\| x \| = \| (\lambda I - A)^{-1}(\lambda I - A)x \| \leq C\| (\lambda I - A)x \|
\]

for all \( x \in D(A) \) and a constant \( C > 0 \). This means that \( \lambda \notin \sigma_{ap}(A) \).

2) Assertion 2) follows from 1).

3) Let \( \lambda \in \partial \sigma(A) \). Then there exist \( \lambda_n \in \rho(A) \) with \( \lambda_n \to \lambda \) as \( n \to \infty \). Due to Theorem 1.13(4), \( \| R(\lambda_n, A) \| \to \infty \) as \( n \to \infty \), and thus there are \( y_n \in X \) such that \( \| y_n \| = 1 \) for all \( n \in \mathbb{N} \) and \( a_n := \| R(\lambda_n, A)y_n \| \to \infty \) as \( n \to \infty \), where we can assume that \( a_n > 0 \) for all \( n \in \mathbb{N} \). Set \( x_n = \frac{1}{a_n} R(\lambda_n, A)y_n \in D(A) \). We then have \( \| x_n \| = 1 \) for all \( n \in \mathbb{N} \) and \( \lambda x_n - Ax_n = (\lambda - \lambda_n)x_n + \frac{1}{a_n} y_n \to 0 \) as \( n \to \infty \). As a result, \( \lambda \in \sigma_{ap}(A) \).

Proposition 1.20. Let \( A \) be closed on \( X \) and \( \lambda \in \rho(A) \). Then the following assertions hold.

1. \( \sigma(R(\lambda, A))\setminus\{0\} = (\lambda - \sigma(A))^{-1} = \{ \frac{1}{\lambda - \mu} : \mu \in \sigma(A) \} \).
2. \( \sigma_j(R(\lambda, A))\setminus\{0\} = (\lambda - \sigma_j(A))^{-1} \) for \( j = p, ap, r \).
3. If \( x \) is an eigenvector for the eigenvalue \( \mu \neq 0 \) of \( R(\lambda, A) \), then \( y = \mu R(\lambda, A)x \) is an eigenvector for the eigenvalue \( \nu = \lambda - 1/\mu \) of \( A \). If \( y \in D(A) \) is an eigenvector for the eigenvalue \( \nu = \lambda - 1/\mu \) of \( A \) with \( \mu \in \mathbb{C}\setminus\{0\} \), then \( x = \mu^{-1}(\lambda y - Ay) \) is an eigenvector for the eigenvalue \( \mu \) of \( R(\lambda, A) \).
4. \( r(R(\lambda, A)) = \frac{1}{d(\lambda, \sigma(A))} \).
5. If \( A \) is unbounded (i.e., \( D(A) \neq X \)), then \( 0 \in \sigma(R(\lambda, A)) \).

Proof. For \( \mu \in \mathbb{C}\setminus\{0\} \) we have

\[
\mu I - R(\lambda, A) = \left( \lambda - \mu \frac{1}{\mu} \right) I - A \mu R(\lambda, A).
\]

Since \( \mu R(\lambda, A) : X \to D(A) \) is bijective, we obtain assertion 2) for \( j = p \), part 3) and the equality of the ranges of the operators \( \mu I - R(\lambda, A) \) and \( (\lambda - 1/\mu)I - A \). Thus the assertions 1) and 2) for \( j = ap, r \) follow from Proposition 1.19. Assertion 4) is a consequence of 1). Finally, 5) is true because of \( R(\lambda, A)^{-1} = \lambda I - A \). \( \square \)
Example 1.21. a) Let $X = L^p(\mathbb{R})$, $1 \leq p \leq \infty$, and the translation $T(t)$ be given by $(T(t)f)(s) = f(s + t)$ for $s \in \mathbb{R}$, $f \in X$, and $t \in \mathbb{R}$. Recall from Example 3.8 in [FA], that $T(t)$ is an isometry on $X$ with inverse $(T(t))^{-1} = T(-t)$ for every $t \in \mathbb{R}$. By Theorem 1.16 we have $\sigma(T(t)) \subset \overline{B}(0, 1)$. Proposition 1.20 further yields $\sigma(T(t))^{-1} = \sigma(T(t)^{-1}) = \sigma(T(-t)) \subset \overline{B}(0, 1)$ so that $\sigma(T(t)) \subset \partial B(0, 1)$ for all $t \in \mathbb{R}$. Fix $t \neq 0$. For every $\lambda \in i\mathbb{R}$, the function $e_\lambda$ belongs to $C_0(\mathbb{R}) \subset L^\infty(\mathbb{R})$ and

$$(T(t)e_\lambda)(s) = e^{\lambda(s + t)} = e^{\lambda t}e_\lambda(s)$$

for all $s \in \mathbb{R}$. Hence, $\sigma(T(t)) = \sigma_p(T(t)) = \partial B(0, 1)$ for $p = \infty$.

If $p \in [1, \infty)$, we use $e_\lambda$ to construct an approximate eigenfunction. For $n \in \mathbb{N}$ set $f_n = n^{-1/2} \mathbb{I}_{[0,n]} e_\lambda$. We then have $\|f_n\|_p = n^{-1/2} \|\mathbb{I}_{[0,n]}\|_p = 1$ and (see above)

$$\|T(t)f_n - e^{\lambda t}f_n\|_p = n^{-1/2} e^{\lambda t}(1 - t_n - t) - 1_{[0,n]}\|_p = n^{-1/2} |2t|^{-1/p} \to 0,$$

as $n \to \infty$. As a result, $\sigma(T(t)) = \partial B(0, 1)$ if $t \neq 0$.

b) Let $X = C_0(\mathbb{R})$ and $Au = u'$ with $D(A) = C^1(\mathbb{R}) := \{f \in C^1(\mathbb{R}) : f, f' \in C_0(\mathbb{R})\}$. As in Example 1.15 one sees that $\lambda \in \rho(A)$ if $\Re \lambda \neq 0$ and

$$R(\lambda, A) f(t) = \int_t^\infty e^{\lambda(t-s)} f(s)ds, \text{ if } \Re \lambda > 0,$$

$$R(\lambda, A) f(t) = -\int_{-\infty}^t e^{\lambda(t-s)} f(s)ds, \text{ if } \Re \lambda < 0,$$

for all $t \in \mathbb{R}$ and $f \in X$. Let $\Re \lambda = 0$. Choose $\varphi_n \in C^1_c(\mathbb{R})$ with $\|\varphi_n\|_{\infty} \leq 1/n$ and $\|\varphi_n\|_{\infty} = 1$, and set $u_n = \varphi_n e_\lambda$ for all $n \in \mathbb{N}$. Then, $\|u_n\|_{\infty} = 1$, $u_n \in D(A)$ and $Au_n = \varphi_n' e_\lambda + \varphi_n e_\lambda + \lambda u_n$. Since $\|\varphi_n' e_\lambda\|_{\infty} \leq 1/n$, we obtain $\lambda \in \sigma_{ap}(A)$ and so $\sigma(A) = i\mathbb{R}$.

Definition 1.22. Let $A$ be a linear operator from $X$ to $Y$ with dense domain. We define its adjoint $A^*$ from $Y^*$ to $X^*$ by setting

$$D(A^*) = \{y^* \in Y^* : \exists z^* \in X^* \forall x \in D(A) : \langle Ax, y^* \rangle = \langle x, z^* \rangle\},$$

$$A^* y^* = z^*.$$

Observe that it holds

$$\langle Ax, y^* \rangle = \langle x, A^* y^* \rangle$$

for all $x \in D(A)$ and $y^* \in D(A^*)$.

Remark 1.23. Let $A$ be linear from $X$ to $Y$ with $\overline{D(A)} = X$.

a) Since $D(A)$ is dense, there is at most one vector $z^* = A^* y^*$ as in Definition 1.22, so that $A^* : D(A^*) \to X^*$ is a map. It is clear that $A^*$ is linear. If $A \in L(X, Y)$, then Definition 1.22 coincides with the definition of $A^*$ in §4.4 of [FA], where $D(A^*) = Y^*$.

b) The operator $A^*$ is closed from $Y^*$ to $X^*$.
1.2. THE SPECTRUM

Proof. Let \( y_n^* \in D(A^*), y^* \in Y^* \), and \( z_n^* \in X^* \) such that \( y_n^* \to y^* \) in \( Y^* \) and \( z_n^* := A^* y_n^* \to z^* \) in \( X^* \) as \( n \to \infty \). Let \( x \in D(A) \). It then holds that

\[
\langle x, z^* \rangle = \lim_{n \to \infty} \langle x, z_n^* \rangle = \lim_{n \to \infty} \langle Ax, y_n^* \rangle = \langle Ax, y^* \rangle.
\]

As a result, \( y^* \in D(A^*) \) and \( A^* y^* = z^* \).

\( \square \)

c) If \( T \in \mathcal{L}(X,Y) \), then the sum \( A + T \) with \( D(A + T) = D(A) \) has the adjoint \( (A + T)^* = A^* + T^* \) with \( D((A + T)^*) = D(A^*) \).

Proof. Let \( x \in D(A) \) and \( y^* \in Y^* \). We obtain

\[
\langle (A + T)x, y^* \rangle = \langle Ax, y^* \rangle + \langle x, T^* y^* \rangle.
\]

Hence, \( y^* \in D((A + T)^*) \) if and only if \( y^* \in D(A^*) \), and then \( (A + T)^* y^* = A^* y^* + T^* y^* \).

\( \square \)

Remark. The operator \( Af = f' \) with \( D(A) = \{ f \in C^1([0,1]) : f(0) = 0 \} \) is not densely defined on \( X = C([0,1]) \) since \( D(A) = \{ f \in X : f(0) = 0 \} \).

Theorem 1.24. Let \( A \) be a closed operator on \( X \) with dense domain. Then the following assertions hold.

1. \( \sigma_r(A) = \sigma_p(A^*) \).
2. \( \sigma(A) = \sigma(A^*) \) and \( R(\lambda, A^*) = R(\lambda, A^*) \) for every \( \lambda \in \rho(A) \).

Proof. 1) Due to a corollary of the Hahn-Banach theorem (see e.g. Proposition 4.11 in [FA]), the set \( (\lambda I - A)D(A) \) is not dense in \( X \) if and only if there is a vector \( y^* \in X^* \setminus \{0\} \) such that \( \langle \lambda x - Ax, y^* \rangle = 0 \) for every \( x \in D(A) \). This fact is equivalent to the equality \( \langle Ax, y^* \rangle = \langle x, \lambda y^* \rangle \) for every \( x \in D(A) \), which in turn means that \( y^* \in D(A^*) \setminus \{0\} \) and \( A^* y^* = \lambda y^* \); i.e., \( \lambda \in \sigma_p(A^*) \).

2) Let \( \lambda \in \rho(A) \). Take \( x \in D(A) \), \( x^* \in X^* \), and set \( y^* = R(\lambda, A)^* x^* \). We then obtain

\[
\langle (\lambda I - A)x, y^* \rangle = \langle R(\lambda, A)(\lambda I - A)x, x^* \rangle = \langle x, x^* \rangle.
\]

Thus, \( y^* \in D(A^*) \) and \( x^* = (\lambda I - A)^* y^* = (\lambda I - A^*) y^* \), where we use Remark 1.23. This means that \( \lambda I - A^* \) is surjective. Further, take \( x^* \in D(A^*) \) and \( x \in X \). We compute

\[
\langle x, R(\lambda, A)^*(\lambda I - A^*)x^* \rangle = \langle R(\lambda, A)x, (\lambda I - A^*)x^* \rangle = \langle (\lambda I - A)R(\lambda, A)x, x^* \rangle = \langle x, x^* \rangle,
\]

using Definition 1.22 and that \( R(\lambda, A)x \) belongs to \( D(A) \). Hence, \( R(\lambda, A)^*(\lambda I - A^*)x^* = x^* \) so that \( \lambda I - A^* \) is also injective. It thus exists \( R(\lambda, A^*) = R(\lambda, A)^* \).

Conversely, let \( \lambda \in \rho(A^*) \). Take \( x \in D(A) \). For every \( x^* \in X^* \), we compute as above

\[
\langle (\lambda I - A)x, R(\lambda, A^*)x^* \rangle = \langle x, (\lambda I - A^*)R(\lambda, A^*)x^* \rangle = \langle x, x^* \rangle.
\]

15
Due to a corollary of the Hahn-Banach theorem (see e.g. Corollary 4.9 in [FA]), there is a $y^* \in X^*$ such that $\|y^*\| = 1$ and $\langle x, y^* \rangle = \|x\|$. Hence,

$$\|x\| = \langle (\lambda I - A)x, R(\lambda, A^*)y^* \rangle \leq \|R(\lambda, A^*)\|\|\lambda x - Ax\|.$$ 

This estimate implies that $\lambda \notin \sigma_{ap}(A)$. Further, $\lambda \notin \sigma_p(A^*) = \sigma_c(A)$ by part a), and so Proposition 1.19 shows that $\lambda \notin \sigma(A)$. \hfill\(\square\)

**Example 1.25.** Let $X = \ell^p$, $1 \leq p < \infty$, or $X = c_0$. Let $Rx = (0, x_1, x_2, \ldots)$ be the right shift on $X$. Then $R^* = L$, where the left shift $L$ acts on $\ell^1$ if $X = c_0$ and on $\ell^p$ otherwise. Since $\sigma(L) = \overline{B}(0,1)$ by Example 1.17, Theorem 1.24 yields

$$\sigma(R) = \sigma(L) = \overline{B}(0,1).$$

Moreover, $\sigma_\ell(R) = \sigma_p(L) = B(0,1)$ if $X = c_0$ or $X = \ell^p$ with $1 < p < \infty$ and $\sigma_\ell(R) = \sigma_\ell(L) = \overline{B}(0,1)$ if $X = \ell^1$. If $X = \ell^\infty$, then $R = L^*$ for $L$ on $\ell^1$ so that again $\sigma(R) = \overline{B}(0,1)$.

Let $A$ be a linear operator. Then a linear operator $B$ from $X$ to $Y$ is called $A$-bounded if $D(A) \subset D(B)$ and $B \in \mathcal{L}([D(A)], Y)$.

**Theorem 1.26.** Let $A$ be a closed operator on $X$ and $\lambda \in \rho(A)$. Further, let $B$ be linear on $X$, $A$-bounded and satisfy $\|BR(\lambda, A)\| < 1$. Then $A + B$ with $D(A + B) = D(A)$ is closed, $\lambda \in \rho(A + B)$,

$$R(\lambda, A + B) = R(\lambda, A) \sum_{n=0}^{\infty} (BR(\lambda, A))^n = R(\lambda, A)(I - BR(\lambda, A))^{-1},$$

and

$$\|R(\lambda, A + B)\| \leq \frac{\|R(\lambda, A)\|}{1 - \|BR(\lambda, A)\|}.$$ 

**Proof.** Set $R_\lambda = R(\lambda, A) \sum_{n=0}^{\infty} (BR(\lambda, A))^n$. It is known that the Neumann series $\sum_{n=0}^{\infty} (BR(\lambda, A))^n$ converges to $(I - BR(\lambda, A))^{-1}$ and that the asserted norm estimate holds, see e.g. Proposition 3.12 in [FA]. Clearly, $R_\lambda$ maps into $D(A)$ and it holds

$$(\lambda I - A - B)R_\lambda = \sum_{n=0}^{\infty} (BR(\lambda, A))^n - \sum_{n=0}^{\infty} (BR(\lambda, A))^{n+1} = I.$$ 

Since $R_\lambda = \sum_{n=0}^{\infty} (R(\lambda, A)B)^n R(\lambda, A)$, we also obtain $R_\lambda(\lambda I - A - B)x = x$ for all $x \in D(A)$. As a result $\lambda \in \rho(A + B)$ and $R_\lambda = R(\lambda, A + B)$. \hfill\(\square\)

Observe that the smallness condition in this theorem is sharp in general: Let $X = \mathbb{C}$, $a \in \mathbb{C} \cong \mathcal{L}(\mathbb{C})$, $a \neq 0$, and $b = a$. Then $a$ is invertible, but $a - a = 0$ is not. Here we have $\lambda = 0$ and $|bR(0,a)| = |\frac{a}{a}| = 1$. 

16
Chapter 2

Spectral theory of compact operators

2.1 Compact operators

A nonempty subset $B \subset X$ is called relatively compact if its closure is compact in $X$. We will often use that this property holds if and only if each sequence in $B$ has a converging subsequence (with limit in $\overline{B}$). In fact, the necessity of the latter condition is clear. Conversely, assume that each sequence in $B$ has a converging subsequence. Let $(x_n)$ be a sequence in $\overline{B}$. Then, for each $n \in \mathbb{N}$ there exists a $y_n \in B$ with $\|x_n - y_n\| \leq 1/n$. By the assumption, we have a subsequence $(y_{n_j})_j$ with limit $y$ in $\overline{B}$. Consequently, $x_{n_j}$ converges to $y$ as $j \to \infty$.

**Definition 2.1.** A linear map $T : X \to Y$ is called compact if $TB(0,1)$ is relatively compact in $Y$. The set of all compact operators is denoted by $L_0(X,Y)$.

**Remark 2.2.** a) If $T$ is compact, then $TB(0,1)$ is bounded and thus $T$ is bounded; i.e., $L_0(X,Y) \subseteq L(X,Y)$.

b) Let $T : X \to Y$ be linear. Then the following assertions are equivalent.

(i) $T$ is compact.

(ii) $T$ maps bounded sets of $X$ into relatively compact sets of $Y$.

(iii) For every bounded sequence $(x_n)_n \subseteq X$ there exists a convergent subsequence $(Tx_{n_j})_j$ in $Y$.

**Proof.** (i) $\Rightarrow$ (ii): If $T$ is compact and $B \subset X$ is bounded, then $B \subseteq \overline{B}(0,r)$ for some $r \geq 0$ and $TB \subseteq \overline{TB}(0,r) = rTB(0,1)$, which is compact. Hence, $TB$ is compact. The implications (ii) $\Rightarrow$ (iii) $\Rightarrow$ (i) are clear.

c) The space of operators of finite rank is defined by

$$L_{00}(X,Y) = \{ T \in L(X,Y) : \dim TX < \infty \}.$$
CHAPTER 2. SPECTRAL THEORY OF COMPACT OPERATORS

The Heine Borel theorem yields that \( \mathcal{L}_{00}(X, Y) \subseteq \mathcal{L}_0(X, Y) \).

d) The identity \( I : X \to X \) is compact if and only if \( \overline{B}(0, 1) \) is compact if and only if \( \dim X < \infty \) (see e.g. Theorem 1.40 in [FA]).

**Proposition 2.3.** \( \mathcal{L}_0(X, Y) \) is a closed linear subspace of \( \mathcal{L}(X, Y) \). Let \( T \in \mathcal{L}(X, Y) \) and \( S \in \mathcal{L}(Y, Z) \). If one of the operators \( T \) or \( S \) is compact, then \( ST \) is compact.

**Proof.** Let \( x_k \in X, k \in \mathbb{N} \), satisfy \( \sup_{k \in \mathbb{N}} \|x_k\| =: c < \infty \).

a) Let \( T, R \in \mathcal{L}_0(X, Y) \) be compact. If \( \alpha \in \mathbb{C} \), then \( \alpha T \) is also compact. There further exists a converging subsequence \( (Tx_{k_j})_j \). Since \( (x_{k_j})_j \) is still bounded, there is another converging subsequence \( (Rx_{k_j})_j \). Hence, \( ((T + R)x_{k_j})_j \) converges and so \( T + R \in \mathcal{L}_0(X, Y) \).

b) Let \( T_n \in \mathcal{L}_0(X, Y) \) converge in \( \mathcal{L}(X, Y) \) to some \( T \in \mathcal{L}(X, Y) \) as \( n \to \infty \). Since \( T_1 \) is compact, there is a converging subsequence \( (T_1x_{\nu_1(j)})_j \). Because \( \|x_{\nu_1(j)}\| \leq c \) for all \( j \), there exists a subsequence \( \nu_2 \subseteq \nu_1 \) such that \( (T_2x_{\nu_2(j)})_j \) converges. Note that \( (T_1x_{\nu_2(j)})_j \) still converges. Iteratively, we obtain subsequences \( \nu_l \subseteq \nu_{l-1} \) such that \( (T_nx_{\nu_l(j)})_j \) converges for all \( l \geq n \).

Set \( u_m = x_{\nu_m(m)} \). Then \( (T_nu_m)_m \) converges as \( m \to \infty \) for each \( n \in \mathbb{N} \). Let \( \varepsilon > 0 \). Fix \( N = N_\varepsilon \in \mathbb{N} \) such that \( \|T_N - T\| \leq \varepsilon \). Let \( M \geq m \geq N \). We then obtain

\[
\|T_{u_M} - T_{u_m}\| \leq \|(T - T_N)u_M\| + \|T_N(u_M - u_m)\| + \|(T_N - T)u_m\| \\
\leq 2\varepsilon c + \|T_N(u_M - u_m)\|.
\]

Therefore \( (T_{u_m}) \) is a Cauchy sequence. So we have shown that \( T \) is compact. Hence, \( \mathcal{L}_0(X, Y) \) is closed.

c) Let \( S \in \mathcal{L}_0(X, Y) \). If \( (Tx_k)_k \) is bounded, then there is a converging subsequence \( (STx_{k_j})_j \), so that \( ST \) is compact.

If \( T \in \mathcal{L}_0(X, Y) \), then there is a converging subsequence \( (Tx_{k_j})_j \) and thus \( STx_{k_j} \) converges. Again, \( ST \) is compact. \( \square \)

**Remark 2.4.** Strong limits of compact operators need not be compact. Consider, e.g., \( X = \ell^2 \) and \( T_n x = (x_1, \ldots, x_n, 0, 0, \ldots) \) for all \( x \in X \) and \( n \in \mathbb{N} \). Then \( T_n \in \mathcal{L}_{00}(X) \subseteq \mathcal{L}_0(X) \) but \( T_n x \to x = 1x \) as \( n \to \infty \) for every \( x \in X \) and \( I \not\in \mathcal{L}_0(X) \).

**Example 2.5.** a) Let \( X \in \{C([0, 1]), L^p([0, 1]), 1 \leq p \leq \infty \}, Y = C([0, 1]), \) and \( k \in C([0, 1]^2) \). Setting

\[
Tf(t) = \int_0^1 k(t, \tau)f(\tau)d\tau,
\]

for \( f \in X \) and \( t \in [0, 1] \), we define the integral operator \( T : X \to Y \) for the kernel \( k \). Since \( \|Tf\| \leq \|k\|\|f\|_1 \leq \|k\|\|f\|_p \) (using that \( \lambda([0, 1]) = 1 \)), we obtain that \( T \in \mathcal{L}(X, Y) \) and thus \( TB_X(0, 1) \) is bounded in \( Y \). Moreover, for \( t, s \in [0, 1] \) and \( f \in B_X(0, 1) \) we have

\[
|Tf(t) - Tf(s)| \leq \int_0^1 |k(t, \tau) - k(s, \tau)||f(\tau)|d\tau.
\]
\subsection*{2.1. Compact Operators}

\begin{align*}
\leq & \sup_{\tau \in [0,1]} |k(t, \tau) - k(s, \tau)| \|f\|_1 \\
\leq & \sup_{\tau \in [0,1]} |k(t, \tau) - k(s, \tau)| \|f\|_p.
\end{align*}

The right hand side tends to 0 as \(|t - s| \to 0\) uniformly in \(f \in \overline{B}(0,1)\), because \(k\) is uniformly continuous. Therefore \(T\overline{B}\) is equicontinuous. The Arzela-Ascoli theorem (see e.g. Theorem 1.45 in [FA]) then implies that \(T\overline{B}\_X(0,1)\) is relatively compact. Hence, \(T \in \mathcal{L}_0(X,Y)\).

b) Let \(Z = L^2([0,1])\) and \(k \in L^2([0,1]^2)\). Define \(T \in \mathcal{L}(Z)\) as in a), cf. Example 2.23 in [FA]. There are \(k_n \in C([0,1]^2)\) converging to \(k \in L^2([0,1]^2)\) (cf. Analysis 3). Let \(T_n\) be the corresponding integral operators in \(\mathcal{L}(Z)\) and \(\mathcal{L}(Z, C([0,1]))\) as in a). Hölder’s inequality yields

\[|Tf(t) - Tnf(t)| \leq \int_0^1 |k(t, \tau) - k_n(t, \tau)| |f(\tau)| d\tau \leq \left( \int_0^1 |k(t, \tau) - k_n(t, \tau)|^2 d\tau \right)^{\frac{1}{2}} \|f\|_2\]

for all \(n \in \mathbb{N}, f \in Z\), and \(t \in [0,1]\). Hence \(|Tf - Tnf|_2 \leq \|k - k_n\|_2 \|f\|_2\) and so \(T_n \to T\) as \(n \to \infty\) in \(\mathcal{L}(Z)\). Let \((f_k) \subseteq Z\) be bounded and \(n \in \mathbb{N}\). Due to a), there is a subsequence \((T_n f_{k_j})\) converging for \(\|\cdot\|_\infty\), and hence for \(\|\cdot\|_2\) (since \(\lambda([0,1]) < \infty\)). So, \(T_n \in \mathcal{L}_0(Z)\) and the compactness of \(T\) then follows from Proposition 2.3.

\begin{proposition}
Let \(T \in \mathcal{L}_0(X,Y)\) and \(x_n \overset{\sigma}{\to} x\) in \(X\) as \(n \to \infty\) (i.e., \(\langle x_n, x^\ast \rangle \to \langle x, x^\ast \rangle\) for every \(x^\ast \in X^\ast\)). Then \(Tx_n\) converges to \(Tx\) in \(Y\) as \(n \to \infty\).
\end{proposition}

\textbf{Proof.} We have \(\langle Tx_n - Tx, y^\ast \rangle = \langle x_n - x, T^\ast y^\ast \rangle \to 0\) as \(n \to \infty\) for each \(y^\ast \in Y^\ast\), hence \(Tx_n \overset{\sigma}{\to} Tx\) as \(n \to \infty\).

Suppose that \(Tx_n\) does not converge to \(Tx\). Then there are \(\delta > 0\) and a subsequence such that

\[\|Tx_{n_j} - Tx\| \geq \delta > 0\]

for all \(j \in \mathbb{N}\). By compactness, there is a subsequence and a \(y \in Y\) such that \(Tx_{n_{j_l}} \to y\) as \(l \to \infty\). On the other hand, \(Tx_{n_{j_l}} \overset{\sigma}{\to} Tx\) and \(Tx_{n_{j_l}} \overset{\sigma}{\to} y\). Since weak limits are unique, it follows that \(y = Tx\), which is impossible.

\begin{theorem}[Schauder]
An operator \(T \in \mathcal{L}(X,Y)\) is compact if and only if \(T^\ast \in \mathcal{L}(Y^\ast, X^\ast)\) is compact.
\end{theorem}

\textbf{Proof.} 1) Let \(T\) be compact and \(y_{n}^\ast \in Y^\ast, n \in \mathbb{N}\), with \(\sup_{n \in \mathbb{N}} \|y_{n}^\ast\| =: \epsilon < \infty\). The set \(K := \overline{T\mathcal{B}}_X(0,1)\) is a compact metric space for the restriction of the norm of \(Y\). Set \(f_n := y_{n}^\ast|K \in C(K)\) for each \(n \in \mathbb{N}\). Putting \(c_1 := \max_{y \in K} \|y\| < \infty\), we obtain

\[\|f_n\|_\infty = \max_{y \in K} \|\langle y, y_{n}^\ast\rangle\| \leq cc_1\]
for every \( n \in \mathbb{N} \). Moreover, \((f_n)_{n \in \mathbb{N}}\) is equicontinuous since
\[
|f_n(y) - f_n(z)| \leq \|y_n\| \|y - z\| \leq c \|y - z\|
\]
for all \( n \in \mathbb{N} \) and \( y, z \in K \).

The Arzela-Ascoli theorem then yields a subsequence \((f_{n_j})_j\) converging in \(C(K)\). We thus deduce that
\[
\|T^*y_{n_j} - T^*y_{n_l}\|_{X^*} = \sup_{\|x\| \leq 1} |\langle x, T^*(y_{n_j} - y_{n_l}) \rangle| = \|f_{n_j} - f_{n_l}\|_{C(K)}
\]
tends to 0 as \( j, l \to \infty \) (using that \(T \overline{B}(0, 1)\) is dense in \(K\)). This means that \((T^*y_{n_j})_j\) converges and so \(T^*\) is compact.

2) Let \(T^*\) be compact. By Step 1), the biadjoint \(T^{**}\) is compact. Let \(J_X : X \to X^{**}\) be the canonical isometric embedding. Due to e.g. Proposition 4.39 in [FA], it holds that
\[
T^{**} \circ J_X = J_Y \circ T : X \to Y^{**}.
\]
Using also Proposition 2.3 we see that \(J_Y \circ T\) is compact. If \((x_n)\) is bounded, we thus obtain a converging subsequence \((J_Y T x_{n_j})_j\) which is Cauchy. Since \(J_Y\) is isometric, also \((T x_{n_j})_j\) is Cauchy and thus converges; i.e., \(T\) is compact. \(\square\)

### 2.2 The Fredholm alternative

We need some fact from functional analysis. For nonempty sets \(M \subseteq X\) and \(N_* \subseteq X^*\) we define the annihilators
\[
M^\perp = \{x^* \in X^* : \forall y \in M \ \langle y, x^* \rangle = 0\},
\]
\[
\perp N_* = \{x \in X : \forall y^* \in N_* \ \langle x, y^* \rangle = 0\}.
\]
These sets are equal to \(X^*\) or \(X\) if and only if \(M = \{0\}\) or \(N_* = \{0\}\), respectively, see e.g. Remark 4.18 in [FA]. For \(T \in \mathcal{L}(X)\) it holds
\[
(2.1) \quad R(T)^\perp = N(T^*), \quad \overline{R(T)} = \perp N(T^*), \quad N(T) = \perp R(T^*), \quad \overline{R(T^*)} \subseteq N(T)^\perp.
\]

In particular, \(R(T)\) is dense if and only if \(T^*\) is injective; and if \(R(T^*)\) is dense, then \(T\) is injective. (See e.g. Proposition 4.41 in [FA].)

The following theorem extends fundamental results on matrices known from Linear Algebra.

**Theorem 2.8** (Riesz 1918, Schauder 1930). Let \(K \in \mathcal{L}_0(X)\) and \(T = I - K\). Then the following assertions hold.

a) \(R(T)\) is closed.
2.2. THE FREDHOLM ALTERNATIVE

b) $\dim N(T) < \infty$ and $\text{codim } R(T) := \dim X/R(T) < \infty$.

c) $T$ is bijective $\iff$ $T$ is surjective $\iff$ $T$ is injective $\iff$ $T^*$ is bijective $\iff$ $T^*$ is injective.

More precisely, we have

\begin{equation}
\dim N(T) = \text{codim } R(T) = \dim N(T^*) = \text{codim } R(T^*).
\end{equation}

**Corollary 2.9** (Fredholm alternative). Let $L \in \mathcal{L}_0(X)$, $\lambda \in \mathbb{C}\setminus\{0\}$, and $y \in X$. Then one of the following alternatives hold:

A) $\lambda x = Lx$ has only the trivial solution $x = 0$.

Then for every $y \in X$ there is a unique solution $x \in X$ of $\lambda x - Lx = y$ given by $x = R(\lambda, L)y$.

B) $\lambda x = Lx$ has an $n$-dimensional solution space $N(\lambda I - L)$ for some $n \in \mathbb{N}$.

Then there are $n$ linearly independent solutions $x_1^*, \ldots, x_n^* \in X^*$ of $\lambda x^* = Lx^*$, and the equation $\lambda x - Lx = y$ has a solution $x \in X$ if and only if $\langle y, x_k^* \rangle = 0$ for every $k = 1, \ldots, n$.

Finally, every $z \in X$ satisfying $\lambda z - Lz = y$ is of the form $z = x + x_0$, where $\lambda x - Lx = y$ and $x_0 \in N(\lambda I - L)$.

**Proof of Corollary 2.9.** Apply Theorem 2.8 to $K := \frac{1}{\lambda} L \in \mathcal{L}_0(X)$ and note that $\lambda x - Lx = y$ is equivalent to $(I - K)x = \frac{1}{\lambda} y$. By Theorem 2.8b), either $\dim N(I - K) = 0$ (case A) or $\dim N(I - K) = n \in \mathbb{N}$ (case B).

In the first case, $I - K$ is bijective due to Theorem 2.8c) which yields part A.

In the second case, Theorem 2.8a) shows that $R(I - K)$ is closed so that $R(I - K) = \frac{1}{\lambda} N(I - K^*)$, by equation (2.1). This gives the solvability condition from case B noting that $\dim N(I - K^*) = n$ due to equation (2.2). If $x - Kx = y$ and $z - Kz = y$, then $z - x$ belongs to $N(I - K^*)$, as required in case B.

**Example 2.10.** Let $X = C([0, 1])$ and $Kf(t) = \int_0^t f(s)ds$ for $t \in [0, 1]$ and $f \in X$. Then

\[ R(K) = \{ g \in C^1([0, 1]) : g(0) = 0 \}, \]

which is not closed in $X$. Moreover, $N(K) = \{0\}$ by Example 1.17. Also, $K$ is compact since $K \overline{B}_X(0, 1)$ is contained in the closed unit ball of $C^1([0, 1])$, which is a compact subset of $X$ due to the Arzela-Ascoli theorem. In this case, $Kf = g$ cannot be solved for all $g \in X$. Summing up, the Fredholm alternative fails for $\lambda = 0$.

**Proof of Theorem 2.8.** 1) The space $N := N(T)$ is closed in $X$. For $x \in N$ we have $Kx = x \in N$, so that $K$ leaves $N$ invariant and its restriction $K_N$ to $N$ coincides with the identity on $N$. On the other hand, $K_N$ is still compact so that $\dim N < \infty$ by Remark 2.2d).

2) Since $\dim N < \infty$, there is a closed subspace $C \subseteq X$ such that $N \cap C = \{0\}$ and $N + C = X$, i.e., $X = N \oplus C$, see e.g. Proposition 4.15 in [FA].
CHAPTER 2. SPECTRAL THEORY OF COMPACT OPERATORS

Let \( \hat{T} : C \to R(T) \) be the restriction of \( T \) to \( C \) and endow \( C \) and \( R(T) \) with the norm of \( X \). (Observe that \( C \) is a Banach space, but we do not know yet whether \( R(T) \) is a Banach space.)

If \( \hat{T}x = 0 \) for some \( x \in C \), then \( x \in N \) and so \( x = 0 \). Let \( y \in R(T) \). Then there is an \( x \in X \) such that \( Tx = y \). We can write \( x = x_0 + x_1 \) with \( x_0 \in N \) and \( x_1 \in C \).

Hence, \( \hat{T}x_1 = T x_1 + T x_0 = y \), and so \( \hat{T} \) is bijective.

A corollary to the open mapping theorem (see e.g. Corollary 3.19 in [FA]) yields that \( R(T) = R(\hat{T}) \) is closed if and only if \( T^{-1} : R(T) \to C \) is bounded. Assume that \( \hat{T}^{-1} \) were unbounded. Then there would exist \( y_n = \hat{T}x_n \in R(T) \) with \( x_n \in C \) such that \( y_n \to 0 \) as \( n \to \infty \) and \( \|x_n\| = \|\hat{T}^{-1}y_n\| \geq \delta \) for some \( \delta > 0 \) and all \( n \in \mathbb{N} \). If there exists an unbounded subsequence \( (x_{n_j}) \), then we set \( \hat{x}_{n_j} := \frac{1}{\|x_{n_j}\|} x_{n_j} \) and note that \( \|\hat{x}_{n_j}\| = 1 \) and that

\[
\hat{y}_{n_j} := \hat{T}\hat{x}_{n_j} = \frac{1}{\|x_{n_j}\|} y_{n_j}
\]

still converges to 0 as \( l \to \infty \). So we can assume that \( (x_n) \) is bounded.

Since \( K \) is compact, there exists a subsequence \( (x_{n_j}) \) and a \( z \in X \) such that \( Kx_{n_j} \to z \) as \( j \to \infty \). Consequently, \( x_{n_j} = y_{n_j} + Kx_{n_j} \to z \) as \( j \to \infty \) and so \( \|z\| \geq \delta > 0 \). We further deduce that \( z \in C \) from the closedness of \( C \) and that \( z \in N \) from

\[
Tz = z - Kz = \lim_{j \to \infty} Kx_{n_j} - K \lim_{j \to \infty} x_{n_j} = 0.
\]

Hence, \( z \in C \cap N = \{0\} \), which contradicts \( \|z\| \geq \delta > 0 \). So, assertion a) has been shown.

3) Theorem 2.7 shows that \( K^* \) is also compact so that \( \dim N(I - K^*) < \infty \) by the first step. Using equation (2.1) and Proposition 4.20 in [FA], we further obtain

\[
N(T^*) = R(T)^\perp \cong (X/R(T))^*.
\]

Since \( N(T^*) \) is finite-dimensional, linear algebra yields that

\[
(*) \quad \dim N(T^*) = \dim (X/R(T))^* = \dim X/R(T) = \text{codim } R(T)
\]

and so \( \text{codim } R(T) < \infty \), i.e., assertion b) is true.

4) Claim A: There is a linear subspace \( \hat{N} \) with \( \dim \hat{N} < \infty \) and a closed linear subspace \( \hat{R} \) of \( X \) such that

\[
X = \hat{N} \oplus \hat{R}, \quad T \hat{N} \subseteq \hat{N}, \quad T \hat{R} \subseteq \hat{R} \quad \text{and} \quad T_2 := T_{|\hat{R}} : \hat{R} \to \hat{R} \text{ is bijective.}
\]

Suppose for a moment that Claim A is true. Setting \( T_1 := T_{|\hat{N}} \in \mathcal{L}(\hat{N}) \), we have the following facts.

(i) \( \dim \hat{N}/R(T_1) = \dim N(T_1) \) (by the dimension formula in \( \mathbb{C}^n \)).

(ii) Writing \( x = x_1 + x_2 \) for \( x \in X \), \( x_1 \in \hat{N} \) and \( x_2 \in \hat{R} \) we deduce that \( Tx = 0 \) if and only if \( T_2x_2 = -T_1x_1 \in \hat{N} \cap \hat{R} = \{0\} \) which is, by the bijectivity of \( T_2 \), equivalent to \( x_2 = 0 \) and \( T_1x_1 = 0 \). Hence, \( x \in N(T) \) if and only if \( x \in \hat{N} \) and \( T_1x = 0 \), which means that \( N(T) = N(T_1) \).
We define the map
\[ \Phi : \hat{N}/R(T_1) \to X/R(T), \quad x + R(T_1) \mapsto x + R(T), \]
for \( x \in \hat{N} \subseteq X \). Because of \( R(T_1) \subseteq R(T) \), the map \( \Phi \) is well defined, and it is linear. We want to show that \( \Phi \) is bijective.

If \( \Phi(x + R(T_1)) = 0 \) for some \( x \in \hat{N} \), then there is a \( y \in X \) such that \( x = Ty \). There are \( y_1 \in \hat{N} \) and \( y_2 \in \hat{R} \) with \( y = y_1 + y_2 \). Hence, \( T_2y_2 = x - T_1y_1 \) belongs to \( \hat{R} \cap \hat{N} = \{0\} \). Since \( T_2 \) is injective, we infer that \( y_2 = 0 \). Thus, \( \Phi \) is injective.

Take \( x \in X \). There are \( x_1 \in \hat{N} \) and \( x_2 \in \hat{R} = T\hat{R} \) with \( x = x_1 + x_2 \). We now conclude that \( x = x_2 - x_1 \in R(T) \) and so
\[ \Phi(x + R(T_1)) = x_1 + R(T) = x + R(T). \]

Hence, \( \Phi \) is bijective, which leads to \( \dim X/R(T) = \dim \hat{N}/R(T_1) \).

The facts (i)-(iii) imply that
\[ (**) \quad \dim N(T) = \dim N(T_1) = \dim \hat{N}/R(T_1) = \dim X/R(T) = \text{codim} R(T). \]
Since also \( K^* \) is compact by Theorem 2.7, we further obtain
\[ (***) \quad \dim N(T^*) = \text{codim} R(T^*). \]
Combining \((*)-(***)\), we arrive at assertion c).

5) **Proof of Claim A.** We set \( N_k = N(T^k) \) and \( R_k = R(T^k) \) for \( k \in \mathbb{N}_0 \). It holds
\[ \{0\} = N_0 \subseteq N_1 \subseteq N_2 \subseteq \ldots, \quad X = R_0 \supseteq R_1 \supseteq R_2 \supseteq \ldots, \]
\[ (+) \quad TN_k \subseteq N_{k-1} \subseteq N_k, \quad \text{and} \quad TR_k = R_{k+1} \subseteq R_k, \]
for all \( k \in \mathbb{N}_0 \). We further have
\[ T^k = (I - K)^k = I - \sum_{j=1}^{k} \binom{k}{j} (-1)^{j+1} K^j =: I - C_k, \]
where \( C_k \) is compact for each \( k \in \mathbb{N} \), due to Proposition 2.3.

Assertions a) and b) thus show that \( R_k \) is closed, \( \dim N_k < \infty \), and \( \text{codim} R_k < \infty \) for every \( k \in \mathbb{N} \). We need four more claims to establish Claim A.

**Claim 1:** There is a minimal \( n \in \mathbb{N}_0 \) such that \( N_n = N_{n+j} \) for all \( j \in \mathbb{N}_0 \).

Indeed, if it were true that \( N_j \nsubseteq N_{j+1} \) for all \( j \in \mathbb{N}_0 \), then Riesz’ Lemma (see e.g. Lemma 1.42 in [FA]) would give \( x_j \in N_j \) with \( \|x_j\| = 1 \) and \( d(x_j, N_{j-1}) \geq 1/2 \) for every \( j \in \mathbb{N}_0 \). Take \( l > k \geq 0 \). Since \( Tx_l + x_k - Tx_k \in N_{l-1} \), we deduce that
\[ \|Kx_l - Kx_k\| = \|x_l - (Tx_l + x_k - Tx_k)\| \geq 1/2. \]
As a result, \((Kx_k)_k\) has no converging subsequence, which contradicts the compactness of \( K \). So there is a minimal \( n \in \mathbb{N}_0 \) with \( N_n = N_{n+1} \). Let \( x \in N_{n+2} \). Then,
\( Tx \in N_{n+1} = N_n \) so that \( x \in N_{n+1} \). This means that \( N_{n+1} = N_{n+2} \), and Claim 1 follows by induction.

**Claim 2**: There is a minimal \( m \in \mathbb{N}_0 \) such that \( R_m = R_{m+j} \) for all \( j \in \mathbb{N}_0 \).

This claim can be shown as Claim 1, see e.g. Lemma VI.2.2 in [Wer05].

**Claim 3**: \( N_n \cap R_n = \{0\} \) and \( N_n + R_n = X \).

Indeed, for the first part, let \( x \in N_n \cap R_n \). Then \( T^nx = 0 \) and there is a \( y \in X \) such that \( T^ny = x \). Hence, \( T^{2n}y = 0 \) and so \( y \in N_{2n} = N_n \) by Claim 1. Consequently, \( x = T^n y = 0 \).

For the second part, let \( x \in X \). By Claim 2 we have \( T^mx \in R_m = R_{2m} \), and thus there exists a \( y \in X \) with \( T^mx = T^{2m}y \). Therefore \( x = (x - T^ny) + T^ny \in N_m + R_m \).

**Claim 4**: It holds \( n = m \).

Indeed, suppose that \( n > m \). Due to Claim 1 and Claim 2, there is an \( x \in N_n \setminus N_m \) and it holds \( R_n = R_m \). Claim 3 further gives \( y \in N_m \subseteq N_n \) and \( z \in R_m = R_n \) such that \( x = y + z \). Therefore, \( z = x - y \in N_n \) so that \( z = 0 \) by Claim 3. As a result, \( x = y \in N_m \), which is impossible. The inequality \( n < m \) can be excluded in a similar way, see e.g. Lemma VI.2.2 in [Wer05].

We can now finish the proof of Claim A, setting \( \tilde{N} := N_n \) and \( \tilde{R} = R_n \). The closedness of \( N_n \) and \( R_n \) have been established before Claim 1. From Claim 3 and 4 we thus deduce that \( X = \tilde{N} \oplus \tilde{R} \). Moreover, equation ( + ) yields \( T\tilde{N} \subseteq \tilde{N} \) and \( T\tilde{R} = \tilde{R} \).

If \( Tx = 0 \) for some \( x = T^ny \in \tilde{R} \) and \( y \in X \), then \( y \in N_{n+1} = N_n \) by Claim 1. Therefore, \( x = 0 \) and \( T|\tilde{R} \) is bijective and thus isomorphic by the open mapping theorem.

We now reformulate the Riesz-Schauder theorem in terms of spectral theory. Observe that the Voltera operator in Example 2.10 is compact and has the spectrum \( \sigma(V) = \{0\} \) with \( \sigma_p(V) = \emptyset \).

**Theorem 2.11.** Let \( \dim X = \infty \) and \( K \in \mathcal{L}(X) \) be compact. Then the following assertions hold.

a) \[ \sigma(K) = \{0\} \cup \{\lambda_j : j \in J\}, \]

where either \( J = \emptyset \), or \( J = \mathbb{N} \), or \( J = \{1, \ldots, n\} \) for some \( n \in \mathbb{N} \).

b) Each \( \lambda \in \sigma(K) \setminus \{0\} \) is an eigenvalue of \( K \) and

\[ \dim N(\lambda I - K) = \text{codim}(\lambda I - K)X < \infty. \]

c) If \( J = \mathbb{N} \), then \( \lambda_j \to 0 \) as \( j \to \infty \). (This means that for all \( \varepsilon > 0 \) the set \( \sigma(K) \setminus B(0, \varepsilon) \) is finite.)
2.2. THE FREDHOLM ALTERNATIVE

Proof. If $0 \in \rho(K)$, then $K$ would be invertible. Proposition 2.3 now shows that $I = K^{-1}K$ would be compact which contradicts $\dim X = \infty$.

Hence, $0$ always belongs to $\sigma(K)$. Observe that thus assertion a) follows from c) by taking $\varepsilon = \frac{1}{n}$, $n \in \mathbb{N}$.

For $\lambda \in \mathbb{C}\setminus\{0\}$ we have $\lambda I - K = \lambda(I - \frac{1}{\lambda}K)$. Since $\frac{1}{\lambda}K \in \mathcal{L}_0(X)$, Theorem 2.8 implies either $\lambda \in \rho(K)$ or $\lambda \in \sigma_p(K)$ with
\[
\dim N(\lambda I - K) = \dim N \left( I - \frac{1}{\lambda}K \right) = \text{codim} \left( I - \frac{1}{\lambda}K \right) X < \infty.
\]
So we have established part b).

To prove assertion c), we suppose that for some $\varepsilon > 0$ we have $\lambda_n \in \sigma(K) \setminus B(0, \varepsilon)$ with $\lambda_n \neq \lambda_m$ for all $n \neq m$ in $\mathbb{N}$ and $x_n \in X \setminus\{0\}$ with $Kx_n = \lambda_n x_n$. In Linear Algebra it is shown that eigenvectors to different eigenvalues are linearly independent. Hence, the subspaces
\[
X_n := \text{lin}\{x_1, \ldots, x_n\}
\]
satisfy $X_n \subseteq X_{n+1}$ for every $n \in \mathbb{N}$. Moreover, $K X_n \subseteq X_n$ and $X_n$ is closed for each $n \in \mathbb{N}$ (since $\dim X_n < \infty$). Riesz’ lemma (see e.g. Lemma 1.42 in [FA]) gives $y_n \in X_n$ such that $\|y_n\| = 1$ and $d(y_n, X_{n-1}) \geq \frac{\varepsilon}{2}$ for each $n \in \mathbb{N}$. There are $\alpha_{n,j} \in \mathbb{C}$ with $y_n = \alpha_{n,1}x_1 + \cdots + \alpha_{n,n}x_n$, and hence the vector
\[
\lambda_n y_n - Ky_n = \sum_{j=1}^{n} (\lambda_n - \lambda_j) \alpha_{n,j} x_j = \sum_{j=1}^{n-1} (\lambda_n - \lambda_j) \alpha_{n,j} x_j
\]
belongs to $X_{n-1}$. For $n > m$ we thus obtain
\[
\|Ky_n - Ky_m\| = |\lambda_n - \lambda_m| \|y_n - \frac{1}{\lambda_n} (\lambda_n y_n - Ky_n + Ky_m)\| \geq \frac{|\lambda_n|}{2} \geq \frac{\varepsilon}{2},
\]
since $\lambda_n y_n - Ky_n + Ky_m \in X_{n-1}$. This fact contradicts the compactness of $K$. \qed

Example 2.12 (The Dirichlet problem and boundary integrals). Let $D \subseteq \mathbb{R}^3$ be open and bounded with $\partial D \in C^2$ (see Analysis 3). Let $\varphi \in C(\partial D)$ be given. We look for $u \in C^2(D) \cap C(\overline{D})$ satisfying

\[
\begin{cases}
\Delta u(x) = 0, & x \in D, \\
u(x) = \varphi(x), & x \in \partial D.
\end{cases}
\]

In the lecture Partial Differential Equations (see e.g. Corollary 2.1.1 in [Jos08]) it is shown that the problem (2.3) has at most one solution. To find a solution, we consider the “Newton potential” in $\mathbb{R}^3$ given by $\gamma(x) = \frac{1}{4\pi|x|^2}$ for $x \in \mathbb{R}^3 \setminus\{0\}$. Let $\nu(y)$ be the outer unit normal at $y \in \partial D$. We define
\[
k(x, y) = \frac{\partial}{\partial \nu(y)} \gamma(x - y) = - (\nabla \gamma)(x - y) \cdot \nu(y) = \frac{(x - y) \cdot \nu(y)}{4\pi|x - y|^3}
\]
for all $x \in \mathbb{R}^3$ and $y \in \partial D$ with $x \neq y$. One defines the “double layer potential” by
\[
Sg(x) = \int_{\partial D} k(x, y) g(y) d\sigma(y)
\]
for all $x \in \mathbb{R}^3 \setminus \partial D$ and $g \in C(\partial D)$. Using results from Analysis 2 or 3, one derives that $Sg \in C^\infty(\mathbb{R}^3 \setminus \partial D)$ and $\Delta Sg = 0$ on $\mathbb{R}^3 \setminus \partial D$ employing that $\Delta \gamma = 0$ on $\mathbb{R}^3 \setminus \{0\}$.

For each $\varphi \in C(\partial D)$ one thus obtains the solution $S\varphi \in \mathbb{C}$ of (2.3) if one can find a $g \in C(\partial D)$ such that

$$\lim_{x \to z} Sg(x) = \varphi(z)$$

for all $z \in \partial D$.

To that purpose we recall from Analysis 3 that the surface integral for a measurable function $h : \partial D \to \mathbb{C}$ is given by

$$\int_{\partial D} h(y) d\sigma(y) = \sum_{j=1}^{m} \int_{U_j} \varphi_j(F_j(t)) h(F_j(t)) \sqrt{\det F'(t)^T \cdot F'(t)} dt,$$

if the right hand side exists. Here, $0 \leq \varphi_j \in C^\infty_c(\mathbb{R}^3)$ satisfy $\sum_{j=1}^{m} \varphi_j = 1$, $\{\tilde{V}_1, \ldots, \tilde{V}_m\}$ is an open cover of $\partial D$ in $\mathbb{R}^3$, $V_j = \tilde{V}_j \cap \partial D$, $\Psi_j : \tilde{V}_j \to U_j$ is a $C^2$-diffeomorphism such that $\partial_k \Psi_j$, $\partial_k \Psi_j$, $\partial_k \Psi_j^{-1}$ and $\partial_k \Psi_j^{-1}$ have continuous extensions to $\partial \tilde{V}_j$ and $\partial \tilde{U}_j$, respectively, and $F_j = \Psi_j^{-1} |_{U_j}$ has the range $V_j$, where $U_j = \tilde{U}_j \cap (\mathbb{R}^2 \times \{0\})$. We identify $U_j$ with a subset of $\mathbb{R}^2$ writing $t \in \mathbb{R}^2$ instead of $(t, 0) \in \mathbb{R}^3$. In the following we omit the index $j \in \{1, \ldots, m\}$.

Recall that for $y = F(t) \in V$, the tangent plane of $\partial D$ at $y$ is spanned by $\partial_1 F(t)$ and $\partial_2 F(t)$, where $t \in U \subseteq \mathbb{R}^2$. Taylor’s formula applied to $\Psi^{-1} \in C^2_b(\tilde{U})$ yields, for $x = \Psi^{-1}(s, 0)$ with $s \in U \subseteq \mathbb{R}^2$, that

$$x = y + (\Psi^{-1})'(t, 0) \begin{pmatrix} s-t \n 0 \end{pmatrix} + O(|s-t|^2) = y + F'(t)(s-t) + O(|s-t|^2).$$

Using that $\nu(y)$ is orthogonal to $\partial_2 F(t)$, we deduce that

$$(x - y) \cdot \nu(y) = \nu(y)^T F'(t)(s-t) + O(|s-t|^2) = O(|s-t|^2).$$

On the other hand, $\Psi^{-1}$ and $\Psi$ are globally Lipschitz so that

$$c |s-t|_2 \leq |x-y|_2 \leq C |s-t|_2$$

for all $x = F(s)$, $y = F(t) \in V$ with $s, t \in U$ and some constants $C, c > 0$. In the following we denote by $c$ various, possibly differing constants.

The above facts can be established for all $j = 1, \ldots, m$, and so we obtain

$$|k(x, y)| \leq \frac{c}{|x-y|_2} \leq \frac{c}{|s-t|_2}$$

for all $x = F(s)$ and $y = F(t)$ in $\partial D$ with $x \neq y$. As a result, the integrands

$$\varphi(F(t)) k(F(s), F(t)) g(F(t)) \sqrt{\det F'(t)^T \cdot F'(t)}$$

of $Sg$ are bounded by a constant times $|s-t|^{-1}_2 \|g\|_\infty$ for all $x = F(s) \subseteq \partial D$ and $y = F(t) \in \partial D$ with $x \neq y$. We now define $k(x, x) = 0$ for $x \in \partial D$ and

$$k_n(x, y) = \begin{cases} k(x, y), & |x-y|_2 > 1/n, \\
n^3(4\pi)^{-1} (x-y) \cdot \nu(y), & |x-y|_2 \leq 1/n, \end{cases}$$

for all $x, y \in \mathbb{R}^3 \setminus \partial D$. Using results from Analysis 2 or 3, one derives that $Sg \in C^\infty(\mathbb{R}^3 \setminus \partial D)$ and $\Delta Sg = 0$ on $\mathbb{R}^3 \setminus \partial D$ employing that $\Delta \gamma = 0$ on $\mathbb{R}^3 \setminus \{0\}$.
2.2. THE FREDHOLM ALTERNATIVE

for \( n \in \mathbb{N} \). Observe that \( |k_n(x, y)| \leq cn^3|s-t|^3 \leq c|s-t|^{-1} \) if \( |x-y| \leq 1/n \) because then \( |s-t| \leq c/n \).

Since \( k_n \) is continuous on \( \partial D \times \partial D \), we see that

\[
T_n g(x) = \int_{\partial D} k_n(x, y)g(y)d\sigma,
\]

for \( x \in \partial D \) and \( g \in C(\partial D) \), defines an operator \( T_n \in L(C(\partial D)) \). As in Example 2.5a), one also shows that \( T_n \) is compact by means of the Arzela-Ascoli theorem. For \( g \in C(\partial D) \) and \( x \in \partial D \) we further estimate

\[
\int_{\partial D} |(k(x, y) - k_n(x, y))g(y)|d\sigma(y) = \int_{\partial D \cap B(x, \frac{1}{n})} |k(x, y) - k_n(x, y)||g(y)|d\sigma(y)
\leq \sum_{j=1}^n c \int_{U_j \cap B(s, \frac{1}{n})} \frac{1}{|s-t|^2} dt
\leq c\|g\|_\infty \int_{B(0, \frac{1}{n})} \frac{dr}{r^2}
\leq c\|g\|_\infty \int_0^{\pi/2} \frac{r}{n} dr = \frac{c\|g\|_\infty}{n},
\]

using polar coordinates in \( \mathbb{R}^2 \).

Hence, for each \( x \in \partial D \) the function \( y \mapsto k(x, y)g(y) \) is integrable for the surface measure \( d\sigma \) on \( \partial D \). Moreover, \( T_n g \) converges uniformly for \( x \in \partial D \) to the function

\[
T g(x) := \int_{\partial D} k(x, y)g(y)d\sigma(y),
\]

as \( n \to \infty \), for every \( g \in C(\partial D) \). This fact yields that \( T g \in C(\partial D) \). Clearly \( T : C(\partial D) \to C(\partial D) \) is linear.

We have even shown that \( T_n \) converges to \( T \) in operator norm so that \( T \) is bounded and, by Proposition 2.3, compact. One can now prove that, for all \( z \in \partial D \) and \( g \in C(\partial D) \), it holds

\[
(2.4) \quad \lim_{x \to z, x \in \partial D} S g(x) = T g(z) - \frac{1}{2} g(z), \quad \text{and} \quad \lim_{x \to z, x \in \mathbb{R}^3 \setminus \overline{D}} S g(x) = T g(z) + \frac{1}{2} g(z),
\]

see e.g. Theorem IX in Chapter VI of [Kel67]. As a result, \( S g \in C^2(D) \cap C(\overline{D}) \) solves (2.3) if and only if there is a \( g \in C(\partial D) \) such that \( \frac{1}{2} g - T g = -\varphi \).

In view of the Fredholm alternative we only have to establish that \( \frac{1}{2} I - T \) is injective. So let \( g \in C(\partial D) \) satisfy \( \frac{1}{2} g_0 = T g_0 \). We have already seen that then \( S g_0 \) solves (2.3) with \( \varphi = 0 \). This problem is also solved by \( u = 0 \). The uniqueness of (2.3) now yields that \( S g_0 = 0 \) on \( D \).

Equation (2.4) and \( T g_0 = \frac{1}{2} g_0 \) show that the function

\[
v(x) = \begin{cases} S g_0(x), & x \in \mathbb{R}^3 \setminus \overline{D}, \\ g_0(x), & x \in \partial D, \end{cases}
\]

27
is continuous on \( \mathbb{R}^3 \setminus D \). Fix \( r_0 > 0 \) such that \( \overline{D} \subseteq B(0, r_0) =: B(r_0) \). For \( r \geq r_0 + 1 \) we can estimate

\[
|k(x, y)| \leq \frac{c}{|x - y|^2} \leq \frac{c}{(r - r_0)^2} \leq \frac{c}{r^2},
\]

\[
|Sg_0(x)| \leq \int_{\partial D} \frac{c}{r^2} \|g_0\|_{\infty} d\sigma \leq \frac{c}{r^2},
\]

for all \( x \in \partial B(r) \) and \( y \in \partial D \). Assume that \( g_0 \neq 0 \). Then we can choose \( r \geq r_0 + 1 \) such that

\[
\|g_0\|_{\infty} > \max_{x \in \partial B(r)} |Sg_0(x)|.
\]

In particular, \( v \) is not constant on \( \overline{B(r)} \setminus D \). Since \( \Delta v = 0 \) on \( B(r) \setminus \overline{D} \), the strong maximum principle says that \( v \) attains no maximum in \( B(r) \setminus \overline{D} \) (see e.g. Theorem 2.1.2 in [Jos08]). As the maximum exists on \( B(r) \setminus D \) it must be attained at a \( y_0 \in \partial D \) and \( v(y_0) > v(x) \) for all \( x \in B(r) \setminus \overline{D} \). Moreover, \( \partial_v v(y_0) \) exists and is equal to 0 since \( Sg_0 = 0 \) on \( D \), see e.g. Theorem X in Chapter VI of [Kel67]. These facts contradict Hopf’s Lemma (see e.g. Lemma 2.1.2 in [Jos08]) so that \( g_0 = 0 \). As a result, \( \frac{1}{2} I - T \) is invertible on \( C(\partial D) \) and

\[
u = \left( -S \left( \frac{1}{2} I - T \right)^{-1} \varphi \right)_{|\overline{D}} \]

solves (2.3). Summing up, we have found a unique \( u \in C^2(D) \cap C(\overline{D}) \) solving the Dirichlet problem (2.3).
Chapter 3

Sobolev spaces and weak derivatives

Throughout, $U \subseteq \mathbb{R}^d$ is open and nonempty.

3.1 Basic properties

We are looking for properties of $C^1$ function and their derivatives which can be generalized to a concept of derivatives suited to $L^p$ spaces, which is in particular not based on pointwise limits. To that purpose, take $f \in C^1(U)$ and $\varphi \in C_\infty(U)$. There is an open set $V$ in $\mathbb{R}^d$ such that $\text{supp}\varphi \subseteq V \subseteq V' \subseteq U$.

Integrating by parts, we deduce

$$
\int_U (\partial_1 f) \varphi dx = \int_{y: V_y \neq \emptyset} \int_{V_y} \partial_1 f(x_1, y) \varphi(x_1, y) dx_1 dy \\
= \int_{y: V_y \neq \emptyset} \left( - \int_{V_y} f(x_1, y) \partial_1 \varphi(x_1, y) dx_1 + [f(x_1, y) \varphi(x_1, y)]|_{x_1 \in \partial V_y} \right) dy \\
= - \int_{y: V_y \neq \emptyset} \int_{V_y} f(x_1, y) \partial_1 \varphi(x_1, y) dx_1 dy = - \int_U f \partial_1 \varphi dx,
$$

since $\varphi$ vanishes on $\partial V_y$ for each $y \in \mathbb{R}^{d-1}$ with $V_y \neq \emptyset$. Inductively one shows that

$$
(3.1) \quad \int_U (\partial^\alpha f) \varphi dx = (-1)^{|\alpha|} \int_U f \partial^\alpha \varphi dx,
$$

for all $f \in C^k(U)$, $\varphi \in C^\infty_c(U)$ and $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \leq k$. 

29
We set

\( L^1_{\text{loc}}(U) = \{ f : U \to \mathbb{C} : f \text{ is measurable}, f|_K \in L^1(K) \text{ for all compact } K \subseteq U \} \).

We extend \( f \in L^1_{\text{loc}}(U) \) by 0 to a measurable function \( f : \mathbb{R}^d \to \mathbb{C} \) without further notice. Convergence in \( L^1_{\text{loc}}(U) \) means that \( (f_n)|_K \) converges in \( L^1(K) \) for all compact \( K \subseteq U \). Observe that \( L^p(U) \subseteq L^1_{\text{loc}}(U) \) for all \( 1 \leq p \leq \infty \).

**Definition 3.1.** Let \( f \in L^1_{\text{loc}}(U) \) and \( \alpha \in \mathbb{N}_0^d \). If there is a function \( g \in L^1_{\text{loc}}(U) \) such that

\[
\int_U g \varphi dx = (-1)^{|\alpha|} \int_U f \partial^\alpha \varphi dx,
\]

for all \( \varphi \in C_c^\infty(U) \), then \( g =: \partial^\alpha f \) is called weak derivative of \( f \). If \( f \) possesses weak derivatives for all \( |\alpha| \leq k \), then we write \( f \in W^k(U) \). Moreover, one defines the Sobolev spaces by

\[
W^k_p(U) = \{ f \in L^p(U) \cap W^k(U) : \partial^\alpha f \in L^p(U) \text{ for all } |\alpha| \leq k \}
\]

for \( k \in \mathbb{N} \) and \( 1 \leq p \leq \infty \) and endows them with

\[
\|f\|_{k,p} = \begin{cases} \left( \sum_{0 \leq |\alpha| \leq k} \|\partial^\alpha f\|_{p}^p \right)^{1/p}, & 1 \leq p < \infty, \\ \max_{0 \leq |\alpha| \leq k} \|\partial^\alpha f\|_{\infty}, & p = \infty, \end{cases}
\]

where \( \partial^\alpha f := f \). We write \( W^0_\infty(U) = L^p(U) \).

As usually, the spaces \( L^1_{\text{loc}}(U) \), \( W^k(U) \) and \( W^k_p(U) \) are spaces of equivalence classes modulo the subspace \( \mathcal{N} = \{ f : U \to \mathbb{R} : f \text{ is measurable}, f = 0 \text{ a.e.} \} \).

**Remark 3.2.** a) We will see in Lemma 3.5 that \( \partial^\alpha f \) is uniquely defined. It is then also clear that the map

\[ \partial^\alpha : W^k(U) \to L^1_{\text{loc}}(U) \]

is linear if \( |\alpha| \leq k \).

b) Formula (3.1) implies that \( C^k(U) + \mathcal{N} \subseteq W^k(U) \) and that weak and classical derivatives coincide for \( f \in C^k(U) \).

c) Let \( 1 \leq p \leq \infty \) and \( k \in \mathbb{N} \). Clearly, \( (W^k_\infty(U), \|\cdot\|_{k,p}) \) is a normed vector space and

\[ J : W^k_p(U) \to L^p(U)^m, \quad f \mapsto (\partial^\alpha f)_{|\alpha| \leq m}, \]

is an isometry where \( m = 1 + d + \ldots + d^k \). We see in the proof of the next proposition that \( W^k_p(U) \) is isometrically isomorphic to a closed subspace of \( L^p(U)^m \).

d) Let \( 1 \leq p \leq \infty \) and \( k \in \mathbb{N} \). Since the \( p \)-norm and the \( 1 \)-norm on \( \mathbb{R}^m \) are equivalent, there are constants \( C_k, c_k > 0 \) such that

\[
c_k \sum_{0 \leq |\alpha| \leq k} \|\partial^\alpha f\|_{p} \leq \|f\|_{k,p} \leq C_k \sum_{0 \leq |\alpha| \leq k} \|\partial^\alpha f\|_{p}
\]
3.1. BASIC PROPERTIES

for all \( f \in W^k_p(U) \).

e) Observe that \( \|f\|_{1,p}^p = \|f\|^p_p + \|\nabla f\|^p_p \) for all \( 1 \leq p < \infty \).

**Proposition 3.3.** For all \( 1 \leq p \leq \infty \) and \( k \in \mathbb{N} \), \( W^k_p(U) \) is a Banach space. It is separable if \( p < \infty \) and reflexive if \( 1 < p < \infty \). Moreover, \( W^k_2(U) \) is a Hilbert space endowed with the scalar product

\[
\langle f, g \rangle_{W^k_2} = \sum_{|\alpha| \leq k} \int_U (\partial^\alpha f) \overline{\partial^\alpha g} dx.
\]

**Proof.** Let \( (f_n)_n \) be a Cauchy sequence in \( W^k_p(U) \). Then \( (\partial^\alpha f_n)_n \) is a Cauchy sequence in \( L^p(U) \) for every \( \alpha \in \mathbb{N}_0^d \) with \( |\alpha| \leq k \) and thus \( \partial^\alpha f_n \to g_\alpha \) in \( L^p(U) \) for some \( g_\alpha \in L^p(U) \) as \( n \to \infty \), where we set \( f_0 := g_0 \).

Let \( \varphi \in C^\infty_c(U) \) and \( |\alpha| \leq k \). Since \( f_n \in W^k_p(U) \), we deduce

\[
\int_U f \partial^\alpha \varphi dx = \lim_{n \to \infty} \int_U f_n \partial^\alpha \varphi dx = \lim_{n \to \infty} (-1)^{|\alpha|} \int_U (\partial^\alpha f_n) \varphi dx
= (-1)^{|\alpha|} \int_U g_\alpha \varphi dx.
\]

This means that \( g_\alpha \) is the weak derivative \( \partial^\alpha f \) so that \( f \in W^k_p(U) \) and \( f_n \to f \) in \( W^k_p(U) \). Hence, \( W^k_p(U) \) is a Banach space. Using Remark 1.60 in [FA], we then deduce from Remark 3.2c) that \( W^k_2(U) \) is isometrically isomorphic to a closed subspace of \( L^p(U) \). The remaining assertions now follow by isomorphy from known results of functional analysis. \( \square \)

**Example 3.4.** a) Let \( f \in C_c(\mathbb{R}) \) be such that \( f_{\pm} := f_{\mathbb{R} \pm} \) belong to \( C^1(\mathbb{R}_\pm) \). We then have \( f \in W^1(\mathbb{R}) \) with

\[
\partial f := \partial_1 f = \left\{ \begin{array}{ll}
   f'_+ & \text{on } [0, \infty) \\
   f'_- & \text{on } (-\infty, 0)
\end{array} \right\} =: g.
\]

For \( f(x) = |x| \), we thus obtain \( \partial f = \mathbb{1}_{\mathbb{R}^+} - \mathbb{1}_{(-\infty, 0)} \).

**Proof.** For every \( \varphi \in C^\infty_c(U) \), we compute

\[
\int_{\mathbb{R}} f \varphi' dt = \int_{-\infty}^0 f_- \varphi' dt + \int_0^\infty f_+ \varphi' dt
= -\int_{-\infty}^0 f'_- \varphi dt + f_- \varphi |^0_{-\infty} - \int_0^\infty f'_+ \varphi dt + f_+ \varphi |^\infty_0
= -\int_{\mathbb{R}} g \varphi dt,
\]

since \( f_+(0) = f_-(0) \) by the continuity of \( f \). \( \square \)

b) The function \( f = \mathbb{1}_{\mathbb{R}^+} \) does not belong to \( W^1(\mathbb{R}) \).
Proof. Assume there would exist $g = \partial f \in L^1_{\text{loc}}(\mathbb{R})$. Then we would obtain for every $\varphi \in C_c^\infty(\mathbb{R})$ that
\[
\int_{\mathbb{R}} g \varphi dt = - \int_{\mathbb{R}} \mathbf{1}_{\mathbb{R}^+} \varphi' dt = - \int_0^\infty \varphi'(t) dt = \varphi(0).
\]
Taking $\varphi$ with $\text{supp } \varphi \subseteq (0, \infty)$, we deduce from Lemma 3.5 below that $g = 0$ on $(0, \infty)$. Similarly, it follows that $g = 0$ on $(-\infty, 0)$. Hence, $g = 0$ and so $\varphi(0) = 0$ for all $\varphi \in C_c^\infty(\mathbb{R})$, which is false.

c) Let $d \geq 2$, $U = B(0, 1)$, $1 \leq p < d$, and $f(x) = \ln|x|/2$ for $x \in U \setminus \{0\}$. Then we have $f \in W^1_p(U)$ with
\[
\partial_j f(x) = \frac{x_j}{|x|^2} =: g_j(x),
\]
for $x \neq 0$. Observe that $f$ is unbounded and has no continuous extension at $x = 0$.

Proof. Using polar coordinates, we obtain
\[
\|f\|_p^p = c \int_0^1 \ln r |r|^{d-1} dr < \infty,
\]
for all $p \in [1, \infty)$. Estimating $|x_j| \leq r$, we further compute
\[
\|g_j\|_p^p \leq c \int_0^1 r^p |r|^{d-1} dr = \int_0^1 r^{d-p-1} dr < \infty,
\]
and thus $g_j \in L^p(U)$, for all $p \in [1, d)$ and $j \in \{1, \ldots, d\}$. Take $j = 1$, $\varepsilon \in (0, 1)$ and $\varphi \in C_c^\infty(U)$. We set $J_\varepsilon = (-1, -\varepsilon) \cup [\varepsilon, 1)$ and write $y = (x_2, \ldots, x_d)$. We then obtain
\[
\int_U f \partial_1 \varphi dx = \int_{[-1,1]^d} f(x_1, y \partial_1 \varphi(x_1, y)) dx_1 dy
\]
\[
= \lim_{\varepsilon \to 0} \int_{[-1,1]^{d-1}} \int_{J_\varepsilon} \left( \ln \sqrt{x_1^2 + |y|^2} \right) \partial_1 \varphi(x_1, y) dx_1 dy
\]
\[
= \lim_{\varepsilon \to 0} \int_{[-1,1]^{d-1}} \left[ \int_{J_\varepsilon} \frac{x_1}{x_1^2 + |y|^2} \varphi(x_1, y) dx_1 - \left( \ln \sqrt{\varepsilon^2 + |y|^2} \right) \varphi(\varepsilon, y) + \left( \ln \sqrt{\varepsilon^2 + |y|^2} \right) \varphi(-\varepsilon, y) \right] dy
\]
\[
= -\int U g_1 \varphi dx.
\]
Here we used the theorem of dominated convergence with majorants $\|\partial_1 \varphi\|_\infty |f|$ and $\|\varphi\|_\infty |g_1|$ and that
\[
\left| \left( \ln \sqrt{\varepsilon^2 + |y|^2} \right) (\varphi(\varepsilon, y) - \varphi(-\varepsilon, y)) \right| \leq 2 \varepsilon \left( |\ln \varepsilon| + \ln \sqrt{d} \right) \|\partial_1 \varphi\|_\infty
\]
converges to 0 as $\varepsilon \to 0$ uniformly in $y \in (-1, 1)^{d-1}$. One similarly sees that $\partial_j f = g_j$ for $j = 2, \ldots, d$. \qed
3.1. BASIC PROPERTIES

We next investigate the properties of mollifiers. Besides duality, they are the basic tool in the study of Sobolev spaces.

Set \( \chi(x) = \exp \frac{-1}{1-|x|^2} \) for \( |x| < 1 \) and \( \chi(x) = 0 \) for \( |x| \geq 1 \), where \( x \in \mathbb{R}^d \).

Observe that \( \chi \in C^\infty(\mathbb{R}^d) \). We define

\[
\Psi(x) = \frac{1}{\|\chi\|_1} \chi(x) \quad \text{and} \quad \Psi_\varepsilon(x) = \varepsilon^{-d} \Psi\left(\frac{1}{\varepsilon}x\right),
\]

for \( x \in \mathbb{R}^d \) and \( \varepsilon > 0 \). We then have \( 0 \leq \Psi_\varepsilon \in C^\infty(\mathbb{R}^d) \), \( \Psi_\varepsilon(x) > 0 \) if and only if \( |x| < \varepsilon \), \( \Psi = 0 \) on \( \mathbb{R}^d \setminus B(0, \varepsilon) \) and \( \|\Psi_\varepsilon\|_1 = 1 \). For \( f \in L^1_\text{loc}(\mathbb{R}^d) \) and \( \varepsilon > 0 \), we now introduce the mollifier \( T_\varepsilon \) by

\[
T_\varepsilon f(x) = (\Psi_\varepsilon * f)(x) = \int_{B(x,\varepsilon)} \Psi_\varepsilon(x-y)f(y)dy = \int_{B(0,\varepsilon)} \Psi_\varepsilon(z)f(x-z)dz,
\]

for \( x \in \mathbb{R}^d \), where we have put \( f(x) := 0 \) for \( x \in \mathbb{R}^d \setminus U \). From e.g. Example 3.9 and Proposition 3.10 in [FA], we recall that

\[
T_\varepsilon f \in C^\infty(\mathbb{R}^d),
\]

\[
\supp T_\varepsilon f \subseteq S_\varepsilon := S + B(0, \varepsilon), \quad \text{if} \quad \supp f = S,
\]

\[
\|T_\varepsilon f\|_{L^p(U)} \leq \|T_\varepsilon f\|_{L^p(\mathbb{R}^d)} \leq \|\Psi_\varepsilon\|_1 \|f\|_p = \|f\|_p, \quad \text{if} \quad f \in L^p(U), \ 1 \leq p \leq \infty,
\]

\[
T_\varepsilon f \to f \text{ in } L^p(U) \text{ as } \varepsilon \to 0, \quad \text{if} \quad f \in L^p(U) \text{ and } 1 \leq p < \infty.
\]

Observe that \( S_\varepsilon \) is compact if \( S \) is compact.

**Lemma 3.5.** Let \( K \subseteq U \) be compact. Then there is a function \( \varphi \in C_c^\infty(U) \) such that \( 0 \leq \varphi \leq 1 \) on \( U \) and \( \varphi = 1 \) on \( K \). Let \( f \in L^1_\text{loc}(U) \) satisfy

\[
\int_U f \psi dx = 0
\]

for all \( \psi \in C_c^\infty(U) \). Then \( f = 0 \) a.e. In particular, weak derivatives are uniquely defined.

**Proof.** Assume that \( f \neq 0 \) on a Borel set \( B \subseteq U \) with \( \lambda(B) > 0 \). By Analysis 3, there is a compact set \( K \subseteq B \subseteq U \) with \( \lambda(K) > 0 \). Fix \( 0 < \delta < \frac{1}{2} \text{dist}(\partial K, \partial U) \) so that \( K_{2\delta} \subseteq U \). The function \( \varphi := T_\delta \chi_{K_{\delta}} \) belongs to \( C_c^\infty(U) \) by (3.4) and (3.5), where \( \supp \varphi \subseteq K_{2\delta} \). Moreover, (3.6) and (3.3) yield that \( 0 \leq \varphi(x) \leq \|\varphi\|_\infty \leq \|\chi_{K_{\delta}}\|_\infty = 1 \) for all \( x \in U \) and that

\[
\varphi(x) = \int_{B(x,\delta)} \Psi_{\varepsilon}(x-y)\chi_{K_{\delta}}(y)dy = \|\Psi_{\varepsilon}\|_1 = 1
\]

for all \( x \in K \). This construction shows the first claim.

Since \( \varphi f \in L^1(U) \), the functions \( T_\varepsilon(\varphi f) \) converge to \( \varphi f \) in \( L^1(U) \) as \( \varepsilon \to 0 \), due to (3.7). Hence, there is a nullset \( N \) and a subsequence \( \varepsilon_j \to 0 \) with \( \varepsilon_j \leq \delta \), such
that \((T_{\varepsilon_j}(\varphi f))(x) \to f(x) \neq 0\) as \(j \to \infty\) for each \(x \in K \setminus N\). For every \(x \in K \setminus N\) and \(j \in \mathbb{N}\), we also deduce
\[
(T_{\varepsilon_j}(\varphi f))(x) = \int_U \Psi_{\varepsilon_j}(x - y)\varphi(y)f(y)dy = 0
\]
from the assumption, since the function \(y \mapsto \Psi_{\varepsilon_j}(x - y)\varphi(y)\) belongs to \(C_c^\infty(U)\). This is a contradiction. 

\(\square\)

Recall that Hölder’s inequality implies that the map
\[
L^p(B) \times L^p(B) \to \mathbb{C}, \quad (f, g) \mapsto \int_B fgdx,
\]
is continuous for all \(1 \leq p \leq \infty\) and Borel sets \(B \subseteq \mathbb{R}^d\).

**Lemma 3.6.** a) Let \(f \in L^p_{\text{loc}}(U)\) possess the weak derivative \(\partial^\alpha f \in L^p_{\text{loc}}(U)\) for some \(\alpha \in \mathbb{N}_0^d\) and \(p \in [1, \infty)\). Then the functions \(T_{\varepsilon}f \in C^\infty(U)\) converge to \(f\) and \(\partial^\alpha(T_{\varepsilon}f)\) converge to \(\partial^\alpha f\) in \(L^p_{\text{loc}}(U)\) as \(\varepsilon \to 0\). We further have \(\partial^\alpha(T_{\varepsilon}f)(x) = T_{\varepsilon}(\partial^\alpha f)(x)\) for all \(x \in U\) and \(\varepsilon < d(x, \partial U)\). Moreover, there is a sequence \(\varepsilon_n \to 0\) such that \(T_{\varepsilon_n}f \to f\) and \(\partial^\alpha(T_{\varepsilon_n}f) \to \partial^\alpha f\) a.e. on \(U\) as \(n \to \infty\). If \(f \in W^k(U)\), we can take the same \(f_n\) for all \(|\alpha| \leq k\).

b) If \(f, g \in L^1(\partial U)\) and there are \(f_n \in W^{[\alpha]}(U)\) such that \(f_n \to f\) and \(\partial^\alpha f_n \to g\) in \(L^1_{\text{loc}}(U)\) as \(n \to \infty\), then \(g\) is the weak derivative \(\partial^\alpha f\). If this convergence holds in \(L^p(U)\) for some \(p \in [1, \infty]\) and all \(\alpha\) with \(|\alpha| \leq k\), then \(f \in W^k_p(U)\).

**Proof.** a) Let \(\varepsilon > 0\) and \(x \in U\). If \(\varepsilon < d(x, \partial U)\), then the function \(y \mapsto \varphi_{\varepsilon,x}(y) = \Psi_{\varepsilon}(x - y)\) belongs to \(C_c^\infty(U)\). Using a corollary to Lebesgue’s theorem (see Analysis 3) and Definition 3.1, we can thus deduce
\[
\partial^\alpha T_{\varepsilon}f(x) = \int_U \partial^\alpha_x \Psi_{\varepsilon}(x - y)f(y)dy = (-1)^{|\alpha|} \int_U (\partial^\alpha \varphi_{\varepsilon,x})(y)f(y)dy
\]
\[
= \int_U \varphi_{\varepsilon,x}(y)(\partial^\alpha f)(y)dy = (T_{\varepsilon}\partial^\alpha f)(x).
\]
Choose a compact subset \(K \subseteq U\) and fix \(\delta > 0\) with \(K_\delta \subseteq U\). Take \(\varepsilon \in (0, \delta]\). Note that the integrand of \((T_{\varepsilon}g)(x)\) is then supported in \(K_\delta\) for every \(x \in K\) and \(g \in L^1_{\text{loc}}(U)\), see (3.3). Due to (3.7), the functions
\[
1_K \partial^\alpha (T_{\varepsilon}f) = 1_K T_{\varepsilon}(\partial^\alpha f) = 1_K (\partial^\alpha (1_K T_{\varepsilon}f))
\]
converge in \(L^p(K)\) to \(1_K \partial^\alpha f = 1_K \partial^\alpha f\) as \(\varepsilon \to 0\). So we have shown the asserted convergence in \(L^p_{\text{loc}}(U)\). The sets
\[
K_m = \{x \in U : d(x, \partial U) \geq \frac{1}{m} \text{ and } |x|_2 \leq m\}
\]
are compact and \(\bigcup_{m \in \mathbb{N}} K_m = U\). Let \(\varepsilon_k \to 0\). Then, for each \(m \in \mathbb{N}\) there is a null set \(N_m \subseteq K_m\) and a subsequence \(\nu_m(j)\) such that \(\partial^\alpha T_{\varepsilon_{\nu_m(j)}} f(x)\) converges to \(\partial^\alpha f(x)\)
and $T_{\varepsilon_n(i,j)}f(x)$ converges to $f(x)$ for all $x \in K_m \setminus N_m$ as $j \to \infty$. By means of a diagonal sequence, one obtains a sequence $\varepsilon_n \to 0$ such that $\partial^\alpha T_{\varepsilon_n} f(x) \to \partial^\alpha f(x)$ and $T_{\varepsilon_n} f(x) \to f(x)$ for $x \in U \setminus \bigcup_{m \in \mathbb{N}} N_m$, as $n \to \infty$, where $\bigcup_{m \in \mathbb{N}} N_m$ is a null set. This procedure can also be done for finitely many $\partial^\alpha f$ at the same time.

b) Let $f_n \in W^{1,1}(U)$ be given such that $f_n \to f$ and $\partial^\alpha f_n \to g$ in $L^1_{\text{loc}}(U)$ as $n \to \infty$. From (3.8) on supp $\varphi$ we deduce that
\[
\int_U f \partial^\alpha \varphi dx = \lim_{n \to \infty} \int_U f_n \partial^\alpha \varphi dx = (-1)^\alpha \lim_{n \to \infty} \int_U (\partial^\alpha f_n) \varphi dx = (-1)^\alpha \int_U g \varphi dx
\]
for all $\varphi \in C_c^\infty(U)$. Hence, $g = \partial^\alpha f$. In the setting of the last assertion we thus obtain $\partial^\alpha f \in L^p(U)$ for all $|\alpha| \leq k$, and hence $f \in W^k_p(U)$.

**Proposition 3.7.** a) Let $f, g \in W^{1}(U) \cap L^\infty(U)$. Then, $f, g \in W^{1}(U) \cap L^\infty(U)$ and
\[
\partial_j (fg) = (\partial_j f) g + f(\partial_j g)
\]
holds for all $j \in \{1, \ldots, d\}$.

b) Let $1 \leq p \leq \infty$, $f \in W^1_p(U)$ and $g \in W^1_{p'}(U)$. Then, $fg \in W^1_1(U)$ and (3.9) holds.

**Proof.** 1) Let $f, g \in W^{1}(U)$. Set $f_n = T_{\varepsilon_n} f \in C^\infty(U)$ and $g_n = T_{\varepsilon_n} g \in C^\infty(U)$ with $\varepsilon_n \to 0$ as in Lemma 3.6a). Fix $m \in \mathbb{N}$ and take $\varphi \in C_c^\infty(U)$ and $j \in \{1, \ldots, d\}$. Choose an open and bounded set $V$ such that supp $\varphi \subseteq V \subseteq \overline{V} \subseteq U$. Since $f_n \to f$ and $\partial_j f_n \to \partial_j f$ on $L^1(\overline{V})$ by Lemma 3.6a), the formulas (3.8) and (3.1) yield
\[
\int_U fg \partial_j \varphi dx = \lim_{n \to \infty} \int_V f_n g_m \partial_j \varphi dx = -\lim_{n \to \infty} \int_V ((\partial_j f_n) g_m + f_n (\partial_j g_m)) \varphi dx
\]
so that the weak derivative $\partial_j (fg_m) = (\partial_j f) g_m + f(\partial_j g_m) \in L^1_{\text{loc}}(U)$ exists.

2) Let $f, g \in W^{1}(U) \cap L^\infty(U)$ and $g_m$ as in 1). Note that $g_m \to g$ and $\partial_j g_m \to \partial_j g$ in $L^1_{\text{loc}}(U)$ as $m \to \infty$. Since $f$ is bounded, we obtain
\[
\int_U fg \partial_j \varphi dx = \lim_{m \to \infty} \int_U fg_m \partial_j \varphi dx = \lim_{m \to \infty} \left[ \int_U (\partial_j f) g_m \varphi dx + \int_U f (\partial_j g_m) \varphi dx \right],
\]
using Step 1). On the right hand side, the second integral converges to $\int_U f \partial_j g \varphi dx$, again because of $f \in L^\infty(U)$. For the first integral we use that $g_m \to g$ a.e. by Lemma 3.6a) and that $\|g_m\|_{L^\infty} \leq \|g\|_{L^\infty}$ by (3.6). The theorem of dominated convergence (with the majorant $|\partial_j f||g|_{L^\infty} \mathbf{1}_{\text{supp} \varphi}$) yields
\[
\int_U fg \partial_j \varphi dx = -\int_U ((\partial_j f) g + f(\partial_j g)) \varphi dx,
\]
as required. Note that $(\partial_j f) g + f(\partial_j g)$ belongs to $L^1_{\text{loc}}(U)$ by our assumptions.

3) Let $f \in W^1_p(U)$ and $g \in W^1_{p'}(U)$. If $p \in (1, \infty]$ we show (3.9) as in Step 2), using (3.8) and that $g_m, \partial_j g_m$ converge in $L^p_{\text{loc}}(U)$ by Lemma 3.6a). If $p = 1$, we replace the roles of $f$ and $g$. Hölder’s inequality and (3.9) then yield $\partial_j (fg) \in L^1(U)$. \hfill \Box
Proposition 3.8. a) Let \( f \in W^1(U) \) be real valued and \( h \in C^1(\mathbb{R}) \) with \( h' \in C^0(\mathbb{R}) \). We then have \( h \circ f \in W^1(U) \) and

\[
\partial_j (h \circ f) = (h' \circ f) \partial_j f
\]

for all \( j \in \{1, \ldots, d\} \).

b) Let \( f \in W^1(U) \), \( V \subseteq \mathbb{R}^d \) be open and \( \Phi : V \to U \) be a diffeomorphism such that \( \Phi' \) and \( (\Phi^{-1})' \) are bounded. We then have \( f \circ \Phi \in W^1(V) \) and

\[
\partial_j (f \circ \Phi) = \sum_{m=1}^d ((\partial_m f) \circ \Phi) \partial_j \Phi_m
\]

for all \( j = 1, \ldots, d \).

In both results we can replace \( W^1(U) \) by \( W^{1,p}_U \) for \( 1 \leq p \leq \infty \), if in a) also \( h(0) = 0 \) holds in the case that \( \lambda(U) = \infty \) and \( p < \infty \).

Proof. a) By Lemma 3.6, there are \( f_n \in C^\infty(U) \) such that \( f_n \to f \) and \( \partial_j f_n \to \partial_j f \) in \( L^1_{\text{loc}}(U) \) and a.e. as \( n \to \infty \), for every \( j \in \{1, \ldots, d\} \). Since

\[
|h(f(x))| \leq |h(f(x)) - h(0)| + |h(0)| \leq ||h'||_\infty |f(x)| + |h(0)|
\]

for all \( x \in U \), the function \( h \circ f \) belongs to \( L^1_{\text{loc}}(U) \) (and to \( L^p(U) \) if \( f \in L^p(U) \) and if \( h(0) = 0 \) in the case that \( \lambda(U) = \infty \) and \( p \neq \infty \)). Let \( K \subseteq U \) be compact. We obtain that

\[
\int_K |h(f_n(x)) - h(f(x))|dx \leq ||h'||_\infty \int_K |f_n(x) - f(x)|dx \to 0,
\]

\[
\int_K |h'(f_n(x))\partial_j f_n(x) - h'(f(x))\partial_j f(x)|dx
\]

\[
\leq ||h'||_\infty \int_K |\partial_j f_n(x) - \partial_j f(x)|dx + \int_K |h'(f_n(x)) - h'(f(x))| |\partial_j f(x)| dx \to 0
\]

as \( n \to \infty \) where we also used Lebesgue’s theorem and the majorant \( 2||h'||_\infty |\partial_j f| \) in the last integral. Since \( h \circ f_n \in C^1(U) \), \( (h' \circ f) \partial_j f \in L^1_{\text{loc}}(U) \) and \( \partial_j (h \circ f_n) = (h' \circ f_n) \partial_j f_n \), Lemma 3.6b) yields assertion a). If \( f \in W^{1,p}_U \) then \( (h' \circ f) \partial_j f \in L^p(U) \) and so \( h \circ f \in W^{1,p}_U \).

Assertion b) can be shown similarly using the transformation rule. \( \square \)

Corollary 3.9. Let \( f \in W^1(U) \) be real valued. Then \( f_+, f_-, |f| \in W^1(U) \) with

\[
\partial_j f_\pm = \pm \mathbb{1}_{\{f \geq 0\}} \partial_j f \quad \text{and} \quad \partial_j |f| = \left( \mathbb{1}_{\{f > 0\}} - \mathbb{1}_{\{f < 0\}} \right) \partial_j f
\]

for all \( j \in \{1, \ldots, d\} \). Here one can replace \( W^1 \) by \( W^{1,p}_U \) for all \( 1 \leq p \leq \infty \).

Proof. We employ the function \( h_\varepsilon \in C^1(\mathbb{R}) \) given by \( h_\varepsilon(t) := \sqrt{t^2 + \varepsilon^2} - \varepsilon \leq t \) for \( t \geq 0 \) and \( h_\varepsilon(t) := 0 \) for \( t < 0 \), where \( \varepsilon > 0 \). Observe that \( ||h'_\varepsilon||_\infty = 1 \) and

36
3.1. BASIC PROPERTIES

$h_\varepsilon(t) \to t$ for $t > 0$ and $h_\varepsilon(t) \to 0$ for $t \leq 0$, as $\varepsilon \to 0$. Proposition 3.8 shows that $h_\varepsilon \circ f \in W^1(U)$ and

$$
\int_U h_\varepsilon(f) \partial_j \varphi dx = -\int_U h'_\varepsilon(f)(\partial_j f) \varphi dx = -\int_{\{f > 0\}} \frac{f}{\sqrt{f^2 + \varepsilon^2}}(\partial_j f) \varphi dx
$$

for each $\varphi \in C_0^\infty(U)$. Thanks to the majorants $\|\partial_j \varphi\|_\infty 1_B|f|$ and $\|\varphi\|_\infty |\partial_j f| 1_B$ with $B = \text{supp} \varphi$, Lebesgue’s convergence theorem shows that

$$
\int_U f_+ \partial_j \varphi dx = -\int_{\{f > 0\}} \frac{f}{|f|}(\partial_j f) \varphi dx = -\int_U 1_{\{f > 0\}}(\partial_j f) \varphi dx.
$$

There thus exists $\partial_j f_+ = 1_{\{f > 0\}} \partial_j f \in L^1_{\text{loc}}(U)$. The other assertions follow from $f_- = -(f)_+$ and $|f| = f_+ + f_-$. \qed

**Theorem 3.10.** Let $J \subseteq \mathbb{R}$ be an open interval and $f \in L^1_{\text{loc}}(J)$. We then have $f \in W^1(J)$ if and only if there is a $g \in L^1_{\text{loc}}(J)$ and a continuous representative of $f$ such that

$$(3.10) \quad f(t) = f(s) + \int_s^t g(\tau) d\tau$$

for all $s, t \in J$. In this case, $g = \partial f$ a.e..

**Proof.** 1) Let $f \in W^1(J)$. Take $f_n = T_{\varepsilon_n} f \in C^\infty(J)$ from Lemma 3.6a). Then for a.e. $t \in J$ and for a.e. $t_0 \in J$ we have

$$
|f(t) - f(t_0)| = \lim_{n \to \infty} |f_n(t) - f_n(t_0)| = \lim_{n \to \infty} \int_{t_0}^t f'_n(\tau) d\tau = \int_{t_0}^t |f(\tau)| d\tau.
$$

Fixing one such $t_0$ and noting that $t \mapsto \int_{t_0}^t |f(\tau)| d\tau$ is continuous, we obtain a continuous representative of $f$ which satisfies (3.10) for $s = t_0$ and $g = \partial f$. Subtracting two such equations for any given $t, s \in J$ and the fixed $t_0$, we deduce (3.10) with $g = \partial f$ for all $t, s \in J$.

2) If (3.10) holds for some $g \in L^1_{\text{loc}}(J)$, take $g_n \in C^\infty(J)$ such that $g_n \to g$ in $L^1_{\text{loc}}(J)$ as $n \to \infty$. For any $s \in J$ and $n \in \mathbb{N}$, the function $f_n(t) := f(s) + \int_s^t g_n(\tau) d\tau$, $t \in J$ belongs to $C^\infty(J)$ with $f'_n = g_n$. Moreover, for $[a, b] \subset J$ with $s \in [a, b]$, we estimate

$$
\|f_n - f\|_{L^1([a, b])} \leq \int_a^b \int_s^t |g_n(\tau) - g(\tau)| d\tau dt \leq (b - a)\|g_n - g\|_{L^1([a, b])},
$$

using (3.10), so that $f_n \to f$ in $L^1_{\text{loc}}(J)$ as $n \to \infty$. Lemma 3.6b) then yields $f \in W^1(J)$ and $\partial f = g$. \qed

**Remark 3.11.** a) Let $J = (a, b)$ for some $a < b$ in $\mathbb{R}$ and $f : J \to \mathbb{C}$. We then have $f \in W^1(J)$ if and only if $f$ is absolutely continuous, i.e.
∀ ε > 0 ∃ δ > 0 ∀ a < α_1 < β_1 < ⋯ < α_n < β_n < b, n ∈ N with \( \sum_{i=1}^{n} (β_j - α_j) \leq δ \) it holds

\[
\sum_{j=1}^{n} |f(β_j) - f(α_j)| \leq ε.
\]

In this case, \( f \) is differentiable for a.e. \( t \in J \) and the pointwise derivative \( f' \) is equal a.e. to the weak derivative \( \partial f \in L^1(J) \).

(Remark: A Lipschitz continuous function is absolutely continuous; an absolutely continuous function is uniformly continuous.)

**Proof.** The implication “⇒” holds since (3.10) yields that

\[
\sum_{j=1}^{n} |f(β_j) - f(α_j)| = \sum_{j=1}^{n} \left| \int_{α_j}^{β_j} f'(τ) dτ \right| \leq \int_{\bigcup_{j=1}^{n} (α_j, β_j)} |f'(τ)| dτ =: S,
\]

where \( S \to 0 \) as \( λ(\bigcup_{j=1}^{n} (α_j, β_j)) \to 0 \).

The other implication “⇐” and the last assertion is shown in Theorem 7.20 of [Rud87] (combined with our Theorem 3.10).

b) There is a continuous increasing function \( f : [0, 1] \to \mathbb{R} \) with \( f(0) = 0 \) and \( f(1) = 1 \) such that \( f'(t) = 0 \) exists for a.e. \( t \in [0, 1] \). Consequently,

\[
1 = f(1) \neq f(0) + \int_{0}^{1} f'(τ) dτ = 0,
\]

and this function is not absolutely continuous, does not belong to \( W^1_1((0, 1)) \) and violates (3.10). (See §7.17 in [Rud87].)

**Proposition 3.12.** Let \( U \) be convex. Then \( W^1_∞(U) \) is isomorphic to

\[
C^1_b(U) = \{ f \in C_b(U) : f \text{ is Lipschitz} \},
\]

and the norm of \( W^1_∞(U) \) is equivalent to

\[
\| f \|_{C^1_b} = \| f \|_∞ + [f]_{Lip},
\]

where \([f]_{Lip}\) is the Lipschitz constant of \( f \).

**Proof.** Let \( f \in W^1_∞(U) \). Take \( ε_n \to 0 \) from Lemma 3.6a). Let \( K \subset U \) be compact. For sufficiently large \( n \in \mathbb{N} \), Lemma 3.6 and (3.6) yield

\[
|\partial_j T_{ε_n} f(z)| = |T_{ε_n} \partial_j f(z)| \leq \| \partial_j f \|_∞ \leq \| f \|_{1,∞},
\]

for all \( j \in \{1, \ldots, d\} \) and \( z \in K \). Using that \( T_{ε_n} f(x) \to f(x) \) as \( n \to ∞ \) for all \( x \in U \setminus N \) and a null set \( N \), we thus estimate

\[
|f(x) - f(y)| = \lim_{n \to ∞} |T_{ε_n} f(x) - T_{ε_n} f(y)|
\]
3.2. Density and Embedding Theorems

\[ = \lim_{n \to \infty} \left| \int_{0}^{1} \nabla T_{\varepsilon_{n}} f(y + \tau(x - y)) \cdot (x - y) \, d\tau \right| \leq \|f\|_{1, \infty} |x - y|/2 \]

for all \(x, y \in U \setminus N\). Hence, \(f\) has a representative with Lipschitz constant \(\|f\|_{1, \infty}\).

Let \(f \in C_0^1(U)\). Take \(\varphi \in C_0^\infty(U)\), \(j \in \{1, \ldots, d\}\), and \(\delta > 0\) such that \((\text{supp } \varphi)_\delta \subseteq U\). For \(\varepsilon \in (0, \delta)\) the difference quotient \(\frac{1}{\varepsilon}(\varphi(x + \varepsilon e_j) - \varphi(x))\) converges uniformly on \(\text{supp } \varphi\) as \(\varepsilon \to 0\), and hence

\[ |\int_{U} f \partial_j \varphi dx| = \lim_{\varepsilon \to 0} \left| \int_{\text{supp } \varphi} f(x) \frac{1}{\varepsilon}(\varphi(x + \varepsilon e_j) - \varphi(x)) \, dx \right| \]

\[ \leq \lim_{\varepsilon \to 0} \int_{\text{supp } \varphi} \frac{1}{\varepsilon}|f(y - \varepsilon e_j) - f(y)| |\varphi(y)| \, dy \]

\[ \leq \|f\|_{\text{Lip}} \|\varphi\|_1. \]

Taking into account that \(C_0^\infty(U)\) is dense in \(L^1(U)\), we see that the map \(\varphi \mapsto - \int_{U} f \partial_j \varphi dx\) has a continuous linear extension \(F_j : L^1(U) \to C\). Therefore there is a function \(g_j \in L^\infty(U) = L^1(U)^*\) with \(\|g_j\|_{\infty} = \|F_j\| \leq \|f\|_{\text{Lip}}\) such that

\[ - \int_{U} f \partial_j \varphi dx = F_j(\varphi) = \int_{U} g_j \varphi dx \]

for all \(\varphi \in C_0^\infty(U)\). This means that \(f\) has the weak derivative \(\partial_j f = g_j \in L^\infty(U)\). As a result, \(f \in W_{\infty}^1(U)\) and \(\|f\|_{W_{\infty}^1(U)} \leq \|f\|_{\infty} + \|f\|_{\text{Lip}}\).

In the above proof convexity is only used to estimate the Lipschitz constant by \(\|\nabla f\|_{\infty}\). Instead of convexity, it suffices to assume that there exist a \(\delta > 0\) such that for all \(x, y \in U\) with \(|x - y| \leq \delta\) the line segment from \(x\) to \(y\) belongs to \(U\). For such \(x\) and \(y\) we can argue as above. If \(|x - y| \geq \delta\) and one chooses bounded representative of \(f\), then one obtains \(|f(x) - f(y)| \leq 2\|f\|_{\infty} \delta^{-1} |x - y|\).

3.2 Density and Embedding Theorems

In this section we prove some of the most important theorems on Sobolev spaces.

Definition 3.13. For \(k \in \mathbb{N}\) and \(1 \leq p < \infty\), the closure of \(C_0^\infty(U)\) in \(W_p^k(U)\) is denoted by \(\bar{W}_p^k(U)\).

Theorem 3.14. Let \(k \in \mathbb{N}\) and \(p \in [1, \infty)\). We then have

\[ \bar{W}_p^k(\mathbb{R}^d) = W_p^k(\mathbb{R}^d). \]

Moreover, the set \(C^\infty(U) \cap W_p^k(U)\) is dense in \(W_p^k(U)\).

Proof. We prove the theorem only for \(k = 1\), the general case can be treated similarly.

1) Let \(f \in W_1^1(\mathbb{R}^d)\). Take any \(\phi \in C^\infty(\mathbb{R})\) such that \(0 \leq \phi \leq 1\), \(\phi = 1\) on \([0, 1]\) and \(\phi = 0\) on \([2, \infty)\). Set

\[ \varphi_n(x) = \phi \left( \frac{1}{n} |x|_2 \right) \text{ ("cut-off function")}. \]
for \( n \in \mathbb{N} \) and \( x \in \mathbb{R}^d \). We then have \( \varphi_n \in C_c^\infty(\mathbb{R}^d) \), \( 0 \leq \varphi_n \leq 1 \) and \( \| \partial_j \varphi_n \|_\infty \leq \| \varphi' \|_\infty \frac{1}{n} \) for all \( n \in \mathbb{N} \), as well as \( \varphi_n(x) \to 1 \) for all \( x \in \mathbb{R}^d \) as \( n \to \infty \). Thus \( \| \varphi_n f - f \|_p \to 0 \) as \( n \to \infty \) by Lebesgue’s convergence theorem. Further, Proposition 3.7 implies that

\[
\| \partial_j (\varphi_n f - f) \|_p = \| (\varphi_n \partial_j f - \partial_j f) + (\partial_j \varphi_n) f \|_p \\
\leq \| \varphi_n \partial_j f - \partial_j f \|_p + \frac{1}{n} \| \varphi' \|_\infty \| f \|_p,
\]

and the right hand side tends to 0 as \( n \to \infty \) for each \( j \in \{1, \ldots, d\} \). Given \( \varepsilon > 0 \), we can thus fix \( m \in \mathbb{N} \) such that \( \| \varphi_m f - f \|_{1,p} \leq \varepsilon \). Due to (3.4) and (3.5), the functions \( T_\frac{1}{n}(\varphi_m f) \) belong to \( C_c^\infty(\mathbb{R}^d) \) for all \( n \in \mathbb{N} \). Equation (3.7) and Lemma 3.6 further yield that

\[
T_\frac{1}{n}(\varphi_m f) \to \varphi_m f \quad \text{and} \quad \partial_j T_\frac{1}{n}(\varphi_m f) = T_\frac{1}{n} \partial_j (\varphi_m f) \to \partial_j (\varphi_m f)
\]

in \( L^p(\mathbb{R}^d) \) as \( n \to \infty \), for each \( j \in \{1, \ldots, d\} \). So there is an \( n \in \mathbb{N} \) such that

\[
\| T_\frac{1}{n}(\varphi_m f) - \varphi_m f \|_{1,p} \leq \varepsilon,
\]

and thus

\[
\| T_\frac{1}{n}(\varphi_m f) - f \|_{1,p} \leq 2\varepsilon.
\]

2) For the second assertion, we can assume that \( \partial U \neq \emptyset \). Let \( f \in W^1_p(U) \). Set

\[
U_n = \left\{ x \in U : |x|_2 < n \text{ and } d(x, \partial U) > \frac{1}{n} \right\}
\]

for all \( n \in \mathbb{N} \). Then \( U_n \subseteq \overline{U_n} \subseteq U_{n+1} \subseteq U \), \( \overline{U_n} \) is compact and \( \bigcup_{n=1}^\infty U_n = U \). Observe that \( U = \bigcup_{n=1}^{\infty} U_{n+1} \setminus U_{n-1} \), where \( U_0, U_1 := \emptyset \). There are functions \( \varphi_n \in C_c^\infty(U) \) such that \( \text{supp } \varphi_n \subseteq U_{n+1} \setminus U_{n-1} \), \( \varphi_n \geq 0 \), and \( \sum_{n=1}^\infty \varphi_n(x) = 1 \) for all \( x \in U \) (see §2.19 in [Alt06] or Analysis 3).

Fix \( \varepsilon > 0 \). As in Step 1), for each \( n \in \mathbb{N} \) there is a \( \delta_n > 0 \) such that \( g_n := T_{\delta_n}(\varphi_n f) \in C_c^\infty(U) \), \( \text{supp } g_n \subseteq (\text{supp } \varphi_n f)_{\delta_n} \subseteq U_{n+1} \setminus U_{n-1} \), \( \| g_n - \varphi_n f \|_{1,p} \leq 2^{-n}\varepsilon \). Define \( g(x) = \sum_{n=1}^\infty g_n(x) \) for all \( x \in U \). Observe that on each ball \( \overline{B(x, r)} \subseteq U \), this sum is finite. Hence, \( g \in C_c^\infty(U) \). Since \( f = \sum_{n=1}^\infty \varphi_n f \), we further have

\[
g(x) - f(x) = \sum_{n=1}^\infty (g_n(x) - \varphi_n(x) f(x)),
\]

for all \( x \in U \) and \( n \in \mathbb{N} \). Due to \( \| g_n - \varphi_n f \|_{1,p} \leq 2^{-n}\varepsilon \), this series converges absolutely in \( W^1_p(U) \), and

\[
\| f - g \|_{1,p} \leq \sum_{n=1}^\infty \| g_n - \varphi_n f \|_{1,p} \leq \varepsilon.
\]
3.2. DENSITY AND EMBEDDING THEOREMS

Remark 3.15. a) If $U$ is bounded, then $W^k_p(U) \neq W^k_p(U)$, see Lemma 6.67 in [Ren04].

b) For “good” $\partial U$ one can replace in $C^\infty(U)$ by $C^\infty(\overline{U})$ in Theorem 3.14, see Corollary 3.22 below.

We now want to study embeddings of Sobolev spaces. We clearly have

$$W^k_p(U) \hookrightarrow W^j_p(U) \quad \text{if } k \geq j \geq 0$$

and

$$W^k_p(U) \hookrightarrow W^j_q(U) \quad \text{if } k \geq j \geq 0, 1 \leq q \leq p \leq \infty \text{ and } \lambda(U) < \infty.$$  

(Here we put $W^0_p(U) = L^p(U)$ for $1 \leq p \leq \infty.$) The embedding $X \hookrightarrow Y$ means that there is an injective map $J \in L(X,Y)$. Above it holds $Jf = f$, and below we also use $Jf = f + N$. Writing $c = ||J||$, one obtains $||f||_Y \leq c||f||_X$ if one identifies $Jf$ with $f$.

Theorem 3.16 (Sobolev, Morrey). Let $k \in \mathbb{N}$ and $1 \leq p < \infty$. We have the following embeddings.

a) If $kp < d$, then

$$p^* := \frac{pd}{d - kp} \in (p, \infty) \quad \text{and} \quad W^k_p(\mathbb{R}^d) \hookrightarrow L^{q}(\mathbb{R}^d)$$

for all $q \in [p, p^*].$

b) If $kp = d$, then

$$W^k_p(\mathbb{R}^d) \hookrightarrow L^{q}(\mathbb{R}^d)$$

for all $q \in [p, \infty).$

c) If $kp > d$, then there are either $j \in \mathbb{N}_0$ and $\beta \in (0,1)$ such that $k - \frac{d}{p} = j + \beta$ or $k - \frac{d}{p} \in \mathbb{N}$. In the latter case we set $j := k - \frac{d}{p} - 1 \in \mathbb{N}_0$ and take any $\beta \in (0,1)$. Then

$$W^k_p(\mathbb{R}^d) \hookrightarrow C^0_b(\mathbb{R}^d) \quad \text{and} \quad |\partial^\alpha f(x) - \partial^\alpha f(y)| \leq c|x - y|^\beta$$

for all $x, y \in \mathbb{R}^d$, $|\alpha| \leq j$, a constant $c > 0$, and a representative $f \in C^0_b(\mathbb{R}^d)$ of $f$, where

$$C^0_b(\mathbb{R}^d) = \{ u \in C^0(\mathbb{R}^d) : \partial^\alpha u(x) \to 0 \text{ as } |x| \to \infty \text{ for all } 0 \leq |\alpha| \leq j \}.$$

Corollary 3.17. Let $k \in \mathbb{N}$ and $p \in [1, \infty)$. If there are $j \in \mathbb{N}_0$ and $q \in [p, \infty)$ with $k - \frac{d}{p} = j - \frac{d}{q}$, then

$$W^k_p(\mathbb{R}^d) \hookrightarrow W^j_q(\mathbb{R}^d).$$

Observe that Theorem 3.16a) is a special case with $j = 0$ and $q = p^*$.

Proof of Corollary 3.17. By assumption, we have $(k - j)p = d - \frac{dp}{q} \in (0,d)$. Theorem 3.16a) thus yields

$$W^{k-j}_p(\mathbb{R}^d) \hookrightarrow L^q(\mathbb{R}^d) \quad \text{since } q = \frac{pd}{d - kp + jp}.$$
Applying this embedding to $\partial^\alpha f \in W^{k-j}_p(\mathbb{R}^d)$ for all $|\alpha| \leq j$ and $f \in W^k_p(\mathbb{R}^d)$, we deduce that

$$\|\partial^\alpha f\|_q \leq c\|\partial^\alpha f\|_{k-j,p} \leq c\|f\|_{k,p},$$

as asserted. \qed

**Example 3.18.** There is an unbounded function $f \in W^1_d(\mathbb{R}^d)$ for $d \geq 2$, showing that Theorem 3.16b) is sharp. In fact, for any $\alpha \in (0, 1 - \frac{1}{d})$ and $\varphi \in C_c^\infty(\mathbb{R}^d)$ with $\text{supp } \varphi \subseteq B(0, 3/4)$ and $\varphi = 1$ on $B(0, 1/2)$, we define

$$f(x) := \begin{cases} \varphi(x)(-\ln|x|)^\alpha, & 0 < |x| \leq 3/4, \\ 0, & |x| > 3/4 \text{ or } x = 0. \end{cases}$$

Arguing as in Example 3.4c), one sees that $f \in L^p(\mathbb{R}^d)$ for all $p < \infty$, $f \notin L^\infty(\mathbb{R}^d)$ and that, for all $j \in \{1, \ldots, d\}$, we have

$$\partial_j f(x) = (\partial_j \varphi(x))(-\ln|x|)^\alpha - \alpha \varphi(x)(-\ln|x|)^{\alpha - 1} \frac{x_j}{|x|^2}$$

for $0 < |x| < 3/4$ and $\partial_j f(x) = 0$ otherwise. Using polar coordinates, we further estimate

$$\left(\int_{\mathbb{R}^d} |\partial_j f|^d dx\right)^{1/d} \leq c\|\partial_j \varphi\|_\infty \left(\int_0^{3/4} (\ln r)^{\alpha d} r^{d-1} dr\right)^{1/d} + c\|\varphi\|_\infty \left(\int_0^{3/4} (\ln r)^{\alpha - 1} d r^{d-1} dr\right)^{1/d}$$

$$\leq c + c \left(\int_0^{3/4} \frac{dr}{r(\ln r)^{(1-\alpha)d}}\right)^{1/d} < \infty$$

for some constants $c > 0$, since $(1 - \alpha)d > 1$. Hence, $f \in W^1_d(\mathbb{R}^d) \setminus L^\infty(\mathbb{R}^d)$.

For the proof of Theorem 3.16 we set $\hat{x}^j = (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_d) \in \mathbb{R}^{d-1}$ for all $x \in \mathbb{R}^d$, $j \in \{1, \ldots, d\}$ and $d \geq 2$. We start with a lemma.

**Lemma 3.19.** Let $d \geq 2$ and $f_1, \ldots, f_d \in L^{d-1}(\mathbb{R}^{d-1}) \cap C(\mathbb{R}^{d-1})$. Set $f(x) = f_1(\hat{x}^1) \cdot \ldots \cdot f_d(\hat{x}^d)$ for $x \in \mathbb{R}^d$. We then have $f \in L^1(\mathbb{R}^d)$ and

$$\|f\|_{L^1(\mathbb{R}^d)} \leq \|f_1\|_{L^{d-1}(\mathbb{R}^{d-1})} \cdot \ldots \cdot \|f_d\|_{L^{d-1}(\mathbb{R}^{d-1})}.$$

**Proof.** If $d = 2$, then Fubini’s theorem shows that

$$\int_{\mathbb{R}^2} |f(x)| dx = \int_{\mathbb{R}} \int_{\mathbb{R}} |f_1(x_2)| |f_2(x_1)| dx_1 dx_2 = \|f_1\|_1 \|f_2\|_1,$$

as asserted. Assume that the assertion holds for some $d \in \mathbb{N}$ with $d \geq 2$.

Take $f_1, \ldots, f_{d+1} \in L^d(\mathbb{R}^d) \cap C(\mathbb{R}^d)$. Write $y = (x_1, \ldots, x_d) \in \mathbb{R}^d$ and $x = (y, x_{d+1}) \in \mathbb{R}^{d+1}$. For a.e. $x_{d+1} \in \mathbb{R}$, the maps $\gamma^j \mapsto |f_j(\gamma^j, x_{d+1})|^d$ are integrable
3.2. DENSITY AND EMBEDDING THEOREMS

on \( \mathbb{R}^{d-1} \) for every \( j \in \{1, \ldots, d\} \) due to Fubini’s theorem. Fix such a \( x_{d+1} \in \mathbb{R} \) and write

\[
\hat{f}(y, x_{d+1}) := \prod_{j=1}^{d} f_j(\hat{x}^j).
\]

Using Hölder’s inequality and \( d' = \frac{d}{d-1} \), we obtain

\[
\int_{\mathbb{R}^d} |f(y, x_{d+1})|dy = \int_{\mathbb{R}^d} |\hat{f}(y, x_{d+1})|f_{d+1}(y)|dy
\leq \|f_{d+1}\|_{L^{d'}(\mathbb{R}^d)} \left( \int_{\mathbb{R}^d} |\hat{f}(y, x_{d+1})|^{d'} dy \right)^{1/d'}.
\]

We set \( g_j(\hat{y}^j) = |f_j(\hat{y}^j, x_{d+1})|^{d'} \) for \( j \in \{1, \ldots, d\} \) and \( x \in \mathbb{R}^{d+1} \). Since \( d'(d-1) = d \), we have \( g_j \in L^{d-1}(\mathbb{R}^{d-1}) \), and the induction hypothesis yields

\[
\int_{\mathbb{R}^d} |\hat{f}(y, x_{d+1})|^{d'} dy = \int_{\mathbb{R}^d} g_1(\hat{y}^1) \cdot \ldots \cdot g_d(\hat{y}^d)dy \leq \|g_1\|_{d-1} \cdot \ldots \cdot \|g_d\|_{d-1}
= \prod_{j=1}^{d} \left( \int_{\mathbb{R}^{d-1}} |f_j(\hat{x}^j, x_{d+1})|dx \right)^{\frac{1}{d}}.
\]

Integrating over \( x_{d+1} \in \mathbb{R} \), we thus arrive at

\[
\int_{\mathbb{R}^{d+1}} |f|dx \leq \|f_{d+1}\|_d \prod_{j=1}^{d} \left( \int_{\mathbb{R}^{d-1}} |f_j(\hat{x}^j)|^{d} dy \right)^{\frac{1}{d'}} dx_{d+1}.
\]

Applying the \( d \)-fold Hölder inequality to the \( x_{d+1} \)-integral, we conclude that

\[
\int_{\mathbb{R}^{d+1}} |f|dx \leq \|f_{d+1}\|_d \prod_{j=1}^{d} \left( \int_{\mathbb{R}^{d-1}} |f_j(\hat{x}^j)|^{d} dy \right)^{\frac{1}{d'}} dx_{d+1}
\]

\[
= \|f_1\|_d \cdot \ldots \cdot \|f_{d+1}\|_d. \quad \Box
\]

Recall from Analysis 3 that for \( f \in L^p(U) \cap L^q(U) \) and \( r \in [p, q] \) with \( 1 \leq p < q \leq \infty \), we have

\[
\|f\|_r \leq \|f\|_p^{\theta} \|f\|_q^{1-\theta} \leq \theta \|f\|_p + (1-\theta) \|f\|_q,
\]

where \( \theta \in [0, 1] \) is given by \( \frac{1}{r} = \frac{\theta}{p} + \frac{1-\theta}{q} \) and we also used Young’s inequality from Analysis 1 and 2.

Proof of Theorem 3.16. We only prove the case \( k = 1 \), the rest can be done by induction, see e.g. §5.6.3 in [Eva10]. Since \( W^{1}_p(\mathbb{R}^d) \hookrightarrow L^p(\mathbb{R}^d) \), estimate (3.11) implies that for assertion a) it suffices to show

\[
W^{1}_p(\mathbb{R}^d) \hookrightarrow L^\infty(\mathbb{R}^d).
\]
1) Let $f \in C^1_c(\mathbb{R}^d)$. Let first $p = 1 < d$, whence $p^* = \frac{d}{d-p}$. For $x \in \mathbb{R}^d$ and $j \in \{1, \ldots, d\}$, we then obtain

$$|f(x)| = \left| \int_{-\infty}^{x_j} \partial_j f(x_1, \ldots, x_{j-1}, t, x_{j+1}, \ldots, x_d) dt \right| \leq \int_{\mathbb{R}} |\partial_j f(x)| dx_j,$$

$$|f(x)|^d \leq \prod_{j=1}^{d} \int_{\mathbb{R}^d} |\partial_j f(x)| dx_j.$$

Setting $g_j(\hat{x}^j) = (\int_{\mathbb{R}} |\partial_j f(x)| dx_j)^{\frac{1}{d-1}}$, we deduce

$$|f(x)|^d \leq \prod_{j=1}^{d} g_j(\hat{x}^j).$$

After integration over $x \in \mathbb{R}^d$, Lemma 3.19 yields

$$\|f\|_{L^\frac{d}{d-p}(\mathbb{R}^d)} \leq \int_{\mathbb{R}^d} g_1(\hat{x}^1) \cdots g_d(\hat{x}^d) dx \leq \prod_{j=1}^{d} \|g_j\|_{L^{d-1}(\mathbb{R}^{d-1})}$$

$$= \prod_{j=1}^{d} \left( \int_{\mathbb{R}^{d-1}} \int_{\mathbb{R}} |\partial_j f(x)| dx_j d\hat{x}^j \right)^{\frac{1}{d-1}},$$

(3.12) $$\|f\|_{L^\frac{d}{d-p}(\mathbb{R}^d)} \leq \prod_{j=1}^{d} \|\partial_j f\|_{L^{d-1}(\mathbb{R}^d)} \leq \|\nabla f\|_1 \leq \|f\|_1.$$

2) Next, let $p \in (1, d)$ and $p^* = \frac{pd}{d-p}$. Set $t_* := \frac{d-1}{d}p^* = \frac{d-1}{d}p > 1$. An elementary calculation shows that $(t_* - 1)p' = t_* \frac{d}{d-p} = p^*$. Set

$$g = f|f|^{t-1} = f(f \overline{f})^{\frac{t-1}{2}}$$

for $t > 1$. We compute

$$\partial_j g = \partial_j f|f|^{t-1} + \frac{f - 1}{2} (f \overline{f})^{\frac{t-1}{2}} ((\partial_j f) \overline{f} + f(\partial_j \overline{f}))$$

$$= \partial_j f|f|^{t-1} + (t - 1)f |f|^{t-3} \text{Re}(f \partial_j \overline{f}),$$

$$|g| = |f|^t, \quad |\partial_j g| \leq t|\partial_j f| |f|^{t-1}.$$

Applying (3.12) to $g$, we estimate

$$\|f\|_{L^\frac{d}{d-p}(\mathbb{R}^d)} \leq \left( \int_{\mathbb{R}^d} |f|^t dx \right)^{\frac{d}{d-t}} = \left( \int_{\mathbb{R}^d} |g|^t dx \right)^{\frac{d}{d-t}}$$

$$\leq \prod_{j=1}^{d} \|\partial_j g\|_1 \leq \prod_{j=1}^{d} \left( \int_{\mathbb{R}^d} |\partial_j f| |f|^{t-1} dx \right)^{\frac{1}{2}}$$

$$\leq t \prod_{j=1}^{d} \left( \int_{\mathbb{R}^d} |\partial_j f| |f|^p dx \right)^{\frac{1}{2p}} \left( \int_{\mathbb{R}^d} |f|^{(t-1)p'} dx \right)^{\frac{1}{2p'}}.$$
Now, we can apply \( (\ast) \) Here and below the constant \( c > f \). By density (see Theorem 3.14), this estimate can be extended to all \( f \in W^1_p(\mathbb{R}^d) \). Then the identity map on \( W^1_p(\mathbb{R}^d) \) is the required embedding.

3) Now let \( f \in C^1_c(\mathbb{R}^d) \), \( p = d \), and \( t > 1 \). Then \( p' = \frac{d}{d-1} \), and Step 2 yields

\[
(\ast) \quad \| f \|_{L^p(\mathbb{R}^d)} \leq \left( \frac{t^\frac{1}{p} - 1}{\frac{d}{d-1} - 1} \right)^\frac{1}{d} \| \nabla f \|_{L^d(\mathbb{R}^d)} \leq c \left( \| f \|_{L^1(\mathbb{R}^d)} + \| \nabla f \|_{L^d(\mathbb{R}^d)} \right)
\]

using Young’s inequality from Analysis 1 and 2. For \( t = d \), this estimate gives \( f \in \mathcal{L}^{\frac{d^d}{d-1}}(\mathbb{R}^d) \) and

\[
\| f \|_{\frac{d^d}{d-1}} \leq c \| f \|_{1,d}.
\]

Here and below the constant \( c > 0 \) does not depend on \( f \). For \( q \in (d, d\frac{d}{d-1}) \), inequality (3.11) further yields

\[
\| f \|_q \leq c(\| f \|_d + \| f \|_{\frac{d^d}{d-1}}) \leq c \| f \|_{1,d}.
\]

Now, we can apply (\ast) with \( t = d + 1 \) and obtain

\[
\| f \|_{\frac{d^{d+1}}{d-1}} \leq c(\| f \|_{\frac{d^d}{d-1}} + \| \nabla f \|_{L^d(\mathbb{R}^d)}) \leq c \| f \|_{1,d}.
\]

As above, we see that \( f \in \mathcal{L}^q(\mathbb{R}^d) \) for \( d < q \leq d\frac{d+1}{d-1} \). We can now iterate this procedure with \( t_n = d + n \) and obtain

\[
\| f \|_q \leq c(q) \| f \|_{1,p}
\]

for all \( q < \infty \). Again, b) follows by density.

4) Let \( f \in C^1_c(\mathbb{R}^d) \), \( p > d \), \( Q(r) = [-\frac{r}{2}, \frac{r}{2}]^d \) for \( r > 0 \), and \( x_0 \in Q(r) \). Set \( M(r) = r^{-d} \int_{Q(r)} f \, dx \). We further put \( \beta := 1 - \frac{d}{p} \in (0, 1) \). Using \( |x - x_0|_\infty \leq r \) for \( x \in Q(r) \), the transformation \( y = t(x - x_0) \) and Hölder’s inequality, we compute

\[
|f(x_0) - M(r)| = \left| r^{-d} \int_{Q(r)} (f(x_0) - f(x)) \, dx \right|
\]

\[
= r^{-d} \int_{Q(r)} \int_0^t \frac{d}{dt} f(x_0 + t(x - x_0)) \, dt \, dx
\]

\[
\leq r^{-d} \int_{Q(r)} \int_0^1 |\nabla f(x_0 + t(x - x_0)) \cdot (x - x_0)| \, dt \, dx
\]
If \( f \) is a Cauchy sequence in \( \mathbb{R}^d \), then \( f\) converges to \( \tilde{f} \) in \( W^1_p(\mathbb{R}^d) \) and it has a representative \( \tilde{f} \). Hence, \( f \) is the required embedding.}

\[ \int_0^1 \int_{Q(\tau)} |\nabla f(x_0 + t(x - x_0))|_1 \, dx \, dt \]

\[ = r^{1-d} \int_0^1 \int_{t(Q(\tau) - x_0)} |\nabla f(x_0 + y)|_1 \, dy \, t^{-d} \, dt \]

\[ \leq r^{1-d} \left( \int_0^1 \left( \int_{t(Q(\tau) - x_0)} |\nabla f(x_0 + y)|_p \, dy \right)^{\frac{1}{p}} \right) \, \text{vol}(t(Q(\tau) - x_0))^{\frac{1}{p}} t^{-d} \, dt \]

\[ \leq cr^{1-d} ||\nabla f||_{L^p(\mathbb{R}^d)} \int_0^1 r^{\frac{d}{p}} t^{\frac{d}{p} - d} \, dt \]

\[ = Cr^{1-d} ||\nabla f||_{L^p(\mathbb{R}^d)} \]

for constants \( C, c > 0 \) only depending on \( d \) and \( p \), using also that \( \frac{d}{p} - d > -1 \) due to \( p > d \). A translation then gives

\[ |f(x_0 + z) - r^{-d} \int_{z+Q(\tau)} f(y) \, dy| \leq Cr^d ||\nabla f||_{L^p(\mathbb{R}^d)} \]

for all \( z \in \mathbb{R}^d \). Taking \( x = z \), \( x_0 = 0 \), \( r = 1 \), and using Hölder’s inequality, we thus obtain

\[ |f(x) - f(y)| \leq |f(x) - \int_{x+Q(1)} f \, dy| + |\int_{x+Q(1)} f \, dy| \leq C ||\nabla f||_{L^p} + ||f||_{L^p} \leq c ||f||_{1,p} \]

for all \( x \in \mathbb{R}^d \), where \( c \) only depends on \( d \) and \( p \). Given \( x, y \in \mathbb{R}^d \), we find a cube \( Q \) of side length \( |x - y|_\infty = r \) such that \( x, y \in Q \) and \( Q \) is parallel to the axes. Hence,

\[ |f(x) - f(y)| \leq |f(x) - r^{-d} \int_Q f \, dy| + |r^{-d} \int_Q f \, dy - f(y)| \leq 2C ||\nabla f||_{L^p} |x - y|_\infty^d \]

\[ \leq 2C ||\nabla f||_{L^p} |x - y|^d. \]

If \( f \in W^1_p(\mathbb{R}^d) \), then there are \( f_n \in C^1_c(\mathbb{R}^d) \) converging to \( f \) in \( W^1_p(\mathbb{R}^d) \). By (+), \( f_n \) is a Cauchy sequence in \( C_0(\mathbb{R}^d) \). Hence, \( f \) has a representative \( \tilde{f} \in C_0(\mathbb{R}^d) \) such that \( f_n \to \tilde{f} \) uniformly as \( n \to \infty \). So the above estimates imply that

\[ \| \tilde{f} \|_{1,p} + \sup_{x \neq y} \frac{|\tilde{f}(x) - \tilde{f}(y)|}{|x - y|_2^d} \leq c \| f \|_{1,p}. \]

The map \( f \mapsto \tilde{f} \) is the required embedding. \( \square \)

**Remark 3.20.** The assertions of Theorem 3.16 hold on an open set \( U \) instead of \( \mathbb{R}^d \) if we replace \( W^k_p(\mathbb{R}^d) \) by \( W^k_p(U) \). In fact, Theorem 3.16 can be applied to the 0-extension of \( \varphi \in C_0^\infty(U) \) and then the result follows by density.

**Problem:** How to extend the Sobolev embedding to \( W^k_p(U) \)?
3.2. DENSITY AND EMBEDDING THEOREMS

**Definition 3.21.** Let $k \in \mathbb{N}$. The open set $U \subseteq \mathbb{R}^d$ has the $k$-extension property if for all $m \in \{0,1,\ldots,k\}$ and $p \in [1,\infty)$, there is an operator

$$E_{m,p} \in \mathcal{L}(W^m_p(U), W^m_p(\mathbb{R}^d))$$

with $E_{m,p}f = f$ on $U$ and $E_{m,p} = E_{l,q}$ for all $f \in W^m_p(U) \cap W^l_q(U)$, $1 \leq m,l \leq k$ and $1 \leq p,q < \infty$. We write $E_U$ instead of $E_{m,p}$.

Observe that $E_0 = f$ for $x \in U$ and $E_0f(x) = 0$ for $x \in \mathbb{R}^d \setminus U$ defines an isometry $E_0 : L^p(U) \to L^p(\mathbb{R}^d)$. Moreover, $P_Uf = f|_U$ defines a contractive map in all spaces $\mathcal{L}(W^k_p(\mathbb{R}^d), W^k_p(U))$.

**Corollary 3.22.** If $U$ has the $m$-extension property for some $m \in \mathbb{N}$, then Theorem 3.16 and Corollary 3.17 hold for $k \in \{1,\ldots,m\}$ with $\mathbb{R}^d$ replaced by $U$ and $C^0_0(\mathbb{R}^d)$ replaced by

$$C^0_0(U) = \{f \in C^0(U) : \frac{\partial^\alpha f}{\partial x^\alpha} \text{ has a continuous extension to } \partial U \text{ and } \frac{\partial^\alpha f}{\partial x^\alpha}(x) \to 0 \text{ as } |x| \to \infty \text{ if } U \text{ is unbounded, for all } 0 \leq |\alpha| \leq j\}.$$

**Proof.** Consider e.g. Theorem 3.16a). We have the embedding

$$J : W^k_p(\mathbb{R}^d) \hookrightarrow L^p(\mathbb{R}^d)$$

given by the identity. Then,

$$R_UJE_U : W^k_p(U) \to L^p(U)$$

is continuous and injective. The other assertions are proved in the same way. \qed

**Corollary 3.23.** Let $U$ possess the $k$-extension property for some $k \in \mathbb{N}$. Then

$$W^k_p(U) \cap C^\infty(\overline{U}) \text{ is dense in } W^k_p(U)$$

for all $1 \leq p < \infty$.

**Proof.** If $f \in W^k_p(U)$, then $E_Uf \in W^k_p(\mathbb{R}^d)$. By Theorem 3.14, there are $g_n \in C^\infty_c(\mathbb{R}^d)$ converging to $E_Uf$ in $W^k_p(\mathbb{R}^d)$. Hence, $R_Ug_n \in W^k_p(U) \cap C^\infty(\overline{U})$ converge to $f = R_UE_Uf$ in $W^k_p(U)$ as $n \to \infty$. \qed

**Theorem 3.24.** Let $U \subseteq \mathbb{R}^d$ be bounded and open with $\partial U \subset C^k$. Then $U$ has the $k$-extension property.

**Sketch of the proof for $k = 1$.** (See also Theorem 5.22 in [AdF07].) 1) Let

$$H_\pm = \{(y,t) \in \mathbb{R}^d : y \in \mathbb{R}^{d-1}, t \geq 0\}$$

and $f \in W^1_p(H_-) \cap C^1(\overline{H}_-)$. Define

$$E_-f(y,r) = \begin{cases} f(y,t), & (y,t) \in \overline{H}_-, \\ 4f(y, -\frac{1}{2}) - 3f(y, -t), & (y,t) \in H_+. \end{cases}$$

47
CHAPTER 3. SOBOLEV SPACES AND WEAK DERIVATIVES

Note that \( E_- f \in C^1(\mathbb{R}^d) \). One can check that \( \|E_- f\|_{W^p_1(\mathbb{R}^d)} \leq c\|f\|_{W^p_1(H_-)} \) for a constant \( c > 0 \).

2) We show that \( W^1_p(H_-) \cap C^1(\overline{H}_-) \) is dense in \( W^1_p(H_-) \), so that \( E_- \) can be extended to a 1-extension operator on \( W^1_p(H_-) \). In fact, let \( f \in W^1_p(H_-) \) and \( \varepsilon > 0 \). By Theorem 3.14, there is a function \( g \in C^\infty(\overline{H}_- \cap W^1_p(H_-)) \) with \( \|f - g\|_{1,p} \leq \varepsilon \).

Setting \( g_n(y,t) = g(y,t - \frac{1}{n}) \) for \( t \leq 0 \), \( y \in \mathbb{R}^{d-1} \) and \( n \in \mathbb{N} \), we define the functions \( g_n \in C^1(\overline{H}_- \cap W^1_p(H_-)) \). Observe that

\[
\partial^\alpha g_n = R_{\overrightarrow{n}} S_n E_0 \partial^\alpha y
\]

for \( 0 \leq |\alpha| \leq 1 \), where \( S_n \in \mathcal{L}(L^p(\mathbb{R}^d)) \) is given by \( S_n h(y,t) = h(y,t - \frac{1}{n}) \) for \( h \in L^p(\mathbb{R}^d) \). One can see that \( S_n h \to h \) in \( L^p(\mathbb{R}^d) \) as in Example 3.8 of [FA]. Hence, \( g_n \) converges to \( g \) in \( W^1_p(H_-) \) implying the claim.

3) Since \( \partial U \in C^1 \), there are bounded open sets \( U_0, U_1, \ldots, U_m \subseteq \mathbb{R}^d \) such that \( U \subseteq U_0 \cup \cdots \cup U_m \) and \( \partial U \subseteq U_1 \cup \cdots \cup U_m \), as well as a diffeomorphism \( \Psi_j : U_j \to V_j \) such that \( \Psi_j' \) and \( (\Psi_j^{-1})' \) are bounded and \( \Psi_j(U_j \cap U) \subseteq H_- \) and \( \Psi_j'(U_j \cap \partial U) \subseteq \mathbb{R}^{d-1} \times \{0\} \), for each \( j \in \{1, \ldots, m\} \). Moreover, there are functions \( 0 \leq \varphi_j \in C^\infty(\mathbb{R}^d) \) with \( \text{supp} \varphi_j \subseteq U_j \) for all \( j = 0, 1, \ldots, m \) and \( \sum_{j=0}^m \varphi_j(x) = 1 \) for all \( x \in U \) (see Analysis 3).

Let \( j \in \{1, \ldots, m\} \). Set \( S_j g(y) = g(\Psi_j^{-1}(y)) \) for \( y \in H_- \cap V_j \) and \( S_j g(y) = 0 \) for \( y \in H_- \setminus V_j \), where \( g \in W^1_p(U_j \cap U) \). For \( h \in W^1_p(\mathbb{R}^d) \), set \( S_j h(x) = h(\Psi_j(x)) \) for \( x \in U_j \) and \( S_j h(x) = 0 \) for \( x \in \mathbb{R}^d \setminus U_j \). Take any \( \tilde{\varphi}_j \in C^\infty(\mathbb{R}^d) \) with \( \text{supp} \tilde{\varphi}_j \subseteq U_j \) and \( \tilde{\varphi}_j = 1 \) on \( \text{supp} \varphi_j \) (see Lemma 3.5).

Let \( f \in W^1_p(U) \). Define

\[
E_U f = E_0 \varphi_0 f + \sum_{j=1}^m \tilde{\varphi}_j S_j E_- S_j \left( R_{(U_j \cap U)} \varphi_j R_{(U_j \cap U)} f \right).
\]

Note that \( \text{supp} E_U f \subseteq U_0 \cup U_1 \cup \cdots \cup U_m \). Using part 2) and Proposition 3.7 and 3.8, we see that \( E_U \in \mathcal{L}(W^1_p(U), W^1_p(\mathbb{R}^d)) \). Let \( x \in U \). If \( x \in U_k \) for some \( k \in \{1, \ldots, m\} \), we have \( \Psi_k(x) \in H_- \). If \( x \notin U_j \), then \( \tilde{\varphi}_j(x) = 0 \). Thus

\[
E_U f(x) = \varphi_0(x)f(x) + \sum_{j=1, \ldots, m} \tilde{\varphi}_j(x)(\varphi_j f)(\Psi_j^{-1}(\Psi_j(x)));
\]

\[
= \sum_{j=0}^m \varphi_j(x)f(x) = f(x).
\]

If \( x \in U_0 \setminus (U_1 \cup \cdots \cup U_m) \), we also have \( E_U f(x) = \varphi_0 f(x) = f(x) \). \( \square \)

We refer to [AdF07] for other extension theorems and a detailed study of the impact of the regularity of \( \partial U \).

Theorem 3.25 (Rellich-Kondrachov). Let \( U \subseteq \mathbb{R}^d \) be bounded and open with \( \partial U \in C^k, k \in \mathbb{N} \), and \( 1 \leq p < \infty \). Then the following assertions hold.

48
3.2. DENSITY AND EMBEDDING THEOREMS

a) If $kp \leq d$ and $1 \leq q < p^*$, then the embedding

$$J : W_p^k(U) \hookrightarrow L_q(U)$$

is compact. (For instance, $q = p$.)

b) If $k - \frac{d}{p} > j \in \mathbb{N}_0$, then the embedding

$$J : W_p^k(U) \hookrightarrow C_j(U)$$

is compact.

Note that a compact embedding $J : Y \hookrightarrow X$ means that any bounded sequence $(y_n)$ in $Y$ has a subsequence such that $(Jy_n)_j$ converges in $X$. In Theorem 3.25a), $J$ is given by the identity, in b) it is given by choosing the representative in $C_j$.

Proof of Theorem 3.25. We prove the result only for $k = 1$ (and thus $j = 0$). Part b) follows from the Arzela-Ascoli theorem since Corollary 3.22 and Theorem 3.16 give constants $\beta, c > 0$ such that $|f(x) - f(y)| \leq c|x - y|^{\beta}$ and $|f(x)| \leq c$ for all $x, y \in U$ and $f \in W_p(U)$ with $\|f\|_1 \leq 1$, where $p > d$.

In the case $p < d$, take $f_n \in W_p^1(U)$ with $\|f_n\|_{1,p} \leq M$ for all $n \in \mathbb{N}$. Set $g_n = E_U f_n \in W_p^1(\mathbb{R}^d)$. From the proof of Theorem 3.24 we know that the support of $g_n$ belongs to a fixed open bounded set $\tilde{U} \subseteq \mathbb{R}^d$ containing $U$. Moreover, $\|g_n\|_{1,p} \leq \|E_U\|_p = M$ for all $n \in \mathbb{N}$. Take $q \in [1, p^*)$ and $\theta \in (0, 1]$ with $\frac{1}{q} = \frac{\theta}{1} + \frac{1-\theta}{p^*}$. Inequality (3.11) and Theorem 3.16 yield that

$$\|f_n - f_m\|_{L^q(U)} \leq \|g_n - g_m\|_{L^q(\tilde{U})} \leq \|g_n - g_m\|_{L^1(\tilde{U})}^{1-\theta} \|g_n - g_m\|_{L^{p^*}(\tilde{U})}^{\theta}$$

for all $n, m \in \mathbb{N}$. So it suffices to construct a subsequence of $g_n$ which converges in $L^1(\tilde{U})$.

For $x \in \tilde{U}$, $n \in \mathbb{N}$ and $\varepsilon > 0$, we compute

$$|g_n(x) - T_\varepsilon g_n(x)| = \left| \int_{\mathbb{R}^d} \Psi_\varepsilon(x-y)(g_n(x) - g_n(y))dy \right|$$

$$\leq \varepsilon^{-d} \int_{B(x,\varepsilon)} \Psi_\varepsilon^\frac{1}{\varepsilon}(x-y) \left| g_n(x) - g_n(\varepsilon y) \right| dy$$

$$= \int_{B(0,1)} \Psi(z) \left| g_n(x) - g_n(x - \varepsilon z) \right| dz$$

$$= \int_{B(0,1)} \Psi(z) \int_0^1 \frac{d}{dt} g_n(x - tz) dt dz$$

$$\leq \int_{B(0,1)} \Psi(z) \int_0^\varepsilon |\nabla g_n(x - tz) \cdot z| dt dz$$
where we have used the transformations $z = \frac{1}{t}(x - y)$ and $y = x - tz$, as well as Fubini’s theorem, Young’s inequality, (3.6) and $L^p(U) \hookrightarrow L^1(\tilde{U})$. We thus obtain

\[ \|g_n - T_\varepsilon g_n\|_{L^1(\tilde{U})} \leq c\lambda(\tilde{U})M_1\varepsilon =: C\varepsilon \]

for all $n \in \mathbb{N}$ and $\varepsilon > 0$. On the other hand, the definition of $T_\varepsilon g_n$ yields

\[ |T_\varepsilon g_n(x)| \leq \|\Psi_\varepsilon\|_{L^\infty}\|g_n\|_{L^1(\tilde{U})}, \quad \text{and} \quad |\nabla T_\varepsilon g_n(x)| \leq \|\nabla \Psi_\varepsilon\|_{L^\infty}\|g_n\|_{L^1(\tilde{U})} \]

for all $x \in \tilde{U}$ and $n \in \mathbb{N}$ and each fixed $\varepsilon > 0$. The Arzela-Ascoli theorem now implies that the set $F_\varepsilon := \{T_\varepsilon g_n : n \in \mathbb{N}\}$ is relatively compact in $C(\overline{U})$ for each $\varepsilon > 0$, and thus in $L^1(\tilde{U})$ since $C(\overline{U}) \hookrightarrow L^1(\tilde{U})$. Let $\delta > 0$ be given and fix $\varepsilon = \frac{\delta}{2\varepsilon}$. Then there are $n_1, \ldots, n_l \in \mathbb{N}$ such that

\[ F_\varepsilon \subseteq \bigcup_{j=1}^l B_{L^1(\tilde{U})}(T_\varepsilon g_{n_j}, \frac{\delta}{2}) =: \bigcup_{j=1}^l B_j. \]

Hence, given $n \in \mathbb{N}$, there is a $n_j$ such that $T_\varepsilon g_n \in B_j$. The estimates (*) and (3.6) then yield

\[ \|g_n - T_\varepsilon g_{n_j}\|_{L^1(\tilde{U})} \leq \|g_n - T_\varepsilon g_n\|_{L^1(\tilde{U})} + \|T_\varepsilon (g_n - g_{n_j})\|_{L^1(\tilde{U})} \leq C\varepsilon + \delta/2 = \delta. \]

We have shown that, for each $\delta > 0$, the set $G := \{g_n : n \in \mathbb{N}\}$ is covered by finitely many open balls $B_j$ of radius $\delta/2$, i.e., $G$ is totally bounded in $L^1(\tilde{U})$. Thus $G$ contains a subsequence converging in $L^1(\tilde{U})$ (see Satz B.1.7 in [Wer05]). In the case $p = d$ one simply replaces $p^*$ by any $r \in (q, \infty)$.

**Remark 3.26.** a) Theorem 3.25 is wrong for unbounded domains. In fact, let $k \in \mathbb{N}$, $p \in [1, \infty)$ and define the functions $f_n = f(\cdot - n)$ in $W^k_p(\mathbb{R})$ for any function $0 \neq f \in C^\infty(\mathbb{R})$ such that $\text{supp}\ f \subseteq (-1/2, 1/2)$. Then $\|f_n\|_{L^p}$ and $\|f_n - f_m\|_q > 0$ do not depend on $n \neq m$ in $\mathbb{N}$ so that $(f_n)$ is bounded in $W^k_p(\mathbb{R})$ and has no subsequence which converges in $L^q$ with $1 \leq q < p^*$.

b) The embedding $W^1_p(U) \hookrightarrow L^p(U)$ is never compact, see Example 6.12 in [Adf07].

**Proposition 3.27.** Let $1 \leq p < \infty$ and either $f \in \dot{W}^2_p(U)$ or $U$ be bounded with $\partial U \subset C^2$ and $f \in W^2_p(U)$. Then there are constants $C, \varepsilon_0 > 0$ such that

\[ \left( \sum_{j=1}^d \|\partial_j f\|_p^p \right)^{1/p} \leq \varepsilon \left( \sum_{i,j=1}^d \|\partial_{ij} f\|_p^p \right)^{1/p} + \frac{C}{\varepsilon} \|f\|_p, \]

50
for all $\varepsilon > 0$ if $f \in \dot{W}^2_p(U)$ and for all $0 < \varepsilon \leq \varepsilon_0$ if $f \in \dot{W}^2_p(U)$.

**Proof.** Let $f \in C^2_\infty(U)$ and extend it to $\mathbb{R}^d$ by 0. Take $j = 1$. Write $x = (t, y) \in \mathbb{R} \times \mathbb{R}^{d-1}$ for $x \in \mathbb{R}^d$. Fix $y \in \mathbb{R}^{d-1}$ and set $g(t) = f(t, y)$ for $t \in \mathbb{R}$. Let $\varepsilon > 0$ and $a, b \in \mathbb{R}$ with $b - a = \varepsilon$. Take any $r \in (a, a + \frac{\varepsilon}{3})$ and $t \in (b - \frac{\varepsilon}{3}, b)$. There there is a $\sigma = \sigma(r, t) \in (a, b)$ such that

$$|g'(\sigma)| = \left| \frac{g(t) - g(r)}{t - r} \right| \leq \frac{3}{\varepsilon} (|g(t)| + |g(r)|).$$

For every $s \in (a, b)$ we thus obtain

$$|g'(s)| = |g'(\sigma) + \int_{\sigma}^{s} g''(\tau) d\tau|$$

$$\leq \frac{3}{\varepsilon} (|g(r)| + |g(t)|) + \int_{a}^{b} |g''(\tau)| d\tau.$$ 

Integrating first over $r$ and then over $t$, we conclude

$$\frac{\varepsilon}{3} |g'(s)| \leq \frac{3}{\varepsilon} \int_{a}^{a + \frac{\varepsilon}{3}} |g(r)| dr + |g(t)| + \frac{\varepsilon}{3} \int_{a}^{b} |g''(\tau)| d\tau,$$

$$\frac{\varepsilon^2}{9} |g'(s)| \leq \int_{a}^{a + \frac{\varepsilon}{3}} |g(r)| dr + \int_{b - \frac{\varepsilon}{3}}^{b} |g(t)| dt + \frac{\varepsilon^2}{9} \int_{a}^{b} |g''(\tau)| d\tau,$$

$$|g'(s)| \leq \frac{9}{\varepsilon^2} \int_{a}^{b} |g(\tau)| d\tau + \int_{a}^{b} |g''(\tau)| d\tau$$

$$\leq \frac{\varepsilon^p}{2} \left( \frac{9}{\varepsilon^p} \int_{a}^{b} |g(\tau)|^p d\tau \right)^{1/p} + \varepsilon^{\frac{p}{q}} \left( \int_{a}^{b} |g''(\tau)|^p d\tau \right)^{1/p},$$

where we used Hölder’s inequality first for the integrals and then in $\mathbb{R}^2$. We take now the $p$-th power and then integrate over $s$ arriving at

$$\int_{a}^{b} |g'(s)|^p ds \leq \frac{9^p}{\varepsilon^2p} \int_{a}^{b} |g(\tau)|^p d\tau + \int_{a}^{b} |g''(\tau)|^p d\tau.$$ 

Now choose $a = a_k = k\varepsilon$ and $b = b_k = (k + 1)\varepsilon$ for $k \in \mathbb{Z}$. Summing the integrals on $[k\varepsilon,(k + 1)\varepsilon)$ for $k \in \mathbb{Z}$ and then integrating over $y \in \mathbb{R}^{d-1}$, it follows that

$$\int_{\mathbb{R}} |g'(\tau)|^p d\tau \leq \varepsilon^{p-2}\left( \frac{9^p}{\varepsilon^2p} \int_{\mathbb{R}} |g(\tau)|^p d\tau + \int_{\mathbb{R}} |g''(\tau)|^p d\tau \right),$$

(3.14) $$\int_{U} |\partial_1 f|^p dx \leq (2\varepsilon)^p \int_{U} |\partial_1 f|^p dx + \frac{36^p}{(2\varepsilon)^p} \int_{U} |f|^p dx.$$ 

By approximation, (3.14) can be established for all $f \in \dot{W}^2_p(U)$. The same result holds for $\partial_j f$ and $\partial_{jj} f$ with $j \in \{2, \ldots, d\}$. We now replace $2\varepsilon$ by $\varepsilon$, sum over $j$ and
take the $p$-th root to arrive at
\[
\left( \sum_{j=1}^{d} \| \partial_j f \|_p^p \right)^{1/p} \leq \varepsilon^p \left( \sum_{j=1}^{d} \| \partial j f \|_p^p + \frac{36p}{\varepsilon^p} \| f \|_p^p \right)^{1/p}
\]
(3.15)
\[
\leq \varepsilon \left( \sum_{j=1}^{d} \| \partial j f \|_p^p \right)^{1/p} + \frac{36}{\varepsilon} \| f \|_p,
\]
for all $f \in W^2_p(U)$, as asserted. If $u \in W^2_p(U)$ and $U$ is bounded with $\partial U \in C^2$, we use the extension operator $E_U \in L(W^2_p(U), W^2_p(\mathbb{R}^d))$ from Theorem 3.24 to deduce from (3.15) with $U = \mathbb{R}^d$ that
\[
\left( \sum_{j=1}^{d} \| \partial_j f \|_{L_p(U)}^p \right)^{1/p} \leq \varepsilon \left( \sum_{j=1}^{d} \| \partial j E_U f \|_{L_p(\mathbb{R}^d)}^p \right)^{1/p}
\]
\[
\leq \varepsilon \left( \sum_{j=1}^{d} \| \partial j E_U f \|_{L_p(\mathbb{R}^d)}^p \right)^{1/p} + \frac{36}{\varepsilon} \| E_U f \|_{L_p(\mathbb{R}^d)}
\]
\[
\leq \varepsilon \| E_U f \|_{W^2_p(\mathbb{R}^d)} + \frac{36}{\varepsilon} \| E_U f \|_{L_p(U)}
\]
\[
\leq c \| f \|_{W^2_p(U)} + \frac{c}{\varepsilon} \| f \|_{L_p(U)}
\]
\[
\leq c_0 \varepsilon \left( \sum_{i,j=1}^{d} \| \partial_{ij} f \|_p^p \right)^{1/p} + c_1 \varepsilon \left( \sum_{j=1}^{d} \| \partial j f \|_p^p \right)^{1/p} + \frac{c}{\varepsilon} \| f \|_p
\]
where we assume that $\varepsilon \in (0, 1]$ and the constants $c, c_0, c_1$ do not depend on $\varepsilon$ or $f$. Choosing $\varepsilon_1 = \min\left\{ \frac{1}{2c_1}, 1 \right\}$ we arrive at
\[
\frac{1}{2} \left( \sum_{j=1}^{d} \| \partial j f \|_p^p \right)^{1/p} \leq c_0 \varepsilon \left( \sum_{i,j=1}^{d} \| \partial_{ij} f \|_p^p \right)^{1/p} + \frac{c}{\varepsilon} \| f \|_p
\]
if $0 < \varepsilon \leq \varepsilon_1$. This inequality implies the assertion for $U$ with $\partial U \in C^2$ after replacing $\varepsilon$ by $\varepsilon/(2c_0)$ and $\varepsilon_1$ by $\varepsilon_0 = c_0/c_1$. \qed
Chapter 4

Selfadjoint operators

4.1 Closed operators and their spectra, revisited

Example 4.1. Let $J \subseteq \mathbb{R}$ be an open interval, $X = L^p(J)$, $1 \leq p < \infty$, and $A_k u = \partial^k u$ with $D(A_k) = W^k_p(J)$ for $k = 1, 2$. By Exercise 11, $A_1$ is closed and $A_2$ is closed on the domain $D_2 = \{ u \in X : \exists \partial^2 u \in X \}$. Exercise 16 further yields $D(A_2) = D_2$.

Example 4.2. Let $1 \leq p < \infty$, $X = L^p(\mathbb{R})$ and $A = \partial$ with $D(A) = W^1_p(\mathbb{R})$. For $\text{Re} \lambda > 0$ we set $(R_\lambda f)(t) = \int_t^\infty e^{\lambda(t-s)} f(s) ds$, for $t \in \mathbb{R}$ and $f \in X$, see Example 1.21. Let $\varphi_\lambda = e^{\text{Re} \lambda} \mathbb{1}_{\mathbb{R}_-}$. Since

$$\left| R_\lambda (f(t)) \right| \leq \int_t^\infty e^{\text{Re} \lambda(t-s)} |f(s)| ds = (\varphi_\lambda * |f|)(t)$$

for all $t \in \mathbb{R}$, we deduce from Young’s inequality (see e.g. formula (3.1) in [FA]) that

$$\| R_\lambda f \|_p \leq \| \varphi_\lambda \|_1 \| f \|_p = \frac{1}{\text{Re} \lambda} \| f \|_p.$$  

Moreover, a slight variant of Theorem 3.10 shows that $R_\lambda f \in W^1(\mathbb{R})$ with $\partial R_\lambda f = \lambda R_\lambda f - f \in L^p(\mathbb{R})$. As a result, $R_\lambda f \in W^1_p(\mathbb{R}) = D(A) \hookrightarrow C_0(\mathbb{R})$ and $(\lambda I - A) R_\lambda = I$. Next, let $\lambda \in \mathbb{C}$ and suppose that there would exist some $u \in D(A)$ and $\lambda u = Au = \partial u$. Then Theorem 3.10 yields that

$$\frac{1}{t-s} (u(t) - u(s)) = \frac{1}{t-s} \int_s^t \partial u(\tau) d\tau = \lambda \frac{1}{t-s} \int_s^t u(\tau) d\tau \longrightarrow \lambda u(s)$$

as $t \to s$, i.e., $u$ is differentiable with the continuous derivative $\lambda u$. (Observe that the above argument shows that if $u \in W^1(\mathbb{R})$ and $\partial u$ is continuous, then $u \in C^1(\mathbb{R})$ for any open interval $J \subseteq \mathbb{R}$.) Consequently, $u$ is equal to a multiple of $e_\lambda$. Since $e_\lambda \notin X$, we obtain $u = 0$ and so $\lambda I - A$ is injective. Summing up, we have shown
that $\sigma_p(A) = \emptyset$ and $\lambda \in \rho(A)$ with $R(\lambda, A) = R_\lambda$ if $\text{Re} \lambda > 0$. In the same way, one sees that $\lambda \in \rho(A)$ if $\text{Re} \lambda < 0$, where the resolvent is given by

$$R(\lambda, A)f(t) = -\int_{-\infty}^{t} e^{\lambda(t-s)} f(s) ds,$$

for $t \in \mathbb{R}$ and $f \in X$. Now, let $\lambda \in i\mathbb{R}$. We take $\varphi_n \in C_c^1(\mathbb{R})$ with $\varphi_n = n^{-1/p}$ on $[-n, n]$, $\varphi_n(t) = 0$ for $|t| \geq n + 1$ and $\|\varphi_n\|_\infty \leq 2n^{-1/p}$, for every $n \in \mathbb{N}$. We then obtain $u_n = \varphi_n e_\lambda \in W_p^1(\mathbb{R})$,

$$\|u_n\|_p \geq \left( \int_{-n}^{n} |\varphi_n(t)e_\lambda(t)|^p dt \right)^{1/p} = n^{-1/p}(2n)^{1/p} = 2^{1/p},$$

$$\|\lambda u_n - Au_n\|_p = \|\varphi_n' e_\lambda\|_p = \left( \int_{\{|n| \leq n+1\}} |\varphi_n'(t)e_\lambda(t)|^p dt \right)^{1/p} \leq 2n^{-1/p}2^{1/p},$$

for every $n \in \mathbb{N}$. As a result, $\lambda \in \sigma_{sp}(A)$ and so $\sigma(A) = i\mathbb{R}$.

**Example 4.3.** Let $X = L_p(\mathbb{R})$, $1 \leq p < \infty$, and $B = \partial^2$ with $D(B) = W_p^2(\mathbb{R})$. For $\mu \notin \mathbb{R}$ there is a $\lambda \in \mathbb{C}$ with $\text{Re} \lambda > 0$ such that $\lambda^2 = \mu$. Using $A$ from Example 4.2, we obtain

$$(\mu I - B)u = (\lambda I - A)(\lambda I + A)u$$

for $u \in W_p^2(\mathbb{R})$. Since $\lambda I - A$ and $\lambda I + A = -(-\lambda I - A)$ are invertible by Example 4.2, we infer that

$$u = (\lambda I + A)^{-1}(\lambda I - A)^{-1}(\mu I - B)u$$

for all $u \in D(B)$. On the other hand, for $v \in W_p^1(\mathbb{R}^d)$ it holds that

$$\partial(\lambda I + A)^{-1}v = A(\lambda I + A)^{-1}v = -\lambda(\lambda I + A)^{-1}v + v \in W_p^1(\mathbb{R}).$$

This fact implies that $(\lambda I + A)^{-1}W_p^1(\mathbb{R}) \subseteq W_p^2(\mathbb{R})$. Therefore we can choose $u = (\lambda I + A)^{-1}(\lambda I - A)^{-1}f$ for any $f \in X$ in $(\ast)$, arriving at

$$(\mu I - B)(\lambda I + A)^{-1}(\lambda I - A)^{-1}f = f$$

for all $f \in X$. Hence, $\mu \in \rho(B)$ and $R(\mu, B) = (\lambda I + A)^{-1}(\lambda I - A)^{-1}$, so that $\sigma(B) \subseteq \mathbb{R}_-$. For $\mu \leq 0$, we have $\mu = \lambda^2$ for some $\lambda \in i\mathbb{R}$. Then $\lambda I - A$ is not surjective (since its range is not closed by Example 4.2). Equation $(\ast)$ thus implies that $\mu I - B$ is not surjective. As a result, $\sigma(B) = \mathbb{R}_- = \sigma(A^2)$.

**Definition 4.4.** A closed operator $A$ in $X$ has a compact resolvent if there exists a $\lambda \in \rho(A)$ such that $R(\lambda, A) \in \mathcal{L}(X)$ is compact.
4.1. CLOSED OPERATORS AND THEIR SPECTRA, REVISITED

Remark 4.5. a) If $A$ has a compact resolvent $R(\lambda, A)$ and $\mu \in \rho(A)$, then

$$R(\mu, A) = R(\lambda, A) + (\lambda - \mu)R(\lambda, A)R(\mu, A)$$

by the resolvent equation, so that $R(\lambda, A)$ is also compact due to Proposition 2.3.

b) For a closed operator $A$ with $\lambda \in \rho(A)$ the following assertions are equivalent, where $[D(A)] = (D(A), \|\cdot\|_A)$.

1) $A$ has a compact resolvent,

2) Each bounded sequence in $[D(A)]$ has a subsequence which converges in $X$,

3) The inclusion map $J : [D(A)] \rightarrow X$ is compact.

Proof. “1) $\Rightarrow$ 2)”: Let $x_n \in D(A)$ with $\|x_n\|_A \leq c$ for all $n \in \mathbb{N}$. Set $y_n = \lambda x_n - Ax_n$. Then $\|y_n\| \leq (|\lambda| + 1)c$ for every $n \in \mathbb{N}$ so that $x_n = R(\lambda, A)y_n$ has a subsequence which converges in $X$.

“2) $\Rightarrow$ 3)” : This implication is clear, see Remark 2.2.

“3) $\Rightarrow$ 1)” : Define $R_\lambda \in \mathcal{L}(X, [D(A)])$ by $R_\lambda x = R(\lambda, A)x$ for $x \in X$. The operator $R(\lambda, A) = JR_\lambda : X \rightarrow X$ is then compact due to Proposition 2.3. \[\square\]

c) If $\dim X = \infty$, then a closed operator $A$ with compact resolvent $R(\lambda, A)$ cannot be bounded since otherwise $\lambda I - A$ were bounded and thus $I = (\lambda I - A)R(\lambda, A)$ were compact by Proposition 2.3.

Theorem 4.6. Let $\dim X = \infty$ and $A$ be a closed operator with compact resolvent. Then $\sigma(A)$ is either empty or contains only at most countably many eigenvalues $\lambda_n$ with $\dim N(\lambda_n I - A) < \infty$. If $A$ has infinitely many eigenvalues $\lambda_n$, then $|\lambda_n| \rightarrow \infty$ as $n \rightarrow \infty$.

Proof. Fix a $\mu \in \rho(A)$. By Theorem 2.11, the spectrum $\sigma(R(\mu, A))$ only contains 0 and either no or finitely many eigenvalues $\mu_n$ or a nullsequence of eigenvalues $\mu_n$. Moreover, $\dim N(\mu_n I - R(\mu, A)) < \infty$ for all $n$.

By Proposition 1.20, we have $(\mu - \sigma(A))^{-1} = \sigma(\mu, A)) \setminus \{0\}$ and $(\mu - \sigma_p(A))^{-1} = \sigma_p(R(\mu, A)) \setminus \{0\}$. As a result, the spectrum of $A$ consists of only the eigenvalues $\lambda_n := \mu - \frac{1}{\mu_n}$ (if $\sigma(R(\mu, A)) \setminus \{0\} \neq \emptyset$). If there are infinitely many $\mu_n$, then $\mu_n \rightarrow 0$ and hence $|\lambda_n| \rightarrow \infty$ as $n \rightarrow \infty$. Since

$$\lambda_n I - A = (\mu_n I - R(\mu, A))\frac{1}{\mu_n}(\mu I - A),$$

linearly independent eigenvectors of $\lambda_n I - A$ give linearly independent eigenvectors $\frac{1}{\mu_n}(\mu I - A)v$ of $\mu_n I - R(\mu, A)$, so that the kernel of $\lambda_n I - A$ is finite dimensional. \[\square\]

Example 4.7. Let $U \subseteq \mathbb{R}^d$ be open and bounded with $\partial U \in C^1$ if $1 \leq p < \infty$. Let $X = L^p(U)$ for $p \in [1, \infty)$ and $X = C(\overline{U})$ for $p = \infty$. Let $A$ be a closed operator such that $[D(A)] \hookrightarrow W^1_p(U)$ if $1 \leq p < \infty$ and $[D(A)] \hookrightarrow C^1(\overline{U})$ if $p = \infty$. Theorem
3.25 of Rellich-Kondrachov and the theorem of Arzela-Ascoli (see e.g. Theorem 1.45 in [FA]) show that $W^1_p(U)$ and $C^1(\overline{U})$ are compactly embedded in $X$, respectively, so that the embedding $[D(A)] \hookrightarrow X$ is compact as a product of a bounded and a compact operator. By Remark 4.5, $A$ has a compact resolvent in $X$ if $\rho(A) \neq 0$.

We stress that for $L^p(0,1)$ and $C([0,1])$ the operators $A_k = \partial^k$ with domains $W^{k}_p(0,1)$ and $C^{k}([0,1])$, respectively, for $k = 1, 2$ have empty resolvent sets since $e_\lambda \in D(A_k)$ is an eigenfunction for the eigenvalue $\lambda$ if $k = 1$ and $\lambda^2$ if $k = 2$, for every $\lambda \in \mathbb{C}$.

To study boundary values of functions in $W^1_p(U)$, we need the following result proved e.g. in §5.5 of [Eva10].

**Theorem 4.8** (Trace theorem). Let $U \subseteq \mathbb{R}^d$ be bounded and open with $\partial U \in C^1$, and let $p \in [1, \infty)$. Then the trace map $f \mapsto f|_{\partial U}$ from $W^1_p(U) \cap C(\overline{U})$ to $L^p(\partial U, d\sigma)$ has a bounded linear extension $T : W^1_p(U) \rightarrow L^p(\partial U, d\sigma)$ whose kernel is $W^1_p(U)$.

**Proof.** 1) Let $u \in C^1(\overline{U})$. By the definition of the surface integral, see Analysis 3, there are finitely many diffeomorphisms $\Psi_j : \tilde{U}_j \rightarrow \tilde{V}_j$ and $\varphi_j \in C^1(\tilde{U}_j)$ with $0 \leq \varphi_j \leq 1$ such that $\|u\|_{L^p(\partial U)}$ is dominated by

$$c \sum_{j=1}^{m} \int_{V_j} \varphi_j \circ \Psi_j^{-1} |u \circ \Psi_j^{-1}|^p dy'$$

where $\tilde{U}_j$ and $\tilde{V}_j$ are open subsets of $\mathbb{R}^d$, the sets $\tilde{U}_j$ cover $\partial U$, the functions $\varphi_j$ form a partition of unity subordinated to $\tilde{U}_j$, $V_j := \{(y', y_d) \in \tilde{V}_j : y_d = 0\}$, $V_{j+} := \{(y', y_d) \in \tilde{V}_j : y_d > 0\}$, $\Psi_j(\tilde{U}_j \cap \partial U) = V_j$, and $\Psi_j(\tilde{U}_j \cap U) = V_{j+}$. We set $v = u \circ \Psi_j^{-1}$ and $\psi = \varphi_j \circ \Psi_j^{-1} \in C^1(\tilde{V}_j)$ and drop the indices $j$ below. By means of Fubini’s theorem and the fundamental theorem of calculus, we compute

$$\int_{V} \psi |v(y')|^p dy' = - \int_{V_{+}} \partial_d(\psi |v|^p) dy = - \int_{V_{+}} [(\partial_d\psi) |v|^p + py |v|^{p-1} \Re(\overline{\partial_dv})] dy \
\leq c \int_{V_{+}} [||v||^p + |v|^{p-1} ||\partial_dv||] dy \leq c ||v||_p^p + c ||v||_p^{p-1} ||\partial_dv||_p \
\leq c (||v||_p^p + ||\partial_dv||_p^p) \leq c ||v||_{W^1_p(U)}.$$

Here we also used Hölder’s and Young’s inequality, Proposition 3.8 and the transformation rule. As a result, the map $T : (C^1(\overline{U}), \| \cdot \|_1, p) \rightarrow L^p(\partial U, d\sigma)$, $Tu = u|\partial U$, is continuous. Corollary 3.23 and Theorem 3.24 allow to extend $T$ to an operator in $\mathcal{L}(W^1_p(U), L^p(\partial U, d\sigma))$. If we start with a $u \in W^1_p(U) \cap C(\overline{U})$, then we can define an approximating sequence of $u_n \in C^1(\overline{U})$ which converge to $u$ in $W^1_p(U)$ and in $C(\overline{U})$, see the proof of Theorem 5.3.3 in [Eva10]. Hence, $Tu_n = u_n|\partial U$ tends to $u|\partial U$ uniformly on $\partial U$ and to $Tu$ in $L^p(\partial U, d\sigma)$, so that $Tu = u|\partial U$.

2a) We next observe that the inclusion $W^1_p(U) \subseteq N(T)$ is a consequence of the continuity of $T$ since $T$ vanishes on $C^\infty_c(U)$ and this space is dense in $W^1_p(U)$ by
4.1. CLOSED OPERATORS AND THEIR SPECTRA, REVISITED

definition. To prove the converse, we start with the model case that \( v \in W_1^p(V_+) \) has a compact support in \( V_+ \) and \( Tv = 0 \). Our density results yield \( v_n \in C^1(V_+) \) converging to \( v \) in \( W_1^p(V_+) \), and hence \( Tv_n = v_n |V \to 0 \) in \( L^p(V) \), as \( n \to \infty \). Observe that

\[
|v_n(y', y_d)| \leq |v_n(y', 0)| + \int_0^{y_d} |\partial_d v_n(y', s)| \, ds,
\]

\[
|v_n(y', y_d)|^p \leq 2 |v_n(y', 0)|^p + 2 \left( \int_0^{y_d} |\partial_d v_n(y', s)| \, ds \right)^p
\]

for \( y' \in V \) and \( y_d \geq 0 \). Integrating over \( y' \) und employing H"older’s inequality, we obtain

\[
\int_V |v_n(y', y_d)|^p \, dy' \leq 2 \int_V |v_n(y', 0)|^p \, dy' + 2x_d^{p-1} \int_V \int_0^{y_d} |\partial_d v_n(y', s)|^p \, ds \, dy'.
\]

We can now let \( n \to \infty \) and arrive at

\[(4.1) \quad \int_V |v(y', y_d)|^p \, dy' \leq 2x_d^{p-1} \int_V \int_0^{y_d} |\partial_d v(y', s)|^p \, ds \, dy'.\]

We next use a cutoff argument to obtain a support in the interior of \( V_+ \). Choose a function \( \chi \in C^\infty(\mathbb{R}_+) \) such that \( \chi = 0 \) on \([0, 1]\) and \( \chi = 1 \) on \([2, \infty)\). Set \( \chi_n(s) = \chi(ns) \) for \( s \geq 0 \) and \( n \in \mathbb{N} \), and define \( w_n = \chi_n v \) on \( V_+ \). Observe that \( w_n \to v \) in \( L^p(V_+) \) as \( n \to \infty \). It holds \( \partial_j w_n = \chi_n \partial_j v \) for \( j = 1, \ldots, d-1 \) and \( \partial_d w_n = \chi_n \partial_d v + n \chi'(n \cdot) v \). Estimate (4.1) further implies

\[
\int_{V_+} |\nabla w_n - \nabla v|^p \, dy \leq c \int_0^{2/n} \int_V |1 - \chi_n|^p |\nabla v|^p \, dy' \, ds + cn^p \int_0^{2/n} \int_V |v(y, s)|^p \, dy' \, ds
\]

\[
\leq c \int_0^{2/n} \int_V |\nabla v|^p \, dy' \, ds + cn^p \int_0^{2/n} s^{p-1} \int_0^s \int_V |\partial_d v(y', \tau)|^p \, dy' \, d\tau \, ds
\]

\[
\leq c \int_0^{2/n} \int_V |\nabla v|^p \, dy' \, ds + c \int_0^{2/n} \int_V |\partial_d v(y', \tau)|^p \, dy' \, d\tau
\]

for some constants \( c > 0 \). Because of \( v \in W_1^p(V_+) \), the above integrals tend to 0 as \( n \to \infty \), and so \( w_n \to v \) in \( W_1^p(V_+) \) as \( n \to \infty \). Since \( w_n = 0 \) for \( y_d \in (0, 1/n] \), we can mollify \( w_n \) to obtain a function \( \tilde{w}_n \in C^\infty_c(V_+) \) such that \( \|\tilde{w}_n - w_n\|_{1,p} \leq 1/n \). This means that \( \tilde{w}_n \to v \) in \( W_1^p(V_+) \) as \( n \to \infty \).

2b) We come back to \( u \in W_1^p(U) \) and consider the sets \( \tilde{U}_j \) and \( \tilde{V}_j \) and the functions \( \Psi_j \) and \( \varphi_j \) from step 1). Let \( v_j = (\varphi_j u) \circ (\Psi^{-1}_j) \). First, observe that the trace of \( v_j \) to the set \( V_j \) is given by \( (Tu) \circ \Psi^{-1}_j \) if \( u \in C(\overline{U}) \), in addition. By continuity one can extend this identity to all \( u \in W_1^p(U) \). Let \( Tu = 0 \). Then we can apply part 2a) to \( v_j \) and obtain \( \tilde{w}_j \in C^\infty_c(V_{j^+}) \) converging to \( v_j \) in \( W_1^p(V_{j^+}) \). The function

\[
\tilde{w}_n = \sum_{j=1}^m \tilde{w}_j \circ (\Psi_j | U \cap \tilde{U}_j)
\]
thus belongs to \( C^1_c(U) \) and converges to \( u \) in \( W^1_1(U) \) as \( n \to \infty \). Since \( \hat{u}_n \) has compact support, we can mollify \( \hat{u}_n \) to a function \( u_n \in C^\infty_c(U) \) with \( \| \hat{u}_n - u_n \|_{1,p} \leq 1/n \). This means that \( u_n \to u \) in \( W^1_1(U) \) as \( n \to \infty \), and hence \( u \in W^1_1(U) \).

We can thus say that \( u \in \dot{W}^1_1(U) \) is 0 at the boundary in the “sense of trace”. Observe that for a Borel set \( \Gamma \in \partial U \) we can define the trace of \( u \in W^1_1(U) \) on \( \Gamma \) by \( T_\Gamma u = R_\Gamma T u \), where \( R_\Gamma : L^p(\partial U, d\sigma) \to L^p(\Gamma, d\sigma) \) is the contractive restriction map. Finally, for \( p > d \) we have \( W^1_1(U) \hookrightarrow C(\overline{U}) \) by Theorem 3.24 and Corollary 3.22 so that Theorem 4.8 is trivial in this case. Theorem 3.10 further implies that \( W^1_1(a, b) \hookrightarrow C([a, b]) \) for \( d = 1 \).

**Example 4.9.** Let \( X = L^p(0, 1), 1 \leq p < \infty \), and \( A = \partial \) with

\[
D(A) = \{ u \in W^1_1(0, 1) : u(0) = 0 \}.
\]

If \( u \in W^1_1(0, 1) \) and \( \lambda u = \partial u \) for some \( \lambda \in \mathbb{C} \), then \( u = e_\lambda \) (cf. Example 4.2). Since \( e_\lambda \not\in D(A) \), we obtain \( \sigma_p(A) = \emptyset \). Further, the Volterra operator \( V \) given by

\[
Vf(t) = \int_0^t f(\tau)d\tau
\]

for \( f \in X \) and \( t \in [0, 1] \) maps into \( D(A) \) and \( AVf = f \), thanks to Theorem 3.10. Hence, \( A \) has the bounded inverse \( V \) and is thus closed due to Lemma 1.5. Theorem 4.6 and Example 4.7 now yield that \( \sigma(A) = \emptyset \).

**Example 4.10.** Let \( X = L^p(0, 1), 1 \leq p < \infty \), and \( A = \partial^2 \) with

\[
D(A) = W^2_p(0, 1) \cap \dot{W}^1_1(0, 1).
\]

(\( A \) is the one-dimensional “Dirichlet-Laplacian”.) Let \( u_n \in D(A) \) and \( u, v \in X \) be given such that \( u_n \to u \) and \( \partial^2 u_n \to v \) in \( X \) as \( n \to \infty \). Example 4.1 shows that \( u_n \to u \) in \( W^2_2(0, 1) \) and \( \partial^2 u = v \). Moreover, \( u \in W^2_2(0, 1) \) since \( W^2_2(0, 1) \cap W^1_1(0, 1) \) is closed in \( W^2_2(0, 1) \). As a result, \( u \in D(A) \) and \( Au = v \) so that \( A \) is closed.

Suppose that \( \lambda = \mu^2 \) is an eigenvalue of \( A \) with eigenfunction \( 0 \neq u \in D(A) \). Then \( \partial u \in W^1_1(0, 1) \to C([0, 1]) \) by Theorem 3.10, so that \( u \in C^1([0, 1]) \) (cf. Example 4.2). Moreover, \( \partial^2 u = \lambda u \in C([0, 1]) \) and so \( u \in C^2([0, 1]) \). It follows that \( u = ae_\mu + be_- \mu \neq 0 \), for some \( a, b \in \mathbb{C} \). The boundary conditions \( u(0) = u(1) = 0 \) imply that \( a + b = 0 \) and \( ae_\mu + be_- \mu = 0 \), which yields \( a = -b \neq 0, e_\mu = e_- \mu \) and \( \mu \neq 0 \). Hence, \( \Re \mu = -\Re \mu \), so that \( \mu \in i\mathbb{R}\setminus\{0\} \). We thus obtain \( e_\mu = \overline{e_- \mu} \) and \( |e_\mu| = 1 \), which gives \( e_\mu \in \{ \pm 1 \} \) and so \( \mu \in i\pi\mathbb{Z}\setminus\{0\} \). Hence, \( \lambda = \mu^2 = \pi^2k^2 \) for some \( k \in \mathbb{N} \). Conversely, \( \lambda = -\pi^2k^2 \) is an eigenvalue with eigenfunction

\[
u(t) = \exp(i\pi kt) - \exp(-i\pi kt) = 2\sin(\pi kt).
\]

Given \( f \in X \) we set

\[
Rf(t) = \frac{1}{2} \int_0^1 |t-s|f(s)ds = \frac{1}{2} t \int_0^t f(s)ds - \frac{1}{2} \int_0^t sf(s)ds - \frac{1}{2} \int_t^1 f(s)ds + \frac{1}{2} \int_t^1 sf(s)ds.
\]
for $t \in [0, 1]$. As in Example 4.2 we see that $Rf \in W^1(0, 1) \cap L^p(0, 1)$ and
\[
\partial Rf(t) = \frac{1}{2} \int_0^t f(s)ds - \frac{1}{2} \int_t^1 f(s)ds
\]
for $t \in [0, 1]$. Consequently, $Rf \in W^2(0, 1)$ and $\partial^2 Rf = f \in L^p(0, 1)$. Thus $Rf \in W^2_p(0, 1)$. Since affine functions $h$ satisfy $h'' = 0$, the function
\[
R_0 f(t) = Rf(t) - (1-t)Rf(0) - tRf(1),
\]
for $t \in (0, 1)$, belongs to $D(A) = W^2_p(0, 1) \cap W^1_2(0, 1)$ and $AR_0 f = f$. Hence, $R_0$ is the inverse of $A$. Theorem 4.6 and Example 4.7 now yield that $\sigma(A) = \sigma_p(A) = \{-\pi^2 k^2 : k \in \mathbb{N}\}$.

**Example 4.11.** Let $X = L^2(\mathbb{R}^d)$ and $A = \triangle = \partial_{11} + \ldots + \partial_{dd}$ with $D(A) = W^2_2(\mathbb{R}^d)$. Then $A$ is closed.

**Proof.** Let $u \in C^3_c(\mathbb{R}^d)$ and $k, l \in \{1, \ldots, d\}$. Integration by parts yields
\[
\|\partial_{kl} u\|_2^2 = \int_{\mathbb{R}^d} \partial_{kl} u \partial_{kl} udx = - \int_{\mathbb{R}^d} \partial_{kl} u \partial_{kl} udx = \int_{\mathbb{R}^d} \partial_{kk} u \partial_{ll} udx,
\]
\[
\sum_{k,l=1}^d \|\partial_{kl} u\|_2^2 = \int_{\mathbb{R}^d} \sum_{k=1}^d \partial_{kk} u \sum_{l=1}^d \partial_{ll} udx = \int_{\mathbb{R}^d} \triangle u \triangle udx = \|\triangle u\|_2^2.
\]
By Exercise 10, we have
\[
\|u\|_{2,2}^2 \leq c_0 \left( \|u\|_2^2 + \sum_{k,l=1}^d \|\partial_{kl} u\|_2^2 \right)
\]
for some constant $c_0 > 0$, so that the above equality yields
\[(*) \quad \|u\|_{2,2}^2 \leq c_0 \left( \|u\|_2^2 + \|\triangle u\|_2^2 \right) \leq \sqrt{2} c_0 \|u\|^2_A \]
for all $u \in C^3_c(\mathbb{R}^d)$. For $v \in W^2_2(\mathbb{R}^d)$ we further have
\[
\|v\|_A^2 \leq \|v\|_2^2 + \left( \sum_{k=1}^d \|\partial_{kk} v\|_2 \right)^2 \leq \|v\|_2^2 + d \sum_{k=1}^d \|\partial_{kk} v\|_2^2 \leq d \|v\|_{2,2}^2,
\]
by Cauchy-Schwarz inequality in $\mathbb{R}^d$. If $u \in W^2_2(\mathbb{R}^d)$, then there are $u_n \in C_c^\infty(\mathbb{R}^d)$ converging to $u$ in $W^2_2(\mathbb{R}^d)$ and thus also for $\|\cdot\|_A$ by the above estimate. Using also $(*)$, we obtain
\[
\|u\|_{2,2}^2 = \lim_{n \to \infty} \|u_n\|_{2,2}^2 \leq \sqrt{2} c_0 \lim_{n \to \infty} \|u_n\|_A^2 = \sqrt{2} c_0 \|u\|_A^2.
\]
As a result, the graph norm of $A$ is equivalent to the complete norm $\|\cdot\|_{2,2}$ on $D(A) = W^2_2(\mathbb{R}^d)$. So $(D(A), \|\cdot\|_A)$ is a Banach space and thus $A$ is closed by Lemma 1.5.
Example 4.12. Let $1 < p < \infty$, $X = L^p(\mathbb{R}^d)$, and $A = \Delta$ with $D(A) = W^2_p(\mathbb{R}^d)$. The Calderón-Zygmund estimate says that the graph norm of $A$ is equivalent to $\|\cdot\|_{2,p}$ on $C^\infty_c(\mathbb{R}^d)$, see e.g. Corollary 9.10 in [Gil98]. As in Example 4.11 one then concludes that $A$ is closed.

There are related results for $X = L^1(\mathbb{R}^d)$ and $X = C_0(\mathbb{R}^d)$, see Theorem 5.8 in [Tan97] and Theorem 31.7 in [Lun95]. In theses case $p \in \{1, \infty\}$, the descriptions of the domains are more complicated and not just Sobolev (or $C^k$-) spaces. To indicate the difficulties, we note that there is a function $u \notin W^2_\infty(B(0, 1))$ such that $\Delta u \in L^\infty(B(0, 1))$, namely

$$u(x, y) = \begin{cases} (x^2 - y^2) \ln(x^2 + y^2) & \text{for } (x, y) \neq (0, 0), \\ 0 & \text{for } (x, y) = (0, 0). \end{cases}$$

Then the second derivative

$$\partial_{xx}u(x, y) = \frac{4x^2}{x^2 + y^2} + \frac{(6x^2 - 2y^2)(x^2 + y^2) - 4x^2(x^2 - y^2)}{(x^2 + y^2)^2}$$

is unbounded around $(0, 0)$, but $\Delta u(x, y) = 8x^2 - y^2$ is bounded.

Example 4.13. Let $1 < p < \infty$, $U \subseteq \mathbb{R}^d$ be bounded and open with $\partial U \in C^2$, $X = L^p(U)$, and $A = \Delta$ with $D(A) = W^2_p(U) \cap \dot{W}^1_p(U)$ ("Dirichlet-Laplacian"). Then $A$ is closed and has a compact resolvent. In fact, the closedness can be proved as in Example 4.10, using Example 4.12. In view of Example 4.7 it suffices to show that $A$ is bijective. This is done e.g. in Theorem 9.15 of [Gil98].

4.2 Selfadjoint operators and their spectra

Let $X$ and $Y$ be Hilbert spaces with scalar product $(\cdot | \cdot)$. Let $T \in \mathcal{L}(X, Y)$ and $A$ be a linear densely defined operator from $X$ to $Y$. We define the Hilbert space adjoints $T'$ and $A'$ as in the Banach space case by

for all $x \in X$, $y \in Y$: $$(x | T'y) = (Tx | y),$$

$$D(A') = \{y \in Y : \exists z \in X \forall x \in D(A) : (Ax | y) = (x | z)\}, \quad A'y := z$$

As in Remark 1.23 one sees that $A'$ is well-defined, linear and closed. We say that $T$ is unitary if $T$ is invertible with $T^{-1} = T'$. If $X = Y$, we call $T$ selfadjoint if $T = T'$ and normal if $TT' = T'T$. Note that if $T$ is unitary or selfadjoint, then $T$ and $\lambda I - T$ are normal for all $\lambda \in \mathbb{C}$. Also, a densely defined linear operator is selfadjoint if $A = A'$ (in particular, $D(A) = D(A')$ and $A$ must be closed) and normal if $AA' = A'A$. 

60
4.2. SELFADJOINT OPERATORS AND THEIR SPECTRA

Let $\Phi : X \to X^*$ be the Riesz isomorphism given by $(\Phi(x))(y) = (y|x)$ for all $x, y \in X$. For $\lambda \in \mathbb{C}$ and $X = Y$, we then obtain

$$\lambda I_X - T^* = \Phi(\lambda I_X - T^*)\Phi^{-1},$$
$$\lambda I_X - A^* = \Phi(\lambda I_X - A^*)\Phi^{-1},$$

with $D(A^*) = \Phi D(A^*)$, see §4.4 in [FA]. So Theorem 1.24 implies that

$$\sigma(A) = \sigma(A^*) = \overline{\sigma(A)} \quad \text{and} \quad \sigma_r(A) = \sigma_p(A^*) = \overline{\sigma_p(A)},$$

where the bars mean complex conjugation. Moreover,

$$R(\lambda, A^*) = \Phi^{-1} R(\lambda, A^*) \Phi = \Phi^{-1} R(\lambda, A)^* \Phi = R(\lambda, A)'$$

for all $\lambda \in \rho(A)$.

Recall that $\|T\| = \|T\|$ for every $T \in \mathcal{L}(X, Y)$, $T'' = T$ and $(ST)' = T'S'$ for $S \in \mathcal{L}(Y, Z)$.

**Proposition 4.14.** Let $X$ and $Y$ be Hilbert spaces. If $T \in \mathcal{L}(X, Y)$, then $\|T'T\| = \|TT''\| = \|T\|^2$. Let $T \in \mathcal{L}(X)$ be normal. We then have $\|T\| = r(T)$, and thus $T = 0$ if $\sigma(T) = \{0\}$.

**Remark.** In $X = \mathbb{C}^2$, the Jordan matrix $T = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ has the spectrum $\sigma(T) = \{0\}$, but $\|T\| = 1$. The equality $\|T\| = r(T)$ will be the key to the deeper properties of selfadjoint (or normal) operators.

**Proof of Proposition 4.14.** For $x \in X$ we have

$$\|Tx\|^2 = (Tx|Tx) = (T'Tx|x) \leq \|T'T\|\|x\|^2,$$
$$\|T\|^2 = \sup_{\|x\| \leq 1} \|Tx\|^2 \leq \|T'T\| \leq \|T\|^2 ||T|| = \|T\|^2.$$

Hence, $\|T\|^2 = \|T'T\|$. Similarly, one obtains $\|T\|^2 = \|T'\|^2 = \|TT''\|$.

Next, let $T$ be normal. From the first part we then deduce that

$$\|T^2\|^2 = \|T^2(T^2)'\| = \|TT'T'^*\| = \|TT'(TT')^*\| = \|TT''\|^2 = \|T\|^4,$$

so that $\|T^2\| = \|T\|^2$. Iteratively it follows that $\|T^{2n}\| = \|T\|^{2n}$ for all $n \in \mathbb{N}$. Using Theorem 1.16, we conclude that

$$r(T) = \lim_{n \to \infty} \|T^m\|^{1/n} = \lim_{n \to \infty} \|T^{2n}\|^{2/n} = \|T\|. \quad \Box$$

**Definition 4.15.** Let $A$ and $B$ be linear operators from a Banach space $X$ to a Banach space $Y$. We say that $B$ extends $A$ (and write $A \subseteq B$) if $D(A) \subseteq D(B)$ and $Ax = Bx$ for all $x \in D(A)$.

Next, let $X$ be a Hilbert space. A linear operator $A$ on $X$ is called symmetric if we have $(Ax|y) = (x|Ay)$ for all $x, y \in D(A)$.
Remark 4.16. a) We have $A \subseteq B$ for linear operators if and only if $\text{gr}(A) \subseteq \text{gr}(B)$. It then holds $A = B$ if and only if $D(B) \subseteq D(A)$.

b) Let $A \subseteq B$ for linear operators on the Banach spaces $X$ and $Y$. If $A$ is surjective and $B$ is injective, then $A = B$. As a consequence, if $X = Y$ and there is a $\lambda \in \mathbb{C}$ such that $\lambda I - A$ is surjective and $\lambda I - B$ is injective, then $A = B$.

Proof. Let $x \in D(B)$ and set $y = Bx$. Then there is a $z \in D(A)$ such that $Bx = y = Az = Bz$. Since $B$ is injective, we obtain $x = z \in D(A)$ so that $A = B$. \qed

c) If $A$ is densely defined and symmetric on a Hilbert space $X$, then $A \subseteq A'$ and $A$ is selfadjoint if $D(A') \subseteq D(A)$.

Lemma 4.17. Let $A$ be symmetric and $\alpha, \beta \in \mathbb{R}$. Set $\lambda = \alpha + i\beta$. For $x \in D(A)$ it holds $(Ax|x) \in \mathbb{R}$ and

$$\|\lambda x - Ax\|^2 = \|\alpha x - Ax\|^2 + |\beta|^2\|x\|^2 \geq |\beta|^2\|x\|^2.$$ 

In particular, if $A$ is closed, then $\sigma_{\text{ap}}(A) \subseteq \mathbb{R}$. If $\lambda \in \rho(A) \setminus \mathbb{R}$, then $\|R(\lambda, A)\| \leq \frac{1}{|\text{Im}\lambda|}$.

Proof. For $x \in D(A)$ we have $(Ax|x) = (x|Ax) = (\overline{Ax|x})$ so that $(Ax|x) = (x|Ax)$ is real. From this fact we deduce that

$$\|\lambda x - Ax\|^2 = (\alpha x - Ax + i\beta x)(\alpha x - Ax + i\beta x)$$

$$= \|\alpha x - Ax\|^2 + 2\text{Re}(i\beta x|\alpha x - Ax) + \|i\beta x\|^2$$

$$= \|\alpha x - Ax\|^2 + 2\text{Re}(i\beta x|\alpha x - Ax) + \|i\beta x\|^2$$

$$= \|\alpha x - Ax\|^2 + |\beta|^2\|x\|^2 \geq |\beta|^2\|x\|^2.$$ 

In particular, $\lambda \notin \sigma_{\text{ap}}(A)$ if $\text{Im}\lambda = \beta \neq 0$.

If $\lambda \in \rho(A) \setminus \mathbb{R}$ and $y \in X$, write $x = R(\lambda, A)y \in D(A)$. We then calculate

$$\|y\|^2 = \|\lambda x - Ax\|^2 \geq |\text{Im}\lambda|^2\|x\|^2 = |\text{Im}\lambda|^2\|R(\lambda, A)y\|^2$$

which yields the final inequality in the lemma. \qed

Theorem 4.18. Let $X$ be a Hilbert space and $A$ be densely defined, closed and symmetric.

a) Then $\sigma(A)$ is either a subset of $\mathbb{R}$ or $\sigma(A) = \mathbb{C}$ or $\sigma(A) = \{\lambda \in \mathbb{C} : \text{Im}\lambda \geq 0\}$ or $\sigma(A) = \{\lambda \in \mathbb{C} : \text{Im}\lambda \leq 0\}$.

b) The following assertions are equivalent.

1) $A = A'$,

2) $\sigma(A) \subseteq \mathbb{R}$,

3) $iI - A'$ and $iI + A'$ are injective,
4.2. SELFADJOINT OPERATORS AND THEIR SPECTRA

(4) \((iI - A)D(A)\) and \((iI + A)D(A)\) are dense.

c) Let \(A\) be selfadjoint. Then we have

\[
(4.3) \quad \|R(\lambda, A)\| \leq \frac{1}{|\text{Im} \lambda|}
\]

for \(\lambda \notin \mathbb{R}\). Further, \(\sigma(A) = \sigma_{ap}(A)\) and \(A\) has no selfadjoint extension \(B \neq A\).

Proof. a) Suppose that there would exists \(\lambda \in \sigma(A)\) and \(\mu \in \rho(A)\) with \(\text{Im} \lambda > 0\) and \(\text{Im} \mu > 0\). The line segment from \(\lambda\) to \(\mu\) must contain a point \(\gamma \in \partial \sigma(A)\). Then, \(\text{Im} \gamma > 0\) and \(\gamma \in \sigma_{ap}(A)\) by Proposition 1.19, which contradicts Lemma 4.17 since \(A\) is symmetric. A similar fact holds if \(\text{Im} \lambda < 0\) and \(\text{Im} \mu < 0\). Assertion a) thus follows.

b) Let \(A\) be selfadjoint. Lemma 4.17 yields \(\sigma_{ap}(A) \subseteq \mathbb{R}\). Due to (4.2) we also have \(\sigma_r(A) = \sigma_p(A') = \sigma_p(A)\). Hence, \(\sigma_r(A) = \sigma_p(A) \subseteq \mathbb{R}\). From Proposition 1.19 we thus deduce \(\sigma(A) \subseteq \mathbb{R}\); i.e., (1) implies (2). The implication (2) \(\Rightarrow\) (3) is obvious. Equation (4.2) also shows that \(\pm i \in \sigma_p(A')\) if and only if \(\mp i \in \sigma_r(A)\) so that (3) and (4) are equivalent. Let (4) (and thus (3)) hold. Due to Lemma 4.17 and Proposition 1.19, the range of \(iA \pm A\) is closed. In view of (4), \(iA \pm A\) is then surjective. Due to (3), \(iA \pm A'\) is injective, and hence \(A = A'\) thanks to Remark 4.16b). We also have verified the implication (4) \(\Rightarrow\) (1).

c) Let \(A = A'\). Then \(\sigma(A) = \sigma(A') \subseteq \mathbb{R}\) so that assertion c) follows from Lemma 4.17 and Remark 4.16.

Example 4.19. a) Let \(X = L^2(0, \infty)\) and \(A = i\partial\) on \(D(A) = \overset{\circ}{W}^1_2(0, \infty)\). For \(u \in D(A)\) and \(v \in C^\infty_c(0, \infty)\), we have

\[
(Au|v) = i \int_0^\infty (\partial u)v dx = -i \int_0^\infty u \partial v dx = \int_0^\infty uv' dx = (u|Av).
\]

For \(v \in D(A)\) there are \(v_n \in C^\infty_c(0, \infty)\) converging to \(v\) in \(W^1_2(0, \infty)\), so that \(v_n \rightarrow v\) and \(Av_n \rightarrow Av\) in \(X\) as \(n \rightarrow \infty\). Hence, \((Au|v) = (u|Av)\) for all \(u, v \in D(A)\), i.e., \(A\) is symmetric. By the exercises, \(\sigma(iA) = \{\lambda \in \mathbb{C} : \text{Re} \lambda \leq 0\}\) so that \(\sigma(A) = -i\sigma(iA) = \{\lambda \in \mathbb{C} : \text{Im} \lambda \geq 0\}\) (use Exercise 4). Consequently, \(A\) is not selfadjoint.

b) Let \(X = L^2(\mathbb{R})\) and \(A = i\partial\) with \(D(A) = W^1_2(\mathbb{R})\). As in part a) one shows that \(A\) is symmetric. Example 4.2 implies that \(\sigma(A) = i\sigma(-iA) = i^2\mathbb{R} = \mathbb{R}\). Hence, \(A\) is selfadjoint.

c) Let \(X = L^2(\mathbb{R})\) and \(A = \partial^2\) with \(D(A) = W^2_2(\mathbb{R})\). For \(u \in D(A)\) and \(v \in C^\infty_c(\mathbb{R})\) we have

\[
(Au|v) = \int_\mathbb{R} (\partial^2 u)v = (-1)^2 \int_\mathbb{R} u \partial^2 v = (u|Av).
\]

If \(v \in D(A)\) there exist some \(v_n \in C^\infty_c(\mathbb{R})\) with \(v_n \rightarrow v\) and \(Av_n = v''_n \rightarrow \partial^2 v = Av\) in \(X\) as \(n \rightarrow \infty\). Hence \((Au|v) = (u|Av)\) and \(A\) is symmetric. Moreover, \(\sigma(A) = \mathbb{R}_-\) by Example 4.3 so that \(A\) is selfadjoint.
CHAPTER 4. SELFADJOINT OPERATORS

Lemma 4.20 (Green’s formula). Let \( U \subseteq \mathbb{R}^d \) be bounded and open with \( \partial U \in C^2 \), \( 1 < p < \infty \), \( u \in W^2_p(U) \) and \( v \in W^2_p(U) \). We then have

\[
\int_U (\triangle u) v dx = \int_U u \triangle v dx + \int_{\partial U} (\partial_n u) v d\sigma - \int_{\partial U} u \partial_n v d\sigma.
\]

(In the boundary integrals we write \( u \) and \( v \) instead of \( Tu \) and \( Tv \), \( \nu \) is the outer unit normal of \( \partial U \), and \( \partial_n u = \sum_{j=1}^d \nu_j T \partial_j u \).)

Proof. Green’s formula holds for \( u,v \in C^2(\overline{U}) \) by Analysis 3. For \( u \in W^2_p(U) \) and \( v \in W^2_p(U) \) there are \( u_n, v_n \in C^2(\overline{U}) \) such that \( u_n \to u \) in \( W^2_p(U) \) and \( v_n \to v \) in \( W^2_p(U) \) as \( n \to \infty \), due to Corollary 3.23 and Theorem 3.24. Theorem 4.8 further yields that \( u_n|\partial U \to Tu \) in \( L^p(\partial U, d\sigma) \) and \( v_n|\partial U \to Tv \) in \( L^p(\partial U, d\sigma) \). Since \( \partial_j : W^2_q(U) \to W^1_q(U) \) is continuous for \( q = p, p' \) and \( j \in \{1, \ldots, d\} \), we further obtain \( \partial_j u_n|\partial U \to T \partial_j u \) in \( L^p(\partial U, d\sigma) \) and \( \partial_j v_n|\partial U \to T \partial_j v \) in \( L^p(\partial U, d\sigma) \) as \( n \to \infty \). The assertion now follows by approximation. \( \square \)

We note that the result is also true for \( p \in \{1, \infty\} \), with a modified proof.

Example 4.21. a) Let \( U \subseteq \mathbb{R}^d \) be open and bounded with \( \partial U \in C^2 \) and let \( A_1 = \triangle \) with \( D(A_1) = W^2_2(U) \cap W^1_2(U) \). For \( u,v \in D(A_1) \) we have \( Tu = Tv = 0 \) so that Lemma 4.20 yields

\[
(A_1 u|v) = \int_U (\triangle u)v dx = \int_U u \triangle v dx = (u|A_1 v);
\]
i.e., \( A_1 \) is symmetric. In Example 4.13 we have seen that \( 0 \in \rho(A_1) \) (see also Example 4.10 for \( d = 1 \)). So \( A_1 \) is selfadjoint due to Theorem 4.18.

b) On \( X = L^2(0,1) \) consider \( A_0 = \partial^2 \) with \( D(A_0) = W^2_2(0,1) \) and \( A_1 \) from part a) with \( U = (0,1) \). As in a) we see that \( A_0 \) is symmetric. Since \( A_0 \not\subseteq A_1 \) and \( \sigma(A_1) \subseteq \mathbb{R} \), Remark 4.16 implies that \( \mathbb{C} \setminus \mathbb{R} \) \( \subseteq \sigma(A_0) \). Hence, \( A_0 \) is not selfadjoint and \( \sigma(A_0) = \mathbb{C} \).

Example 4.22. Let \( X = L^2(\mathbb{R}^d) \) and \( A = \triangle \) with \( D(A) = W^2_2(\mathbb{R}^d) \). For \( u,v \in C^\infty_c(\mathbb{R}^d) \) we have \( (Au|v) = (u|Av) \). Since \( C^\infty_c(\mathbb{R}^d) \) is dense in \( W^2_2(\mathbb{R}^d) \), we obtain as before the symmetry of \( A \). To compute the spectrum of \( A \) we use the Fourier transform \( \mathcal{F} \) given by

\[
\hat{u}(\xi) = \mathcal{F} u(\xi) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{-i\xi \cdot x} f(x) dx,
\]
for \( \xi \in \mathbb{R}^d \) and \( f \in L^1(\mathbb{R}^d) \). It is known that \( \mathcal{F} \) can be extended from \( L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d) \) to a unitary operator \( \mathcal{F} \) on \( L^2(\mathbb{R}^d) \). Let \( u \in C^\infty_c(\mathbb{R}^d) \). It then holds

\[
(4.4) \quad \mathcal{F} \partial_i \partial_j u = -\xi_i \xi_j \hat{u},
\]
see e.g. §5 in [FA]. These facts imply

\[
\|u\|_{L^2}^2 \simeq \|u\|_{L^2}^2 + \sum_{i,j=1}^d \|\partial_i \partial_j u\|_{L^2}^2 = \|\hat{u}\|_{L^2}^2 + \sum_{i,j=1}^d \|\mathcal{F} \partial_i \partial_j u\|_{L^2}^2
\]

64
(4.5) \[
\int_{\mathbb{R}^d} \left((\hat{u}(\xi))^2 + \sum_{i,j=1}^{d} |\xi_i \xi_j|^2 |\hat{u}(\xi)|^2\right) d\xi \simeq \|1 + |\xi|^2_2 \hat{u}\|_2^2,
\]
where \(a \simeq b\) means that \(c_1 a \leq b \leq c_2 a\) for some constants \(c_1, c_2 > 0\) and all \(a, b\).

Let \(u \in W^2_2(\mathbb{R}^d)\). There are \(u_n \in C_0^\infty(\mathbb{R}^d)\) converging to \(u\) in \(W^2_2(\mathbb{R}^d)\). Then \(\hat{u}_n \to \hat{u}\) in \(L^2(\mathbb{R}^d)\), and by (4.5), the sequence \(((1 + |\xi|^2_2)\hat{u}_n)\) is Cauchy in \(L^2(\mathbb{R}^d)\).

Hence, the functions \((1 + |\xi|^2_2)\hat{u}_n\) converge to \((1 + |\xi|^2_2)\hat{u}\) in \(L^2(\mathbb{R}^d)\) as \(n \to \infty\). Consequently, (4.5) and (4.4) hold on \(W^2_2(\mathbb{R}^d)\) by approximation. If \(u \in L^2(\mathbb{R}^d)\) satisfies \(|\xi|^2_2 \hat{u} \in L^2(\mathbb{R}^d)\), we have for \(\varphi \in C_0^\infty(\mathbb{R}^d)\) that

\[
\int_{\mathbb{R}^d} u \partial_i \partial_j \varphi dx = (u|\partial_i \partial_j \varphi) = (\mathcal{F}u|\mathcal{F}\partial_i \partial_j \varphi) = (\hat{u}|(-\xi_i \xi_j) \varphi) = (-\xi_i \xi_j \hat{u} | \varphi) = \int_{\mathbb{R}^d} -\mathcal{F}^{-1}(\xi_i \xi_j \hat{u}) | \varphi dx.
\]

In other words, there exists the weak derivatives \(\partial_i \partial_j u = -\mathcal{F}^{-1}(\xi_i \xi_j \hat{u}) \in L^2(\mathbb{R}^d)\) for all \(i, j \in \{1, \ldots, d\}\), and thus \(u \in W^2_2(\mathbb{R}^d)\). We have shown the important characterization

\[
W^2_2(\mathbb{R}^d) = \{u \in L^2(\mathbb{R}^d) : |\xi|^2_2 \hat{u} \in L^2(\mathbb{R}^d)\}
\]

for the Sobolev space with \(p = 2\).

Let \(f \in L^2(\mathbb{R}^d)\) and \(\lambda \in \mathbb{C}\). By the above results, \(u \in L^2(\mathbb{R}^d)\) belongs to \(W^2_2(\mathbb{R}^d)\) and satisfies \(\lambda u - \triangle u = f\) on \(\mathbb{R}^d\) if and only if \(|\xi|^2_2 \hat{u} \in L^2(\mathbb{R}^d)\) and \((\lambda + |\xi|^2_2) \hat{u} = \hat{f}\) on \(\mathbb{R}^d\). The latter implies that \(\hat{u} = \frac{1}{\lambda + |\xi|^2_2} \hat{f} = m \hat{f}\). If \(\lambda \notin \mathbb{R}_-\), then \(m \in L^\infty(\mathbb{R}^d)\) so that the function \(u := \mathcal{F}^{-1}(\frac{1}{\lambda + |\xi|^2_2} \hat{f})\) belongs to \(L^2(\mathbb{R}^d)\), \(|\xi|^2_2 \hat{u} = \frac{|\xi|^2_2}{\lambda + |\xi|^2_2} \hat{f} \in L^2(\mathbb{R}^d)\) and \((\lambda + |\xi|^2_2) \hat{u} = \hat{f}\). As a result, \(\lambda \in \rho(A)\) and \(\sigma(A) \subseteq \mathbb{R}_-\).

If \(\lambda < 0\), then \(m\) is unbounded. Suppose that \(mg \in L^2(\mathbb{R}^d)\) for all \(g \in L^2(\mathbb{R}^d)\). Then the (closed) multiplication operator \(M : g \mapsto mg\) in \(L^2(\mathbb{R}^d)\) were defined on \(L^2(\mathbb{R}^d)\) and thus bounded by the closed graph theorem. But then \(\sigma(M) = \overline{m(\mathbb{R}^d)}\) would be bounded, which is wrong. So there is an \(f \in L^2(\mathbb{R}^d)\) such that \(m \hat{f} \notin L^2(\mathbb{R}^d)\) and so \(\lambda \in \sigma(A)\). Summing up, \(\sigma(A) = \mathbb{R}_-\).

**Theorem 4.23.** Let \(A\) be densely defined and selfadjoint on the Hilbert space \(X\) and let \(B\) be symmetric with \(D(A) \subseteq \mathcal{D}(B)\). Assume there are constants \(c > 0\) and \(\delta \in [0, 1/2]\) such that \(\|Bx\| \leq c\|x\| + \delta \|Ax\|\) for all \(x \in D(A)\). Then \(A + B \in \mathcal{L}(X)\) is selfadjoint.

**Proof.** By Theorem 4.18, we have \(it \in \rho(A)\) for all \(t \in \mathbb{R}\setminus \{0\}\). Take \(\varepsilon \in (0, 1 - 2\delta) \subseteq (0, 1)\) and \(x \in X\). Using (4.3) we estimate

\[
\|B(itI - A)^{-1}x\| \leq \delta\|A(itI - A)^{-1}x\| + c\|i(itI - A)^{-1}x\|
\]

\[
= \delta\|it(itI - A)^{-1}x - x\| + c\|i(itI - A)^{-1}x\|
\]

\[
\leq \delta\left(\frac{|t|} {|t|} + 1\right)\|x\| + c\|x\| \leq (1 - \varepsilon)\|x\|,
\]
whenever $|t| \geq \frac{\varepsilon}{1-2\delta-\varepsilon}$. Theorem 1.26 now implies that $\pm it \in \rho(A + B)$ for such $t$. Finally, $A + B$ is symmetric since

$$(A + B)x|y) = (Ax|y) + (Bx|y) = (x|Ay) + (x|By) = (x|(A + B)y)$$

for all $x, y \in D(A)$. So, $A + B$ is selfadjoint due to Theorem 4.18. \hfill \Box

Actually, in the above theorem it suffices to assume that $\delta < 1$, see Theorem X.13 in [Ree75].

**Example 4.24.** Let $X = L^2(\mathbb{R}^3)$, $A = \triangle$ with $D(A) = W^2_2(\mathbb{R}^3)$ and $V : D(A) \to X$ be given by $Vu(x) = \frac{b}{|x|^2}u(x)$ for some $b \in \mathbb{R}$. Since $kp = 4 > d = 3$ we have $D(A) \hookrightarrow C_0(\mathbb{R}^3)$ due to Sobolev’s embedding.

Let $0 \leq \varepsilon \leq 1$. Using polar coordinates and Example 4.11, we compute

$$\int_{\mathbb{R}^3} |Vu|^2 dx = b^2 \int_{B(0, \varepsilon)} \frac{|u(x)|^2}{|x|^2} dx + b^2 \int_{\mathbb{R}^3 \setminus B(0, \varepsilon)} \frac{|u(x)|^2}{|x|^2} dx$$

$$\leq c \int_0^\varepsilon r^2 \frac{\varepsilon^2}{r^2} dr \|u\|_\infty^2 + \frac{b^2}{\varepsilon^2} \int_{\mathbb{R}^3 \setminus B(0, \varepsilon)} |u|^2 dx$$

$$\leq c \varepsilon \|u\|_2^2 + \frac{b^2}{\varepsilon^2} \|u\|^2_2 \leq c \varepsilon \|Au\|_2^2 + c \varepsilon \|u\|_2^2 + \frac{b^2}{\varepsilon^2} \|u\|^2_2,$$

for constants $c > 0$ independent of $u$ and $\varepsilon$. Moreover, $V$ is symmetric on $D(A)$ since

$$(Vu|v) = \int_{\mathbb{R}^3} \frac{b}{|x|^2} u(x) v(x) dx = (u|Vv)$$

for all $u, v \in D(A)$. So Theorem 4.23 implies that $A + V$ is selfadjoint if we take a small $\varepsilon > 0$.

The spectra $\sigma(A + V)$ and $\sigma_p(A + V)$ and the eigenfunctions of $A + V$ are computed in §7.3.4 of [Tri97], where $b > 0$. The above operator $A = \triangle + V$ is used in physics to describe the hydrogen atom, see also Example 4.33 below.

### 4.3 The spectral theorems

The following results and proofs hold for normal operators with minor modifications. Also, the separability assumption made below can be removed. See e.g. Corollaries X.2.8 and X.5.4 in [DS88b] or Theorems 13.24, 13.30 and 13.33 in [Rud91].

Let $T \in \mathcal{L}(Z)$ for a Banach space $Z$ and $p(z) = a_0 + a_1 z + \ldots + a_n z^n$ be a complex polynomial. We then define the operator polynomial

$$p(T) = a_0 I + a_1 T + \ldots + a_n T^n \in \mathcal{L}(Z).$$

This gives a map $p \mapsto p(T)$ from the space of polynomials to $\mathcal{L}(Z)$. For selfadjoint $T$ on a Hilbert space one can extend this map to all $f \in C(\sigma(T))$, as seen in the next theorem. We set $p_1(z) = z$. 

66
4.3. THE SPECTRAL THEOREMS

Theorem 4.25 (Continuous functional calculus). Let $T \in \mathcal{L}(X)$ be selfadjoint on a Hilbert space $X$. There is exactly one map $\Phi_T : C(\sigma(T)) \to \mathcal{L}(X)$, $f \mapsto f(T)$, such that

(C1) $(\alpha f + \beta g)(T) = \alpha f(T) + \beta g(T),$

(C2) $\|f(T)\| = \|f\|_\infty$ (hence, $\Phi_T$ is injective),

(C3) $1(T) = I$ and $p_1(T) = T$,

(C4) $(fg)(T) = f(T)g(T) = g(T)f(T),$

(C5) $f(T)' = f'(T),$

hold for all $f, g \in C(\sigma(T))$ and $\alpha, \beta \in \mathbb{C}$. In particular, we have $\Phi_T(p) = p(T)$ for each polynomial $p$, where $p(T)$ is given by (4.6).

Proof. We first show the properties (C1)-(C5) for polynomials $p(t) = a_0 + a_1 t + \ldots + a_n t^n$ and $q(t) = b_0 + b_1 t + \ldots + b_m t^m$ with $t \in \mathbb{R}$ and the map $p \mapsto p(T)$ defined by (4.6). Here, any $a_j, b_j \in \mathbb{C}$ may be equal to 0. Clearly, (C1) and (C3) hold in this case, and $p(T)' = \sum_{j=0}^{n} a_j T^j' = p(T)$ since $T' = T'$. Moreover,

$$(pq)(T) = \sum_{m=0}^{2n} \left( \sum_{0 \leq j, k \leq n, j+k=m} a_j b_k \right) T^m = \sum_{j=0}^{n} a_j T^j \sum_{k=0}^{n} b_j T^k = p(T)q(T).$$

It also holds that $(pq)(T) = (qp)(T) = q(T)p(T)$, so that (C4) is shown for polynomials. Properties (C4) and (C5) imply that $p(T)$ is normal. Hence, Proposition 4.14 and Lemma 4.26 below yield

$$\|p(T)\| = \max\{|\lambda| : \lambda \in \sigma(p(T))\} = \max\{|\lambda| : \lambda \in \sigma(p(T))\} = \|p\|_\infty.$$

Since $\sigma(T)$ is compact, the polynomials are dense in $C(\sigma(T))$ by Weierstraß approximation theorem. We can thus extend the map $p \mapsto p(T)$ to a linear isometry $\Phi_T : f \mapsto f(T)$ from $C(\sigma(T))$ to $\mathcal{L}(X)$. By continuity, also (C4) and (C5) hold for $\Phi_T$ on $C(\sigma(T))$.

If there is another map $\Psi : C(\sigma(T)) \to \mathcal{L}(X)$ satisfying (C1)-(C5), then $\Psi(p) = p(T) = \Phi_T(p)$ for all polynomials by (C1), (C3) and (C4), so that $\Psi = \Phi_T$. $\square$

Lemma 4.26. Let $T \in \mathcal{L}(Z)$ for a Banach space $Z$ and let $p$ be a polynomial. Then

$$\sigma(p(T)) = p(\sigma(T)).$$

Proof. See Theorem 5.3 below. $\square$

Corollary 4.27. Let $T \in \mathcal{L}(X)$ be selfadjoint and $f \in C(\sigma(T))$. Then the following assertions hold.

(C6) If $Tx = \lambda x$ for some $x \in X$ and $\lambda \in \mathbb{C}$, then $f(T)x = f(\lambda)x$.
(C7) \( f(T) \) is normal,

(C8) \( \sigma(f(T)) = f(\sigma(T)) \) (spectral mapping theorem),

(C9) \( f(T) \) is selfadjoint if and only if \( f \) is real valued.

Proof. Take a sequence of polynomials \( p_n \) converging to \( f \) uniformly. Let \( Tx = \lambda x \). Property (C6) holds for a polynomial since \( p(T)x = \sum_{j=0}^{n} a_j T^j x = \sum_{j=0}^{n} a_j \lambda^j x = p(\lambda)x \). Hence,

\[
f(T)x = \lim_{n \to \infty} p_n(T)x = \lim_{n \to \infty} p_n(\lambda)x = f(\lambda)x.
\]

From (C5) and (C4) we deduce that \( f(T)f(T)' = f(T)f(T) = f(T)'f(T) \) so that \( f(T) \) is normal.

We next show (C8). If \( \mu \notin f(\sigma(T)) \), then \( g = \frac{1}{\mu - T} \in C(\sigma(T)) \). Thus (C3) and (C4) yield

\[
(\mu I - f(T))g(T) = g(T)(\mu I - f(T)) = g(T)(\mu I - f(T)) = (\mu I - f(T))g(T) = 1(T) = I.
\]

Hence, \( \mu \in \rho(f(T)) \). Let \( \mu = f(\lambda) \) for some \( \lambda \in \sigma(T) \). Then \( \mu_n := p_n(\lambda) \) belongs to \( \sigma(p_n(T)) \) for all \( n \in \mathbb{N} \) by Lemma 4.26. Moreover, \( \mu_n I - p_n(T) \to \mu I - f(T) \).

For the last assertion, observe that \( f(T) = f(T)' \) if and only if \( (f - T')(T) = 0 \) if and only if \( f - T = 0 \), because \( \Phi_T \) is injective.

**Corollary 4.28.** Let \( n \in \mathbb{N} \) and \( T = T' \in \mathcal{L}(X) \) with \( \sigma(T) \subseteq \mathbb{R}_+ \) (in this case one writes \( T = T' \geq 0 \)). Then there is a uniquely determined selfadjoint operator \( W \in \mathcal{L}(X) \) with \( \sigma(W) \subseteq \mathbb{R}_+ \) such that \( W^n = T \).

**Proof.** Consider \( w_n(t) = t^{n/2} \) for \( t \in \sigma(T) \subseteq \mathbb{R}_+ \) and define \( W := w(T) \). Then \( W^n = w(T)^n = w(T) \cdots w(T) = (w \cdots w)(T) = p_1(T) = T \). Properties (C8) and (C9) imply that \( W = W' \geq 0 \). For the proof of the uniqueness of \( W \) we refer to Korollar VII.1.16 in [Wer05].

**Theorem 4.29** (The compact case). Let \( X \) be a Hilbert space with \( \dim X = \infty \), \( T \in \mathcal{L}(X) \) be compact and selfadjoint and \( A \) be densely defined, closed and selfadjoint on \( X \) having a compact resolvent. Then the following assertions hold.

1. There is an index set \( J \in \{0, \mathbb{N}, \{1, \ldots, N\} : N \in \mathbb{N}\} \) and eigenvalues \( \lambda_n \neq 0 \), \( n \in J \), such that

\[
\sigma(T) = \{0\} \cup \{\lambda_n : n \in J\} \subseteq \mathbb{R}
\]

where \( \lambda_n \to 0 \) as \( n \to \infty \) if \( J = \mathbb{N} \). Further, there is an orthonormal basis of \( N(T)^⊥ = R(T) \) consisting of eigenvectors of \( T \) for \( \lambda_n \), \( n \in J \). The eigenspaces \( E_n(T) = N(\lambda_n I - T) \) are finite-dimensional and the orthogonal projection \( P_n \) onto \( E_n(T) \) commutes with \( T \), for all \( n \in J \). Finally, the sum

\[
T = \sum_{n \in J} \lambda_n P_n
\]
4.3. THE SPECTRAL THEOREMS

converges in \( \mathcal{L}(X) \).

b) It holds

\[
\sigma(A) = \sigma_p(A) = \{ \mu_n : n \in \mathbb{N} \} \subseteq \mathbb{R}
\]

with \(|\mu_n| \to \infty \) as \( n \to \infty \), and there is an orthonormal basis of \( X \) consisting of eigenvectors of \( A \). The eigenspaces \( E_n(A) = N(\mu_n I - A) \) are finite-dimensional, and we have \( Q_n D(A) \subseteq D(A) \) and \( A Q_n x = Q_n Ax \) for all \( x \in D(A) \) and \( n \in \mathbb{N} \), where \( Q_n \) is the orthogonal projection onto \( E_n(A) \). Finally, the sum

\[
Ax = \sum_{n=1}^{\infty} \mu_n Q_n x
\]

converges in \( X \) for all \( x \in D(A) \).

\textbf{Proof.} 1) Theorem 2.11, 4.6 and 4.18 show the assertions on the spectra (except for the claim that \( \sigma(A) \) is infinite) and that \( \dim E_n(T) \) and \( \dim E_n(A) \) are finite for all \( n \). Let \( x, y \in D(A) \) be eigenvectors for \( \mu_j \neq \mu_k \) of \( A \). Then

\[
\mu_j \langle x | y \rangle = \langle Ax | y \rangle = \langle x | Ay \rangle = \mu_k \langle x | y \rangle,
\]

so that \( \langle x | y \rangle = 0 \). Similarly, one sees that \( E_j(T) \perp E_k(T) \) if \( j \neq k \).

2) Concerning \( T \), we first consider the case \( J = \mathbb{N} \). Using the Gram-Schmidt algorithm in each \( E_n(T) \), we can construct an orthonormal system \( S \) consisting of eigenvectors of \( T \) for the eigenvalues \( \lambda_n \neq 0 \), \( n \in \mathbb{N} \). Note that \( S \subset R(T) \). Let \( 1_n = 1_{\{\lambda_n\}} \in C(\sigma(T)) \) and \( \varphi_n = 1_1 + \ldots + 1_n \) for each \( n \in \mathbb{N} \). Then \( 1_n(T)^2 = 1_n(T) = 1_n(T) \). Moreover, (C4) and (C9) yield that \( T 1_n(T) = 1_n(T) T \), \( 1_n(T)^t = 1_n(T) \) and

\[
(\lambda_n I - T) 1_n(T) = ((\lambda_n - \lambda_n) 1_n(T)) = 0.
\]

If \( \lambda_n v = T v \) for some \( v \in X \), we further deduce \( 1_n(T)v = 1_n(\lambda_n)v = v \) from (C6). As a result, \( 1_n(T) \) is a selfadjoint projection onto \( E_n(T) \). For \( x \in X \) and \( y \in N(1_n(T)) \), we thus have \( (1_n(T)x | y) = (x | 1_n(T)y) = 0 \) so that \( 1_n(T) \) is orthogonal, i.e., \( 1_n(T) = P_n \) and \( \varphi_n(T) = P_1 + \ldots + P_n \) for all \( n \in \mathbb{N} \). Employing also (C2), we further conclude

\[
\|T - \varphi_n(T) T\| = \|T - T \varphi_n(T)\| = \|(1 - \varphi_n)p_1(T)\| = \|(1 - \varphi_n)p_1\|_{\infty} = \|p_1\|_{\infty} \sup_{j \geq n+1} |\lambda_j| \to 0,
\]

as \( n \to \infty \). Since \( T 1_n(T) = \lambda_n P_n \), we obtain \( T = \sum_{n=1}^{\infty} \lambda_n P_n \) with convergence in \( \mathcal{L}(X) \). Moreover, it follows that \( \varphi_n(T)y \in \operatorname{lin} S \) converges to \( y \) as \( n \to \infty \) for all \( y \in R(T) \). Therefore, \( S \) is an orthonormal basis of \( R(T) \) due to e.g. Theorem 2.16 in [FA]. Finally, it holds \( N(T)^{\perp} = R(T) = R(T) \) (see e.g. Proposition 4.41 in [FA]) since \( T = T^* \) and \( X \) is reflexive.

If \( J = \emptyset \), then \( T = 0 \) because of Proposition 4.14 and the assertions hold trivially. If \( J = \{1, \ldots , N\} \) is finite, then \( (1 - \varphi_N)p_1 = 0 \) on \( \sigma(T) \) and thus the above arguments work also in this case with a finite sum. So the assertions concerning \( T \) are shown.
3) Fix \( t \in \rho(A) \cap \mathbb{R} \) and note that \( 0 \notin \sigma_p(R(t, A)) \). Let \( x \in X \) be orthogonal to all eigenvectors \( v \) of \( A \). If \( v \) is an eigenvector of \( A \) with eigenvalue \( \mu \), then \( \lambda^{-1}(tI - A)v =: w \) is an eigenvector of \( R(t, A) \) for the eigenvalue \( \lambda = \frac{1}{\mu} \). Vice versa any eigenvector \( w \) of \( R(t, A) \) with eigenvalue \( \lambda \) defines the eigenvector \( v = \lambda R(t, A)w \) of \( A \) with eigenvalue \( \mu = t - 1/\lambda \). (See Proposition 1.20.) Observe that \( R(t, A)' = R(t, A') = R(t, A) \) is selfadjoint, compact and injective. By Step 2, the operator \( R(t, A) \) possesses an orthonormal base of eigenvectors \( w \). In particular, \( R(t, A) \) and thus \( A \) have infinitely many eigenvalues since \( \dim X = \infty \) and the eigenspaces of \( R(t, A) \) are finite dimensional. Moreover, \( 0 = (x|v) = \lambda (R(t, A)x|w) \) for all these eigenvectors \( w \). Since these vectors span \( X \), we obtain \( R(t, A)x = 0 \) and hence \( x = 0 \). Consequently, the eigenvectors \( v \) of \( A \) span \( X \). As for self-adjoint matrices one shows the orthogonality of eigenvectors of \( A \) to different eigenvalues. One can then construct an orthonormal basis of \( X \) consisting of eigenvectors \( v \) of \( A \), see Lemma 2.14 and Theorem 2.16 of [FA].

Let \( v_{n,j} \) \( (j = 1, \ldots, m_n) \) be eigenvectors of \( A \) forming an orthonormal basis of \( E_n(A) \) so that \( Q_nx = \sum_{j=1}^{m_n} (x|v_{n,j}) v_{n,j} \) is the orthogonal projection onto \( E_n(A) \), for each \( n \in \mathbb{N} \). As a result, \( \{v_{n,j} : j = 1, \ldots, m_n; n \in \mathbb{N} \} \) is an orthonormal basis of \( X \) and \( x = \sum_{n=1}^{\infty} Q_nx \) for all \( x \in X \). Next, for \( x \in D(A) \) we have

\[
Q_nAx = \sum_{j=1}^{m_n} (Ax|v_{n,j}) v_{n,j} = \sum_{j=1}^{m_n} (x|Av_{n,j}) v_{n,j} = \sum_{j=1}^{m_n} (x|v_n v_{n,j}) v_{n,j}
\]

\[
\sum_{j=1}^{m_n} (x|v_{n,j}) \mu_n v_{n,j} = \sum_{j=1}^{m_n} (x|v_{n,j}) Av_{n,j} = AQ_nX.
\]

We thus obtain

\[
A \sum_{k=1}^{\infty} Q_kx = \sum_{k=1}^{\infty} Q_kAx \rightarrow \sum_{k=1}^{\infty} Q_kAx = \sum_{k=1}^{\infty} AQ_kx = \sum_{k=1}^{\infty} \mu_k Q_kx,
\]

as \( n \to \infty \), so that the closedness of \( A \) yields the last assertion. \( \square \)

**Remark 4.30.** a) In the above proof we also obtain that

\[
Ax = \sum_{n=1}^{\infty} \mu_n \sum_{j=1}^{m_n} (x|v_{n,j}) v_{n,j}
\]

for all \( x \in D(A) \). An analogous result holds for \( T \).

b) Let \( T \) be selfadjoint and compact such that \( N(T) \) is separable. Let \( z_k, k \in J_0, \) be an orthonormal basis for \( N(T) \), where \( J_0 \subseteq \mathbb{Z}_- \) could be empty. Denote by \( \lambda_l \) the non-zero eigenvalues of \( T \) (repeated according to their multiplicity) with corresponding orthonormal basis of eigenvectors \( \{w_l : l \in J_1\} \). The union \( \{b_j : j \in J'\} \) of \( \{z_k : k \in J_0\} \) and \( \{w_l : l \in J_1\} \) is an orthonormal basis of \( X \), where \( J' = J_0 \cup J_1 \). By e.g. Theorem 2.18 in [FA], the map

\[
\Phi : X \to \ell^2(J'), \quad \Phi x = ((x|b_j)_{j \in J'}),
\]

70
4.3. THE SPECTRAL THEOREMS

is unitary with $\Phi^{-1}((\xi_j)_{j \in J'}) = \sum_{j \in J'} \xi_j b_j$. Moreover, the transformed operator $\Phi T \Phi^{-1}$ acts on $\ell^2(J')$ as the multiplication operator

$$
\Phi T \Phi^{-1}(\xi_j) = \Phi T \sum_{j \in J'} \xi_j b_j = \Phi \sum_{j \in J'} \lambda_j \xi_j b_j = (\lambda_j \xi_j)_{j \in J'},
$$

where $\lambda_j = 0$ if $j \in J_0$. Hence, $\Phi T \Phi^{-1}$ is represented as infinite diagonal matrix with diagonal elements $\lambda_j$. Analogous results hold for $A$ from Theorem 4.29, see e.g. Theorems 4.5.1-4.5.3 in [Tri97].

c) We point out that many of the important orthonormal bases in analysis are constructed by means of part b) of the above theorem applied to specific differential operators $A$, see [Tri97].

There are (at least) two ways to generalize the above “diagonalization” result to the non-compact case. Below we present one of them.

**Theorem 4.31** (The multiplication representation). Let $T \in \mathcal{L}(X)$ be selfadjoint on a separable Hilbert space $X$. Then there is a measure space $(\Omega, \mathcal{A}, \mu)$, a measurable function $h : \Omega \to \sigma(T)$ and a unitary operator $U : X \to L^2(\mu)$ such that

$$
Tx = U^{-1}hUx,
$$

for all $x \in X$.

**Proof (partly sketched).** 1) Let $v_1 \in X$. We define the linear subspaces

$$
Y_1 := \{f(T)v_1 : f \in C(\sigma(T))\} \quad \text{and} \quad X_1 = \overline{Y_1}
$$

of $X$. Since $T f(T)v_1 = (p_1 f)(T)v_1 \in Y_1$ for every $f \in C(\sigma(T))$, we obtain $TY_1 \subseteq Y_1$ and so $TX_1 \subseteq X_1$. We introduce the map

$$
\varphi_1 : C(\sigma(T)) \to \mathbb{C}, \quad \varphi_1(f) = (f(T)v_1 | v_1),
$$

which is linear and bounded because $|\varphi_1(f)| \leq \|f(T)||v_1\|_1^2 = \|f\|_{\infty}$ due to (C2). If $f \geq 0$, then $\sigma(f(T)) \subseteq \mathbb{R}_+$ by (C8). So we can deduce from Corollary 4.28 that

$$
(f(T)v_1 | v_1) = \left(\int (f(T)^{1/2}v_1 | f(T)^{1/2}v_1)\right) = \|f(T)^{1/2}v_1\|_2^2 \geq 0.
$$

The Riesz representation theorem of $C(\sigma(T))^*$ now gives a positive measure $\mu_1$ on $B(\sigma(T))$ such that

$$
\varphi_1(f) = \int_{\sigma(T)} f d\mu_1,
$$

for all $f \in C(\sigma(T)) \subseteq L^2(\mu_1)$, see e.g. Theorem 2.14 of [Rud87]. For any $x = f(T)v_1 \in Y_1$, we define $V_1 x := f \in L^2(\mu_1)$. We compute

$$
\|V_1 x\|_{L^2(\mu_1)}^2 = \int_{\sigma(T)} |f|^2 d\mu_1 = \varphi_1(\overline{f} f) = ((\overline{f} f)(T)v_1 | v_1)
$$
In particular, if \( x = f(T)v_1 = g(T)v_1 \) for some \( g \in \mathcal{C}(\sigma(T)) \), then \( \| f - g \|_X^2 = \| (f(T) - g(T))v_1 \|_X^2 = 0 \), and so \( f = g \) in \( L^2(\mu_1) \). As a result, \( V_1 : Y_1 \rightarrow L^2(\mu_1) \) is a linear isometric map and can be extended to a linear isometry \( U_1 : X_1 \rightarrow L^2(\mu_1) \). Observe that \( \mathcal{C}(\sigma(T)) \subseteq R(U_1) \). Riesz’ theorem also yields that \( \mu_1 \) is regular\(^1\) so that \( \mathcal{C}(\sigma(T)) \) is dense in \( L^2(\mu_1) \) due to e.g. Theorem 3.14 in [Rud87]. Hence, the isometry \( U_1 \) is bijective and thus unitary by e.g. Proposition 2.24 in [FA]. Finally, we have

\[
U_1Tf(T)v_1 = V_1(p_1f)(T)v_1 = p_1f = p_1U_1f(T)v_1,
\]

for all \( f \in \mathcal{C}(\sigma(T)) \). By density, we obtain that \( Tx = U_1^{-1}p_1U_1x \) for all \( x \in X_1 \).

2) We are done if there is an \( v_1 \in X \) such that \( X_1 = X \). In general this is not true. Using Zorn’s Lemma (see Theorem I.2.7 of [DS88a]), we instead find an orthogonal system of spaces \( X_i \) as in step 1) which span \( X \). To that aim, we introduce the collection \( \mathcal{E} \) of all sets \( E \) having as elements at most countably many closed subspaces \( X_j \subset X \) of the type constructed in step 1) such that \( X_i \perp X_j \) for all \( X_i \neq X_j \) in \( E \). The system \( \mathcal{E} \) is ordered via inclusion of sets. Let \( C \) be a chain in \( \mathcal{E} \), i.e., a subset of \( \mathcal{E} \) such that \( E \subset F \) or \( F \subset E \) for all \( E, F \in \mathcal{E} \). We put \( C = \bigcup_{E \in \mathcal{C}} E \). Clearly, \( E \subset C \) for all \( E \in \mathcal{C} \). Let \( Y, Z \in C \). Then \( Y \) and \( Z \) are closed subspaces of \( X \) as constructed in 1) and there are \( E, F \in \mathcal{C} \) such that \( Y \in E \) and \( Z \in F \). We may assume that \( E \subset F \) and thus \( Y, Z \in F \) are orthogonal (if \( Y \neq Z \)). As a result, \( C \) contains pairwise orthogonal subspaces of \( X \). If \( x \perp y \) have norm 1, then \( \| x - y \|^2 = \| x \|^2 + \| y \|^2 = 2 \). The separability of \( X \) thus implies that \( C \) contains at most countably many subspaces, i.e., \( C \in \mathcal{E} \) and so \( C \) is an upper bound of \( \mathcal{C} \). Zorn’s Lemma now gives a maximal element \( M = \{ X_j : j \in J \} \) in \( \mathcal{E} \), where \( J \subset \mathbb{N} \) and \( X_j \) are pairwise orthogonal subspaces as constructed in step 1).

Assume that there is a \( v \in X \setminus \{ 0 \} \) being orthogonal to all these \( X_j \). Let \( Z \) be the closure of the subspace of all vectors \( f(T)v \) with \( f \in \mathcal{C}(\sigma(T)) \). Let \( g \in \mathcal{C}(\sigma(T)) \) and \( x = g(T)v_i \in X_i \) for some \( i \in J \), where \( v_i \) generates \( X_i \) as in step 1). We then observe

\[
(x|f(T)v) = (f(T)g(T)v_i|v) = ((f(T)g(T))v_i|v) = 0
\]

since \( (f(T)g(T)v_i|v) \in X_i \). By density, it follows that \( Z \) is orthogonal to all \( X_i \) and thus \( M \cup \{ Z \} \in \mathcal{E} \). The maximality on \( M \) now implies that \( Z \in M \) which is only possibly if \( v = 0 \), which was excluded. As a result, \( X \) is the closed linear span of the elements in the orthogonal subspaces \( X_j \). (One writes \( X = \bigoplus_{j \in J} X_j \).) We now define

\[
\Omega = \bigcup_{j \in J} \sigma(T) \times \{ j \} \subseteq \mathbb{R}^2, \quad A = \mathcal{B}(\Omega), \quad \mu(A) = \sum_{j \in J} \mu_j(A_j)
\]

with \( A_j \times \{ j \} = A \cap (\sigma(T) \times \{ j \}) \), \( h(\omega) = \lambda \) if \( \omega = \lambda \times \{ j \} \in \Omega \) and \( Ux = \sum_{j \in J} U_j x_j \) if \( x = \sum_{j \in J} x_j \) and \( x_j \in X_j \). It can be seen that \( \Omega, A, \mu, h \) and \( U \) satisfy the assumptions. \( \square \)

\(^1\)A positive measure \( \mu \) is called regular, if

\[
\mu(B) = \inf \{ \mu(O) : B \subseteq O, \text{ } O \text{ is open} \} = \sup \{ \mu(K) : K \subset B, K \text{ is compact} \}
\]

holds for all \( B \) in the corresponding Borel sigma algebra.
4.3. THE SPECTRAL THEOREMS

We add an observation to the above proof. Let \( \lambda \in \sigma(T) \setminus \sigma_p(T) \). Then \( \Lambda = \bigcup_{j \in J} \{ \lambda \times \{ j \} \} \subset \Omega \) is a \( \mu \)-null set. In fact, otherwise the characteristic function \( f \) of \( \Lambda \) would be nonzero and so \( x = U^{-1}f \neq 0 \) would be an eigenvector of \( T \) for the eigenvalue \( \lambda \), since \( Tx = U^{-1}\lambda f = \lambda x \). We further note that in the proof of Theorem 4.31 one could also take \( \Omega = \bigcup_{j \in J} \sigma(T_j) \times \{ j \} \), where \( T_j = T|_{X_j} \). It can be shown that \( \sigma(T) = \bigcup_{j \in J} \sigma(T_j) \).

The above representation of bounded selfadjoint operators as multiplication operators now leads to a multiplication representation and to a \( B_b \)-functional calculus for (possibly) unbounded selfadjoint operators \( A \). Here \( B_b(\sigma(A)) \) is the Banach space of bounded Borel functions on \( \sigma(A) \) endowed with the supremum norm. We use this space instead of \( L^\infty(\sigma(A)) \) to avoid certain technical problems.

**Theorem 4.32** (Unbounded \( A \)). Let \( A \) be a closed, densely defined and selfadjoint operator on a separable Hilbert space \( X \). Then the following assertions hold.

a) There is a measure space \( (\Omega, A, \mu) \), a measurable function \( h : \Omega \to \sigma(A) \) and a unitary operator \( U : X \to L^2(\mu) \) such that

\[
D(A) = \{ x \in X : hUx \in L^2(\mu) \} \quad \text{and} \quad Ax = U^{-1}hUx.
\]

b) There is a contractive map \( \Psi : B_b(\sigma(A)) \to \mathcal{L}(X) \), \( \Psi(f) = f(A) \), satisfying (C1) and (C3)-(C5) where “\( p_1(T) = T \)” in (C3) is replaced by \( r_\lambda(A) = R(\lambda, A) \) for \( r_\lambda(z) = (\lambda - z)^{-1} \) and every \( \lambda \in \rho(A) \). Moreover, if \( f_n \in B_b(\sigma(A)) \) are uniformly bounded and converge to \( f \in B_b(\sigma(A)) \) pointwise, then \( f_n(A)x \to f(A)x \) as \( n \to \infty \) for all \( x \in X \). Finally, for \( x \in D(A) \) and \( f \in B_b(\sigma(A)) \) we have \( f(A)x \in D(A) \) and \( Af(A)x = f(A)Ax \).

**Proof.** a) Let \( t \in \rho(A) \cap \mathbb{R} \). (If \( \sigma(A) = \mathbb{R} \), take \( t = i \) and use the version of Theorem 4.31 for the normal operator \( R(i, A) \), see e.g. Satz VII.1.25 in [Wer05].) Then \( R(t, A) \in \mathcal{L}(X) \) is selfadjoint and can be represented as \( R(t, A) = U^{-1}mU \) on a space \( L^2(\Omega, \mu) \) as in Theorem 4.31. Recall that Proposition 1.20 yields \( \sigma(A) = t - \{ \sigma(R(t, A)) \setminus \{ 0 \} \}^{-1} \). Set

\[
h(\lambda \times \{ j \}) = t - \frac{1}{m(\lambda \times \{ j \})} = t - \frac{1}{\lambda} \in \sigma(A),
\]

for \( j \in J \) and \( \lambda \in \sigma(R(t, A)) \setminus \{ 0 \} \). The sets \( \{ 0 \times \{ j \} \} \) have \( \mu \)-measure 0 in view of the observation above the theorem, due to the injectivity of \( R(t, A) \). We can thus extend \( h \) by 0 to a measurable function on \( \Omega \). Let \( x \in D(A) \). We put \( y = tx - Ax \in X \). Since \( x = R(t, A)y = U^{-1}mUy \), we obtain

\[
hUx = hmUy = (tm - \mathbb{1})Uy \in L^2(\mu) \quad \text{(with equalities a.e.)},
\]

\[
U^{-1}hUx = tuU^{-1}mUy - y = tx - y = Ax.
\]

If \( x \in X \) satisfies \( hUx \in L^2(\mu) \), then we put \( y = U^{-1}(t\mathbb{1} - h)Ux \in X \) and obtain \( mUy = (tm - mh)Ux = Ux \). Therefore, \( x = U^{-1}mUy = R(t, A)y \in D(A) \) and part a) is proved.
b) We define a functional calculus $\Psi : f \mapsto f(A)$ for $A$ by setting
\[ f(A)x = U^{-1}((f \circ h)U)x \]
for $f \in B_b(\sigma(A))$ and $x \in X$. We further put $M_f \varphi = (f \circ h)\varphi$ for $\varphi \in L^2(\mu)$. It is straightforward to check that $f(A) \in \mathcal{L}(X)$, $\Psi$ is linear, $1(A) = I$ and (C5) is true. Let $\lambda \in \rho(A)$. We have $hUr_\lambda(A)x = h(r_\lambda \circ h)Ux = h(\lambda I - h)^{-1}Ux \in L^2(\mu)$ for all $x \in X$. So part a) yields that $r_\lambda(A)X \subset D(A)$ and $(\lambda I - A)r_\lambda(A) = I$. Similarly, one sees that $r_\lambda(A)(\lambda x - Ax) = x$ for all $x \in D(A)$, and so (C3) is shown. The contractivity and property(C4) hold since $\|f\| = \|M_f\| \leq \|f\|_\infty$ (cf. Example 1.54 in [FA]) and
\[ (fg)(A)x = U^{-1}(f \circ h)(g \circ h)Ux = U^{-1}(f \circ h)UU^{-1}(g \circ h)Ux = f(A)g(A)x, \]
where $g \in B_b(\sigma(A))$ and $x \in X$.

Let $f, f_n \in B_b(\sigma(A))$ be uniformly bounded by $c$ such that $f_n \to f$ pointwise as $n \to \infty$. For every $x \in X$, we have $f_n(A)x - f(A)x = U^{-1}((f_n - f) \circ h)Ux$. Since $(f_n - f) \circ h \to 0$ pointwise and $\|(f_n - f) \circ h)Ux\| \leq 2c\|Ux\|$, Lebesgue’s convergence theorem shows that $((f_n - f) \circ h)Ux$ tends to $0$ in $L^2(\mu)$ and so $f_n(A)x \to f(A)x$ in $X$ as $n \to \infty$.

Let $x \in D(A)$. The above results yield that
\[ g := h(f \circ h)Ux = (f \circ h)UU^{-1}hUx = (f \circ h)UAx \in L^2(\mu). \]
On the other hand, $g = hUU^{-1}(f \circ h)Ux = hUf(A)x$. Part a) thus implies that $f(A)x \in D(A)$ and
\[ Af(A)x = U^{-1}hUf(A)x = U^{-1}g = U^{-1}(f \circ h)UAx = f(A)Ax. \]

**Example 4.33** (Schrödinger’s equation). Let $X$ be a Hilbert space and $H$ be a closed, densely defined and selfadjoint operator on $X$. For a given $x \in D(H)$ we look for functions $u \in C^1(\mathbb{R}, X) \cap C(\mathbb{R}, \{D(H)\})$ satisfying
\[ \frac{d}{dt} u(t) = -iHu(t), \quad t \in \mathbb{R}, \quad u(0) = x. \]
($H$ is called Hamiltonian.) An example for this setting is $X = L^2(\mathbb{R}^3)$ and $H = -(-\Delta + \frac{b}{|x|^2})$ with $D(H) = W^2_2(\mathbb{R}^3)$, see Example 4.24. For $t \in \mathbb{R}$, we consider the bounded function $f_t(\xi) = e^{-it\xi}$, $\xi \in \mathbb{R}$. Theorem 4.32 allows us to define $U(t) = f_t(H) \in \mathcal{L}(X)$. From (C4) and (C5) it follows that
\[ U(t)U(s) = (f_t f_s)(H) = f_{t+s}(H) = U(t+s) = U(s)U(t), \]
\[ U(0) = f_0(H) = I, \]
\[ U(t)' = f_t'(H) = f_{-t}(H) = U(-t). \]
Taking $s = -t$, we see that there exists $U(t)^{-1} = U(-t) = U(t)'$, and so $U(t)$ is unitary. Since $\|f_t\|_1 = 1$ and $t \mapsto f_t(\xi)$ is continuous for all $t, \xi \in \mathbb{R}$, the function $\mathbb{R} \ni t \mapsto U(t)z \in X$ is continuous with $U(0)z = z$ for each $z \in X$. 

74
4.3. THE SPECTRAL THEOREMS

Let \( x \in D(H) \). We set \( y = \tau x - Hx \) for some \( \tau \in \rho(H) \), so that \( x = R(\tau, H)y = r_\tau(H)y \). We then obtain

\[
\frac{1}{t - s} (U(t)x - U(s)x) = \frac{1}{t - s} (f_t(H) - f_s(H))r_\tau(H)y = \left( \frac{1}{t - s}(f_t - f_s)r_\tau \right)(H)y =: g_{t,s}(H)y
\]

for all \( t \neq s \). Observe that \( g_{t,s}(\xi) \to \frac{\xi}{t - s}f_s(\xi) =: m(\xi)f_s(\xi) \) as \( t \to s \) for all \( \xi \in \sigma(H) \) and \( \|g_{t,s}\|_\infty \leq \|m\|_\infty = \sup_{\xi \in \sigma(H)} \frac{1}{|t - \xi|} < \infty \) for all \( t \neq s \). So Theorem 4.32 shows that there exists

\[
\frac{d}{dt} U(t)x = m(H)U(t)(\tau x - Hx) = U(t)m(H)(\tau x - Hx).
\]

Moreover, \( m(T)y = U(m \circ h)U^{-1}y = U((-ip_1r_\tau \circ h)U^{-1}y = -UihU^{-1}U(r_\tau \circ h)U^{-1}y = -iHR(\tau, H)y = -iHx \) due to Theorem 4.32. Hence, we arrive at

\[
\frac{d}{dt} U(t)x = -iU(t)Hx = -HU(t)x,
\]

using Theorem 4.32 once more. Due to these equations, \( u = U(\cdot)x \) belongs to \( C^1(\mathbb{R}_+, X) \cap C(\mathbb{R}_+, [D(H)]) \) and solves (4.7).

Let \( v \in C^1(\mathbb{R}_+, X) \cap C(\mathbb{R}_+, [D(H)]) \) be another solution of (4.7). For \( 0 \leq s \leq t \) and \( h \neq 0 \) we compute

\[
\frac{1}{h} (U(t - s - h)v(s + h) - U(t - s)v(s)) - U(t - s)(v'(s) + iHv(s)) = U(t - s - h) \left( \frac{1}{h} (v(s + h) - v(s)) - v'(s) \right) + (U(t - s - h) - U(t - s))v'(s) - \frac{1}{h} (U(t - s - h) - U(t - s))v(s) - U(t - s)iHv(s) \to 0
\]

as \( h \to 0 \). For any \( y \in X \) we thus obtain \( \frac{d}{ds} (U(t - s)v(s)|y) = 0 \), since \( v' = -iHv \). Consequently

\[
(U(t)x|y) = (U(t)v(0)|y) = (U(0)v(t)|y) = (v(t)|y),
\]

which gives \( u(t) = v(t) \) for all \( t \geq 0 \). Thus the “strongly continuous unitary group” \( (U(t))_{t \in \mathbb{R}} \) solves (4.7) uniquely.
Chapter 5

Holomorphic functional calculus

We come back to the case of Banach spaces $X$ and $Y$.

5.1 The bounded case

Let $U \subseteq \mathbb{C}$ be open, $f : U \to Y$ be holomorphic (i.e., complex differentiable) and $\Gamma \subseteq U$ be a piecewise $C^1$ curve with parametrization $\gamma : [a, b] \to U$. We define the curve integral

$$\int_{\Gamma} f \, dz = \int_{a}^{b} f(\gamma(t))\gamma'(t) \, dt$$

as a Banach space-valued Riemann integral (having the same definition, results and proofs as for $Y = \mathbb{R}$ in Analysis 1). It also holds $T \int_{\Gamma} f \, dz = \int_{\Gamma} T f \, dz$ for all $T \in \mathcal{L}(Y, Z)$. The index of a closed curve (i.e., $\gamma(a) = \gamma(b)$) is given by

$$n(\Gamma, z) = \frac{1}{2\pi i} \int_{\Gamma} \frac{dw}{w - z}$$

for all $z \in \mathbb{C} \setminus \Gamma$.

If $\Gamma$ is closed, $Y = \mathbb{C}$ and $n(\Gamma, z) = 0$ for all $z \notin U$, then Cauchy’s integral theorem and formula

(5.1) \[ \int_{\Gamma} f \, dz = 0, \]

(5.2) \[ \frac{1}{2\pi i} \int_{\Gamma} \frac{f(w)}{w - z} \, dw = n(\Gamma, z) f(z), \]

hold for all $z \in U \setminus \Gamma$. (This is the so-called homologuous version of these results, see e.g. Theorem IV.5.4 and IV.5.7 in [Con78].) For a general Banach space (5.1) and (5.2) hold for the functions $z \mapsto \langle f(z), y^* \rangle$ for every $y^* \in Y^*$. Hence, we have e.g. $\int_{\Gamma} f \, dz, y^* = 0$ for all $y^* \in Y^*$ so that (5.1) holds in $Y$ due to a corollary of the Hahn-Banach theorem. Similarly one deduces (5.2) for a general Banach space $Y$.

For a compact $K \subseteq \mathbb{C}$ we introduce the space

$$H(K) = \{ f : \text{there exists an open set } D(f) \subseteq \mathbb{C} \text{ such that } K \subseteq D(f) \}$$
5.1. THE BOUNDED CASE

and \( f : D(f) \to \mathbb{C} \) is holomorphic.

Let \( K \subseteq U \subseteq \mathbb{C} \), \( K \) be compact, and \( U \) be open. By Proposition VIII.1.1 in [Con78] and its proof there exists an admissible curve \( \Gamma \) for \( K \) and \( U \) (or, in \( U \setminus K \)) which means that \( \Gamma \subseteq U \setminus K \) is piecewise \( C^1 \), \( n(\Gamma, z) = 1 \) for all \( z \in K \) and \( n(\Gamma, z) = 0 \) for all \( z \in \mathbb{C} \setminus U \). Here it is allowed that \( \Gamma = \Gamma_1 \cup \cdots \cup \Gamma_n \) is the finite union of closed curves \( \Gamma_j \), where we set

\[
\int_\Gamma f\,dz = \int_{\Gamma_1} f\,dz + \cdots + \int_{\Gamma_n} f\,dz.
\]

Let \( T \in \mathcal{L}(X) \), \( f \in H(\sigma(T)) \), and \( \Gamma \) be admissible for \( \sigma(T) \) and \( D(f) \). We then define

\[
(5.3) \quad f(T) := \frac{1}{2\pi i} \int_\Gamma f(\lambda)R(\lambda, T)d\lambda \in \mathcal{L}(X).
\]

This integral exists in the Banach space \( \mathcal{L}(X) \) since \( \lambda \mapsto f(\lambda)R(\lambda, T) \) is holomorphic on \( \rho(T) \cap D(f) \supseteq \Gamma \). Writing \( R(\lambda, T) \) as \( \frac{1}{\lambda - \frac{1}{w-z}} \) one sees the similarity of (5.3) and (5.2), but here \( R(\lambda, T) \) does not exist on \( \sigma(T) \), whereas in (5.2) the function \( w \mapsto \frac{1}{w-z} \) is defined on \( \mathbb{C} \setminus \{z\} \).

If \( \Gamma' \) is another admissible curve for \( \sigma(T) \) and \( D(f) \), then we set \( \Gamma'' = \Gamma \cup (-\Gamma') \), where \( "-" \) denotes the inversion of the orientation. We then have

\[
n(\Gamma'', z) = n(\Gamma, z) - n(\Gamma', z) = \begin{cases} 1 - 1 = 0, & z \in \sigma(T), \\ 0 - 0 = 0, & z \in \mathbb{C} \setminus D(f). \end{cases}
\]

So we can apply (5.1) on \( U = D(f) \setminus \sigma(T) \) obtaining

\[
0 = \int_{\Gamma''} f(\lambda)R(\lambda, T)d\lambda = \int_\Gamma f(\lambda)R(\lambda, T)d\lambda - \int_{\Gamma'} f(\lambda)R(\lambda, T)d\lambda.
\]

Consequently, (5.3) does not depend on the choice of the admissible curve. We recall that \( r_\lambda(z) = \frac{1}{\lambda - z} \) and \( p_1(z) = z \) for \( \lambda, z \in \mathbb{C} \) with \( \lambda \neq z \).

**Theorem 5.1.** Let \( T \in \mathcal{L}(X) \) for a Banach space \( X \). Then the map

\[
\Phi : H(\sigma(T)) \to \mathcal{L}(X), \quad f \mapsto f(T),
\]

defined by (5.3) is linear and satisfies

\[
(P1) \quad \|f(T)\| \leq c(\Gamma) \sup_{\lambda \in \Gamma}|f(\lambda)| \quad \text{for a constant } c(\Gamma) > 0,
\]

\[
(P2) \quad 1(T) = I, \quad p_1(T) = T, \quad r_\lambda(T) = R(\lambda, T),
\]

\[
(P3) \quad f(T)g(T) = g(T)f(T) = (fg)(T),
\]

\[
(P4) \quad f(T)^* = f(T^*),
\]

\[
(P5) \quad \text{if } f_n \to f \text{ uniformly on compact subsets of } D(f), \text{ then } f_n(T) \to f(T) \text{ in } \mathcal{L}(X) \text{ as } n \to \infty,
\]

77
for all \( \lambda \in \rho(T) \) and \( f, g, f_n \in H(\sigma(T)) \) with \( D(f_n) = D(f) \) for every \( n \in \mathbb{N} \).

\( \Phi \) is the only linear map from \( H(\sigma(T)) \) to \( \mathcal{L}(X) \) satisfying (P1)-(P3). For a polynomial \( p, \) the operators \( p(T) \) in (5.3) and in (4.6) coincide.

**Proof.** It is clear that \( f \mapsto f(T) \) is linear. Property (P1) follows from

\[
\|f(T)\| \leq \frac{1}{2\pi} \int_{\Gamma} f(\lambda) \sup_{\lambda \in \Gamma} \|R(\lambda, T)\| \sup_{\lambda \in \Gamma} |f(\lambda)| =: c(\Gamma) \sup_{\lambda \in \Gamma} |f(\lambda)|.
\]

Replacing here \( f(T) \) by \( f(T) - f_n(T) = (f - f_n)(T) \) we also deduce (P5). To check (P4), we recall that \( \sigma(T) = \sigma(T^*) \) and \( R(\lambda, T)^* = R(\lambda, T^*) \) from Theorem 1.24. Hence,

\[
f(T)^* = \frac{1}{2\pi i} \int_{\Gamma} f(\lambda) R(\lambda, T)^* d\lambda = f(T^*).
\]

We next show (P3). We choose a bounded open set \( U \subseteq \mathbb{C} \) such that \( \sigma(T) \subseteq U \subseteq \overline{U} \subseteq D(f) \cap D(g) \) and admissible curves \( \Gamma_f \) in \( U \setminus \sigma(T) \) and \( \Gamma_g \) in \( (D(f) \cap D(g)) \setminus \overline{U} \). We then have \( n(\Gamma_f, \mu) = 0 \) for all \( \mu \in \Gamma_g \subseteq \mathbb{C} \setminus U \) and \( n(\Gamma_g, \lambda) = 1 \) for all \( \lambda \in \Gamma_f \subseteq \overline{U} \). Using the resolvent equation, Fubini’s theorem in \( \mathcal{L}(X) \) (see e.g. Theorem X.6.16 in [AmE08]) and (5.2) in \( \mathbb{C} \), we compute

\[
f(T)g(T) = \frac{1}{2\pi i} \int_{\Gamma} f(\lambda) R(\lambda, T) \frac{1}{2\pi i} \int_{\Gamma_g} g(\mu) R(\mu, T) d\mu d\lambda
\]

\[
= \left( \frac{1}{2\pi i} \right)^2 \int_{\Gamma_f} \int_{\Gamma_g} f(\lambda) g(\mu) \frac{1}{\mu - \lambda} (R(\lambda, T) - R(\mu, T)) d\mu d\lambda
\]

\[
= \frac{1}{2\pi i} \int_{\Gamma_f} f(\lambda) R(\lambda, T) \frac{1}{2\pi i} \int_{\Gamma_g} g(\mu) \frac{1}{\mu - \lambda} d\mu d\lambda
\]

\[
+ \frac{1}{2\pi i} \int_{\Gamma_g} g(\mu) R(\mu, T) \frac{1}{2\pi i} \int_{\Gamma_f} f(\lambda) \frac{1}{\lambda - \mu} d\mu
\]

\[
= \frac{1}{2\pi i} \int_{\Gamma_f} f(\lambda) g(\lambda) R(\lambda, T) d\lambda = (fg)(T).
\]

This identity also yields \( (fg)(T) = (gf)(T) = g(T)f(T) \).

To check (P2), we take \( f = 1 \) with \( D(f) = \mathbb{C} \). We choose \( \Gamma_0 = \partial B(0, 2\|T\|) \). Theorem 1.16 then leads to

\[
1(T) = \frac{1}{2\pi i} \int_{\Gamma_0} R(\lambda, T) d\lambda = \frac{1}{2\pi i} \int_{\Gamma_0} \sum_{n=0}^{\infty} T^n \lambda^{-n-1} d\lambda
\]

\[
= \sum_{n=0}^{\infty} T^n \frac{1}{2\pi i} \int_{\Gamma_0} \lambda^{-n-1} d\lambda = I,
\]

since the series converges in \( \mathcal{L}(X) \) uniformly on \( \Gamma_0 \) and \( \int_{\Gamma_0} \lambda^{-m} d\lambda \) is equal to \( 2\pi i \) if \( m = 1 \) and equal to 0 for \( m \in \mathbb{Z} \setminus \{1\} \). The property \( p_1(T) = T \) is shown similarly. For \( \lambda \in \rho(T) \), consider \( f_\lambda(z) = \lambda - z \). The previous results imply \( f_\lambda(T) = \lambda I - T \) and \( f_\lambda(T)r_\lambda(T) = r_\lambda(T)f_\lambda(T) = (r_\lambda f_\lambda)(T) = 1(T) = I \) so that \( r_\lambda(T) = R(\lambda, T) \).
5.1. THE BOUNDED CASE

Let \( \Psi : H(\sigma(T)) \rightarrow \mathcal{L}(X) \) be any linear map satisfying the assertions (P1)-(P3). The linearity, (P2) and (P3) imply that \( \Psi(p) = p(T) = a_0I + a_1T + \ldots + a_mT^m \) for every polynomial \( p \) with \( p(z) = a_0 + a_1z + \ldots + a_mz^m \). If \( f \in H(\sigma(T)) \) and \( \Gamma \subseteq D(f) \setminus \sigma(T) \) is admissible, then there are polynomials \( p_n \) converging uniformly to \( f \) on the compact set \( \Gamma \). Hence, \( p_n(T) \rightarrow \Psi(f) \) by (P1) and thus \( \Psi = \Phi \). □

Example 5.2. Let \( E = C(K) \) for a compact set \( K \subseteq \mathbb{R}^d \) and let \( m \in C(K) \). We define \( M\varphi = m\varphi \) for \( \varphi \in E \). Proposition 1.14 shows that \( M \in \mathcal{L}(E) \), \( \sigma(M) = m(K) \), and \( R(\lambda, M)\varphi = \frac{1}{\lambda - m}\varphi \) for all \( \lambda \in \rho(M) \).

Let \( f \in H(m(K)) \), \( \Gamma \) be an admissible curve in \( D(f) \setminus m(K) \), \( \varphi \in E \), and \( x \in K \).

Using that the map \( \psi \mapsto \psi(x) \) is continuous and linear from \( E \) to \( \mathbb{C} \) and Cauchy’s formula (5.2) for \( \mu = m(x) \in m(K) \), we compute

\[
(f(M)\varphi)(x) = \frac{1}{2\pi i} \int_{\Gamma} f(\lambda)R(\lambda, M)\varphi d\lambda(x) = \frac{1}{2\pi i} \int_{\Gamma} f(\lambda)(R(\lambda, M)\varphi)(x)d\lambda
\]

\[
= \frac{1}{2\pi i} \int_{\Gamma} f(\lambda)\frac{1}{\lambda - m(x)}d\lambda \varphi(x) = f(m(x))\varphi(x).
\]

As a result, \( f(M)\varphi = (f \circ m)\varphi \) is also a multiplication operator.

Theorem 5.3. Let \( T \in \mathcal{L}(X) \) and \( f \in H(\sigma(T)) \). Then the spectral mapping theorem holds:

\[
\sigma(f(T)) = f(\sigma(T)).
\]

Proof. Let \( \mu \notin f(\sigma(T)) \). Then \( g = \frac{1}{\mu - f(T)} \in H(\sigma(T)) \) and \( g(T)(\mu I - f(T)) = (\mu I - f(T))g(T) = (g(\mu I - f(T)) = I \) so that \( \mu \in \rho(f(T)) \).

Conversely, let \( \mu = f(\lambda) \) for some \( \lambda \in \sigma(T) \). We set \( g(z) = \frac{f(z) - f(\lambda)}{z - \lambda} \) for \( z \in D(f) \setminus \{\lambda\} \) and \( g(\lambda) = f'(\lambda) \). Since \( f \in C^2 \) by its holomorphy, \( g \) is holomorphic on \( D(f) \) and \( g(z)(z - \lambda) = f(z) - \mu \) for all \( z \in D(f) \). Suppose that \( \mu \in \rho(f(T)) \). Since \( g(T)(\lambda I - T) = (\lambda I - T)g(T) = (g(\lambda I - T - p_I))(T) = (\mu I - f(T)) = \mu I - f(T) \),

\( (\lambda I - T) \) would then have the inverse \( g(T)(\mu I - f(T))^{-1} \) which is impossible. Hence, \( \mu \in \sigma(f(T)) \). □

Example 5.4. Let \( A \in \mathcal{L}(X) \). For \( t \in \mathbb{R} \) we set \( f_t(z) = e^{tz} \) and define \( e^{tA} = f_t(A) \in \mathcal{L}(X) \). As in Example 4.33, one sees that \( e^{(t+s)A} = e^{tA}e^{sA} = e^{sA}e^{tA} \), \( e^{0A} = I = e^{tA}e^{-tA} \) for all \( t, s \in \mathbb{R} \). Moreover, \( t \mapsto e^{tA} \) belongs to \( C^1(\mathbb{R}, \mathcal{L}(X)) \) with \( \frac{d}{dt} e^{tA} = Ae^{tA} = e^{tA}A \) and the map \( u(t) = e^{tA}x \) is the unique solution in \( C^1(\mathbb{R}, X) \) of

\[
\frac{d}{dt} u(t) = Au(t) \quad (t \in \mathbb{R}), \quad u(0) = x,
\]

where \( x \in X \) is given. (See also the exercises.) Theorem 5.3 further yields

\[
r(e^{tA}) = \max\{|\mu| : \mu \in \sigma(e^{tA}) = e^{t\sigma(A)}\} = \max\{e^{tRe\lambda} : \lambda \in \sigma(A)\} = e^{ts(A)},
\]

79
where \( s(A) := \max \{ \Re \lambda : \lambda \in \sigma(A) \} \) is the spectral bound of \( A \). Now, if \( s(A) < 0 \)
(i.e. \( \sigma(A) \subseteq \{ \lambda : \Re \lambda < 0 \} \)), then we deduce from Theorem 1.16 that
\[
1 > r(e^A) = \lim_{n \to \infty} \|(e^A)^n\|^{1/n} = \lim_{n \to \infty} \|e^{nA}\|^{1/n}.
\]
So we can fix some \( N \in \mathbb{N} \) such that \( \|e^{NA}\| =: q < 1 \). Writing any given \( t \geq 0 \) as
\( t = hN + \tau \) for some \( h \in \mathbb{N}_0 \) and \( 0 < \tau \leq N \) we estimate
\[
\|e^{tA}\| = \|e^{NA}\|^h |e^{\tau A}| \leq q^h \|e^{\tau A}\| \leq \max_{0 \leq \tau \leq N} \|e^{\tau A}\| \exp(Nh \frac{\ln q}{N}) \leq Me^{-wt} \to 0
\]
as \( t \to \infty \), where \( w := -\frac{\ln q}{N} > 0 \) and \( M := \max_{0 \leq \tau \leq N} \|e^{\tau A}\|^{\ln q} \). So spectral information on the given operator \( A \) implies the exponential decay \( \|u(t)\| \leq Me^{-wt}\|x\| \) of the solutions \( u \).

**Remark.** Let \( S \in \mathcal{L}(X) \) and \( P = P^2 \in \mathcal{L}(X) \) be a projection with \( SP = PS \). Set \( X_1 = R(P) \) and \( X_2 = N(P) \). Then, \( X = X_1 \oplus X_2 \) by e.g. Lemma 1.63 of [FA]. Moreover, if \( y = Px \in R(P) \), then \( Sy = SPx = PSx \in R(P) \). If \( x \in N(P) \), then \( PSx = SPx = 0 \) and so \( Sx \in N(P) \). As a result, \( SX_j \subseteq X_j \) and the restrictions \( S_j|_{X_j} \in \mathcal{L}(X_j) \) are well defined for \( j = 1, 2 \).

**Theorem 5.5 (Spectral projection).** Let \( T \in \mathcal{L}(X) \) and \( \sigma(T) = \sigma_1 \cup \sigma_2 \) for two disjoint closed sets \( \sigma_j \neq \emptyset \) in \( \mathbb{C} \). Then there is a projection \( P \in \mathcal{L}(X) \) with \( f(T)P = Pf(T) \) for all \( f \in H(\sigma(T)) \) such that \( \sigma(T_j) = \sigma_j \) for \( j = 1, 2 \), where \( T_j = T|_{X_j} \in \mathcal{L}(X_j) \), \( X_1 = R(P) \) and \( X_2 = N(P) \). Moreover, \( X = X_1 \oplus X_2 \) and \( R(\lambda, T_j) = R(\lambda, T)|_{X_j} \) for \( \lambda \in \rho(T) = \rho(T_1) \cap \rho(T_2) \). We further have
\[
(5.4) \quad P = \frac{1}{2\pi i} \int_{\Gamma_1} R(\lambda, T) d\lambda,
\]
where \( \Gamma_1 \) is an admissible curve for \( \sigma_1 \) and any open set \( U_1 \supseteq \sigma_1 \) such that \( \overline{U_1} \cap \sigma_2 = \emptyset \).

**Proof.** There are open sets \( U_j \) with \( \overline{U_1} \cap \overline{U_2} = \emptyset \) and \( \sigma_j \subseteq U_j \) for \( j = 1, 2 \). Define \( h \in H(\sigma(T)) \) by \( h = 1 \) on \( U_1 \) and \( h = 0 \) on \( U_2 \). We set \( \tilde{P} = h(T) \in \mathcal{L}(X) \). We then deduce \( P^2 = h^2(T) = h(T) = P \) and \( f(T)P = Pf(T) \) for all \( f \in H(\sigma(T)) \) from (P3) and (P2). The above remark shows that \( X = X_1 \oplus X_2 \) holds and that the operators \( T_j = T|_{X_j} \in \mathcal{L}(X_j) \) are well defined.

The formula (5.4) follows by choosing \( \Gamma = \Gamma_1 \cup \Gamma_2 \), where \( \Gamma_j \) are admissible curves for \( \sigma_j \) and \( U_j \) for \( j = 1, 2 \). Let \( \lambda \notin \sigma_1 \). We may shrink \( U_1 \) so that \( \lambda \notin U_1 \) since \( P \) does not depend on the choice of \( \Gamma \) and thus not on the choice of \( U_1 \). We define \( g(z) = \frac{1}{\lambda z} \) for \( z \in U_1 \) and \( g(z) = 0 \) for \( z \in U_2 \). Then \( g \in H(\sigma(T)) \) and
\[
g(T)(\lambda I - T) = (\lambda I - T)g(T) = ((\lambda I - P_1)g(T) = h(T) = P.
\]
Setting \( R = g(T)|_{X_1} \in \mathcal{L}(X_1) \), we thus obtain
\[
R(\lambda I_{X_1} - T_1) = (\lambda I_{X_1} - T_1)R = I_{X_1}.
\]

80
This means that $\lambda \in \rho(T_1)$, and so $\sigma(T_1) \subseteq \sigma_1$. Similarly, one sees that $\sigma(T_2) \subseteq \sigma_2$. In particular, $\sigma(T_1)$ and $\sigma(T_2)$ are disjoint. Let $\lambda \in \rho(T_1) \cap \rho(T_2)$. Given $x \in X$, we have unique $x_1 \in X_1$ and $x_2 \in X_2$ such that $x = x_1 + x_2$. If $\lambda x - Tx = 0$, then $0 = \lambda x_1 - T_1 x_1 + \lambda x_2 - T_2 x_2 \in X_1 \oplus X_2$ so that $x_j \in N(\lambda I - T_j)$ for $j = 1, 2$ and so $x = 0$. Given $y \in X$, we define $x_j = R(\lambda, T_j)y_j \in X_j$ for $j = 1, 2$. Setting $x = x_1 + x_2$, we derive

$$\lambda x - Tx = \lambda x_1 - T_1 x_1 + \lambda x_2 - T_2 x_2 = y_1 + y_2 = y.$$ 

We have shown that $\lambda \in \rho(T)$, $R(\lambda, T)|_{X_j} = R(\lambda, T_j)$ and

$$\sigma(T) = \sigma_1 \cup \sigma_2 \subseteq \sigma(T_1) \cup \sigma(T_2) \subseteq \sigma_1 \cup \sigma_2,$$

which implies that $\sigma(T_j) = \sigma_j$ for $j = 1, 2$. \hfill \Box

**Example 5.6** (Exponential dichotomy). In the setting of Example 5.4 assume that $\sigma(A) \cap i\mathbb{R} = \emptyset$. Hence, $\sigma(A) = \sigma_1 \cup \sigma_2$ where $\sigma_1 \subseteq \{ \lambda \in \mathbb{C} : \text{Re}\, \lambda < 0 \}$ and $\sigma_2 \subseteq \{ \lambda \in \mathbb{C} : \text{Re}\, \lambda > 0 \}$. Let $P$ be the spectral projection of $A$ for $\sigma_1$ and define $A_1$ and $A_2$ as the restrictions of $A$ to $X_1 = R(P)$ and $X_2 = N(P)$, respectively, as in Theorem 5.5. Let $\Gamma_1$ be given as in Theorem 5.5. Observe that for $x \in X_1$ we have

$$e^{tA}x = e^{tA}P x = (f_i h)(A)x = \frac{1}{2\pi i} \int_{\Gamma_1} e^{t\lambda} R(\lambda, A) x d\lambda$$

$$\quad = \frac{1}{2\pi i} \int_{\Gamma_1} e^{t\lambda} R(\lambda, A_1) x d\lambda = e^{tA_1}x,$$

where $f_i(\lambda) = e^{t\lambda}$ for $t \in \mathbb{R}$ and $h$ is given by the proof of Theorem 5.5. In the same way one derives $e^{tA}x = e^{tA_2}x$ for all $x \in X_2$ and $t \geq 0$. Since $\sigma(A_1) = \sigma_1$, we have $s(A_1) < 0$ and so Example 5.4 shows that $\|e^{tA_1}x_1\| \leq Me^{-\omega t}\|x_1\|$ for all $t \geq 0$ and $x_1 \in X_1$ and some constants $M, \omega > 0$.

We further have $\sigma(A_2) = \sigma_2$ and so $s(-A_2) < 0$. Note that the curve $\tilde{\Gamma} = \{-\lambda : \lambda \in \Gamma\}$ is admissible for $\sigma(-A) = -\sigma(A)$. Substituting $\mu = -\lambda$, we conclude that

$$e^{-tA} = \frac{1}{2\pi i} \int_{\tilde{\Gamma}} e^{-t\lambda}(\lambda I - A)^{-1} d\lambda = \frac{1}{2\pi i} \int_{\Gamma} e^{t\mu}(\mu I - (-A))^{-1} d\mu = e^{t(-A)}$$

for all $t \in \mathbb{R}$. For $x_2 \in X_2$ we thus obtain

$$e^{-tA_2}x_2 = e^{-tA}x_2 = e^{t(-A)}x_2 = e^{t(-A_2)}x_2,$$

so that $\|e^{-tA_2}x_2\| \leq M'e^{-\omega' t}\|x_2\|$ for all $t \geq 0$ and some constants $M', \omega' > 0$. Summing up, we have decomposed $X$ into $e^{tA}$-invariant subspaces where $e^{tA}$ decays exponentially in forward and in backward time, respectively.

### 5.2 Sectorial operators

For $\phi \in (0, \pi]$ we define the open sector

$$\Sigma_\phi = \{ \lambda \in \mathbb{C} \setminus \{0\} : |\text{arg}\, \lambda| < \phi \}.$$ 

We also set $\Sigma_{\pi/2} = \mathbb{C}_-$. Note that $\Sigma_{\pi} = \mathbb{C} \setminus \mathbb{R}_-$. 

81
 CHAPTER 5. HOLOMORPHIC FUNCTIONAL CALCULUS

**Definition 5.7.** A closed operator $A$ is called sectorial of angle $\phi \in (0, \pi]$ if there is a constant $K > 0$ such that $\Sigma_{\phi} \subseteq \rho(A)$ and

$$
\|R(\lambda, A)\| \leq \frac{K}{|\lambda|}
$$

holds for all $\lambda \in \Sigma_{\phi}$. If there is an $\omega \in \mathbb{R}$ such that $A - \omega I$ is sectorial of angle $\phi$, then we say that $A$ is sectorial with constants $(\phi, K, \omega)$.

Note that a sectorial operator of angle $\phi$ is also sectorial of angle $\phi' \in (0, \phi)$.

**Example 5.8.** Let $X$ be a Hilbert space and $A$ be closed, densely defined and selfadjoint on $X$. We further suppose that $\sigma(A) \subseteq (-\infty, \omega]$ for some $\omega \in \mathbb{R}$. Then $A - \omega I =: A_\omega$ is selfadjoint with $\sigma(A_\omega) \subseteq \mathbb{R}^-$. Take $\phi \in (\frac{\pi}{2}, \phi)$ and $\lambda \in \Sigma_{\phi}$. Since $R(\lambda, A_\omega)' = R(\lambda, A_\omega)$, the operator $R(\lambda, A_\omega)$ is normal. Propositions 4.14 and 1.20 then yield

$$
\|R(\lambda, A_\omega)\| = r(R(\lambda, A_\omega)) = \max\{|\mu| : \mu \in (\lambda - \sigma(A_\omega))^{-1} \cup \{0\}\}
$$

$$
= d(\lambda, \sigma(A_\omega))^{-1} \lesssim \begin{cases} 1, & \text{Re} \lambda \geq 0, \\ \frac{1}{|\text{Im} \lambda|}, & \text{Re} \lambda < 0. \end{cases}
$$

If $\text{Re} \lambda < 0$, we can write $\lambda = |\lambda|e^{\pm i\theta}$ for some $\theta \in (\frac{\pi}{2}, \phi)$. We then have $|\text{Im} \lambda| = |\sin(\pi - \theta)| \geq \sin(\pi - \phi) > 0$, and thus

$$
\|R(\lambda, A_\omega)\| \leq \frac{1}{|\sin(\pi - \phi)|} =: \frac{K_\phi}{|\lambda|}
$$

for all $\lambda \in \Sigma_{\phi}$. As a result, $A$ is sectorial with constants $(\phi, K_\phi, \omega)$ for all $\phi < \pi$.

**Example 5.9.** Let $X = C([0, 1])$ and $A_k u = u''$ for $k = 0, 1$ with

$$
D(A_0) = \{u \in C^2([0, 1]) : u(0) = u(1) = 0\},
$$

$$
D(A_1) = \{u \in C^2([0, 1]) : u'(0) = u'(1) = 0\}.
$$

Let $\lambda \in \Sigma_{\pi}$ and $\lambda = \mu^2$ for some $\mu \in \mathbb{C}$ with $\text{Re} \mu > 0$. Let $f \in X$. As in Example 4.10, it holds

$$
u \in D(A_0), \lambda u - A_0 u = f
$$

if and only if $u \in C^2([0, 1]), u'' = \mu^2 u - f$ on $[0, 1], u(0) = u(1) = 0$.

If and only if

$$
u(t) = ae^{\mu t} + be^{-\mu t} + \frac{1}{2\mu} \int_0^1 e^{-\mu |t-s|} f(s) ds
$$

and

$$
u(0) = a + b + \frac{1}{2\mu} \int_0^1 e^{-\mu s} f(s) ds = 0
$$

and

$$
u(1) = ae^{\mu} + be^{-\mu} + \frac{e^{-\mu}}{2\mu} \int_0^1 e^{\mu s} f(s) ds = 0.
$$
5.2. SECTORIAL OPERATORS

The numbers \( a = a(f, \mu) \) and \( b = b(f, \mu) \) satisfy the two boundary conditions above if and only if

\[
\begin{pmatrix}
  a(f, \mu) \\
  b(f, \mu)
\end{pmatrix} = \frac{1}{e^{-\mu} - e^\mu} \begin{pmatrix}
  e^{-\mu} & -1 \\
  -e^\mu & 1
\end{pmatrix} \begin{pmatrix}
  -\frac{1}{2\mu} \int_0^1 e^{\mu s} f(s) ds \\
  -\frac{1}{2\mu} \int_0^1 e^{-\mu s} f(s) ds
\end{pmatrix}
\]

\[
= \frac{1}{2\mu(e^{-\mu} - e^\mu)} \begin{pmatrix}
  e^{-\mu} \int_0^1 (e^{\mu s} - e^{-\mu s}) f(s) ds \\
  e^\mu \int_0^1 (e^{\mu s} - e^{-\mu s}) f(s) ds
\end{pmatrix}.
\]

We thus obtain \( \lambda \in \rho(A_0) \) and

\[
R(\mu^2, A_0)f(t) = a(f, \mu)e^{\mu t} + b(f, \mu)e^{-\mu t} + \frac{1}{2\mu} \int_0^1 e^{-\mu|t-s|} f(s) ds
\]

for all \( \mu^2 = \lambda \in C \setminus \mathbb{R}_-, \) \( Re \mu > 0, \) \( f \in X \) and \( t \in [0, 1] \).

Fix \( \phi \in (\pi, 2\pi) \). Take \( \lambda \in \Sigma_{\phi} \) and thus \( \mu \in \Sigma_{\phi}^{\pi} \) i.e., \( \mu = |\mu|e^{i\theta} \) with \( 0 \leq |\theta| < \phi/2 \).

We then obtain \( Re \mu = |\mu| \cos \theta \geq |\mu| \cos \pi/2 \). So we can estimate

\[
\| R(\lambda, A_0) f \| \leq |a(f, \mu)|e^{Re \mu} + |b(f, \mu)| + \| f \|_{L^1(\mathbb{R})} \sup_{t \in [0,1]} \int_{-t^{-1}}^t e^{-Re \mu r} dr
\]

\[
\leq \frac{1}{\lambda} \left( \int_0^1 (e^{Re \mu s} + e^{-Re \mu s}) ds + \frac{1}{2\mu} \int_0^1 e^{-Re \mu s} ds \right) + \frac{f(\infty)}{|\mu|Re \mu}
\]

\[
= \frac{1}{\lambda \mu} \left( (e^{Re \mu} - 1 + 1 - e^{-Re \mu}) + \frac{f(\infty)}{|\mu|Re \mu} \right)
\]

\[
\leq \frac{1}{\lambda \mu^2} \| f \| \left( \frac{(e^{Re \mu} - e^{-Re \mu}) + (e^{Re \mu} - e^{-Re \mu})}{2(e^{Re \mu} - e^{-Re \mu})} + 1 \right)
\]

\[
= \frac{1}{\lambda^{\frac{5}{2}}} \| f \| \| \mu \|_2.
\]

Hence, \( A_0 \) is sectorial for all angles \( \phi < \pi \), where \( K_\phi \to \infty \) as \( \phi \to \pi \).

As in Example 4.10 one sees that \( \sigma(A_0) = \sigma_0(A_0) = \{-\pi^2 k^2 : k \in \mathbb{N}\} \) with eigenfunctions \( u_k(t) = \sin(k\pi t) \). Note that \( D(A_0) = \{ u \in X : u(0) = u(1) = 0 \} \neq X \).

In a similar way, one shows that \( A_1 \) is sectorial for every angle \( \phi < \pi \) with \( \sigma(A_1) = \sigma_0(A_1) = \{-\pi^2 k^2 : k \in \mathbb{N}_{\geq 0}\} \) with eigenfunctions \( u_k(t) = \cos(k\pi t) \). Here \( D(A_1) \) is dense in \( X \). (See the exercises.)

**Example 5.10.** Let \( X = L^p(\mathbb{R}), \) \( 1 \leq p < \infty \), and \( Au = u' \) for \( D(A) = W_{o}^1(\mathbb{R}) \). In Example 4.2 we have seen that \( \sigma(A) = i\mathbb{R} \) and \( \| R(\lambda, A) \| \leq \frac{1}{Re \lambda} \) for \( Re \lambda > 0 \). If \( \phi \in (0, \pi/2) \) and \( \lambda \in \Sigma_{\phi} \) we have \( |Re \lambda| = \lambda \cos \phi \) so that \( A \) is sectorial for each angle \( \phi < \pi/2 \). If \( \lambda = s + i \) for \( s > 0 \), then \( \| R(s + i, A) \| \geq 1/s \) due to Theorem 1.13 and so \( A \) is not sectorial of angle \( \phi = \pi/2 \).
CHAPTER 5. HOLOMORPHIC FUNCTIONAL CALCULUS

Remark. Let \(1 < p < \infty\), \(U \subseteq \mathbb{R}^d\) be open with \(\partial U \in C^2\), and \(X = L^p(\mathbb{R}^d)\). The operators
\[
A_0 = \triangle, \quad D(A_0) = W^2_p(U) \cap W^1_p(U),
\]
\[
A_1 = \triangle, \quad D(A_1) = \{u \in W^2_p(U) : \partial_j u := \sum_{j=1}^d \nu_j T \partial_j u = 0 \text{ on } \partial U\},
\]
are sectorial on \(X\) with angle \(\phi > \pi/2\). Here \(T : W^1_p(U) \to L^p(\partial U)\) is the trace operator and \(\nu\) is the outer unit normal. There are variants for the spaces \(X = L^1(\mathbb{R}^d)\) and \(X = \mathcal{C}(\overline{U})\) as well as for more general differential operators and boundary conditions. See e.g. Chapter 3 in [Lun95] and also [Tan97].

Let \(A\) be sectorial of angle \(\phi \in (\pi/2, \pi)\) with constant \(K\). Take any \(r > 0\) and \(\theta \in (\pi/2, \phi)\). We define
\[
\Gamma_1 = \{\lambda = \gamma_1(s) = (-s)e^{-i\theta} : -\infty < s \leq -r\},
\]
\[
\Gamma_2 = \{\lambda = \gamma_2(\alpha) = re^{i\alpha} : -\theta \leq \alpha \leq \theta\},
\]
\[
\Gamma_3 = \{\lambda = \gamma_3(s) = se^{i\theta} : r \leq s < \infty\},
\]
\[
\Gamma = \Gamma(r, \theta) = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3.
\]
For \(t > 0\), we introduce the operator
\[
e^{tA} = \frac{1}{2\pi i} \int_{\Gamma} e^{t\lambda} R(\lambda, A) d\lambda = \lim_{R \to \infty} \frac{1}{2\pi i} \int_{\Gamma_R} e^{t\lambda} R(\lambda, A) d\lambda,
\]
where \(\Gamma_R = \Gamma \cap B(0, R)\) for \(R > r\). We first have to show that the limit in (5.5) exists in \(\mathcal{L}(X)\).

Lemma 5.11. Under the above assumptions, the integral in (5.5) converges absolutely in \(\mathcal{L}(X)\) and gives an operator \(e^{tA} \in \mathcal{L}(X)\) which does not depend on the choice of \(r > 0\) and \(\theta \in (\pi/2, \phi)\). Moreover, \(\|e^{tA}\| \leq M\) for all \(t > 0\) and a constant \(M = M(K, \theta) > 0\).

Proof. Since \(\|R(\lambda, A)\| \leq \frac{K}{|\lambda|}\) on \(\Gamma\), we can estimate
\[
\left| \int_{\Gamma_R} \|e^{t\lambda} R(\lambda, A)\| d\lambda \right| \leq K \int_{r}^{R} \frac{\exp(ts \Re e^{-i\theta})}{|se^{-i\theta}|} |e^{-i\theta}| ds
\]
\[
+ K \int_{-\theta}^{\theta} \frac{\exp(tr \Re e^{i\alpha})}{|re^{i\alpha}|} |re^{i\alpha}| d\alpha + K \int_{r}^{R} \frac{\exp(ts \Re e^{i\theta})}{|se^{i\theta}|} |e^{i\theta}| ds
\]
\[
\leq K \left(2 \int_{r}^{\infty} \frac{e^{sl \cos \theta}}{s} ds + \int_{-\theta}^{\theta} e^{tr \cos \alpha} d\alpha \right)
\]
\[
\leq K \left(2 \int_{rt \cos \theta}^{\infty} \frac{e^{-\sigma}}{\sigma} \frac{d\sigma}{-t \cos \theta} + 2\theta e^{tr} \right)
\]
\[
= : Kc(r, t, \theta),
\]
84
for all $R,t > 0$, where we substituted $\sigma = -st \cos \theta$. Thus the limit in (5.5) exists absolutely in $\mathcal{L}(X)$ by the majorant criterium, and $\|e^{tA}\| \leq KC(r,t,\theta)$. If we take $r = 1/t$, then $c(1/t, t, \theta) = c(\theta)$ does not depend on $t > 0$.

So it remains to check that the integral in (5.5) is independent of $r > 0$ and $\theta \in (\frac{\pi}{2}, \phi)$. To this aim, we define $\Gamma' = \Gamma(r', \theta')$ for some $r' > 0$ and $\theta' \in (\frac{\pi}{2}, \phi)$. We further set $\Gamma'' = \Gamma''(\emptyset, R)$ and choose $R > r, r'$. Let $C_R^+$ and $C_R^-$ be the circle arcs from the endpoint of $\Gamma_R$ to that of $\Gamma''_R$ in $\{\text{Im} \lambda > 0\}$ and $\{\text{Im} \lambda < 0\}$, respectively. (If $\theta = \theta'$, then $C_R^+$ contain just one point.) Then $S_R = \Gamma_R \cup C_R^+ \cup (-\Gamma_R) \cup (-C_R^-)$ is a closed curve in the starshaped domain $\Sigma$. So (5.1) shows that

$$
\int_{S_R} e^{i\lambda} R(\lambda, A)d\lambda = 0.
$$

Let $\theta = \min\{\theta, \theta'\} \in (\frac{\pi}{2}, \pi)$. It holds

$$
\|\int_{C_R^+} e^{i\lambda} R(\lambda, A)d\lambda\| \leq \left| \int_{\theta}^{\theta'} e^{iRRe^{i\alpha}} \frac{K}{|Re^{i\alpha}|} \right| RRe^{i\alpha} |d\alpha | \leq K|\theta' - \theta| e^{iR\cos \theta} \longrightarrow 0,
$$
as $R \to 0$. So we conclude that

$$
\int_{\Gamma} e^{i\lambda} R(\lambda, A)d\lambda = \lim_{R \to \infty} \int_{\Gamma_R} e^{i\lambda} R(\lambda, A)d\lambda = \lim_{R \to \infty} \int_{\Gamma''_R} e^{i\lambda} R(\lambda, A)d\lambda = \int_{\Gamma'} e^{i\lambda} R(\lambda, A)d\lambda,
$$
as asserted. \hfill \Box

**Theorem 5.12.** Let $A$ be sectorial of angle $\phi > \frac{\pi}{2}$. Define $e^{tA}$ as in (5.5) for $t > 0$, and set $e^{0A} = I$. Then the following assertions hold.

a) $e^{tA} e^{sA} = e^{sA} e^{tA} = e^{(t+s)A}$ for all $t, s \geq 0$.

b) The map $t \mapsto e^{tA}$ belongs to $C^1((0, \infty), \mathcal{L}(X))$. Moreover, $e^{tA} \lambda \subseteq D(A)$, \(\frac{d}{dt} e^{tA} = Ae^{tA} \) and $\|Ae^{tA}\| \leq C \|e^{tA}\|$ for a constant $C > 0$ and all $t > 0$. We also have $Ae^{tA} x = e^{tA} Ax$ for all $x \in D(A)$ and $t \geq 0$.

c) $e^{tA} x$ converges as $t \to 0$ if and only if $x \in D(A)$. In this case, it holds $e^{tA} x \to x$ as $t \to 0$.

**Proof.** a) Let $t,s > 0$. Take $0 < r < r'$ and $\frac{\pi}{2} < \theta' < \theta < \phi$. Set $\Gamma = \Gamma(r, \theta)$ and $\Gamma' = \Gamma(r', \theta')$. Using the resolvent equation and Fubini’s theorem, we compute

$$
e^{tA} e^{sA} = \frac{1}{(2\pi i)^2} \int_{\Gamma} e^{t\lambda} \int_{\Gamma'} e^{s\mu} R(\lambda, A)R(\mu, A)d\mu d\lambda
$$

$$
= \frac{1}{2\pi i} \int_{\Gamma} e^{t\lambda} R(\lambda, A) \frac{1}{2\pi i} \int_{\Gamma'} \frac{e^{s\mu}}{\mu - \lambda} d\mu d\lambda
$$

$$
+ \frac{1}{2\pi i} \int_{\Gamma'} e^{s\mu} R(\mu, A) \frac{1}{2\pi i} \int_{\Gamma} \frac{e^{t\lambda}}{\lambda - \mu} d\mu.
$$

85
CHAPTER 5. HOLOMORPHIC FUNCTIONAL CALCULUS

Fix \( \lambda \in \Gamma \) and take \( R > \max\{r, r', |\lambda|\} \). We define \( C'_R = \{z = Re^{i\alpha} : \theta \leq \alpha \leq 2\pi - \theta'\} \) and \( S'_R = \Gamma'_R \cup C'_R \). Since \( n(S'_R; \lambda) = 1 \), Cauchy’s formula (5.2) yields

\[
\frac{1}{2\pi i} \int_{S'_R} \frac{e^{s\mu}}{\mu - \lambda} d\mu = e^{s\lambda}.
\]

We further have

\[
\int_{\Gamma'_R} \frac{e^{s\mu}}{\mu - \lambda} d\mu \longrightarrow \int_{\Gamma'} \frac{e^{s\mu}}{\mu - \lambda} d\mu \quad \text{and}
\]

\[
|\int_{C'_R} \frac{e^{s\mu}}{\mu - \lambda} d\mu| \leq 2\pi R \sup_{\mu \in C'_R} \frac{e^{sR\cos \theta'}}{|\mu - \lambda|} \leq \frac{e^{sR\cos \theta'}}{R - |\lambda|} \longrightarrow 0
\]

as \( R \to \infty \). Consequently,

\[
e^{s\lambda} = \frac{1}{2\pi i} \int_{\Gamma'} \frac{e^{s\mu}}{\mu - \lambda} d\mu.
\]

Closing \( \Gamma_R \) with the circle arc \( C_R = \{z = Re^{i\alpha} : \theta \leq \alpha \leq 2\pi - \theta\} \), one verifies in the same way that

\[
0 = \int_{\Gamma} \frac{e^{\lambda \mu}}{\lambda - \mu} d\lambda.
\]

We thus conclude that

\[
e^{tA} e^{sA} = \frac{1}{2\pi i} \int_{\Gamma} \frac{e^{\lambda \mu}}{\lambda - \mu} R(\lambda, A) d\lambda = e^{(t+s)A} = e^{sA} e^{tA}.
\]

b) Let \( x \in X, \ t > 0, \ \varepsilon > 0, \) and \( R > r \). Observe that the Riemann sums for \( \int_{\Gamma_R} e^{t\lambda R(\lambda, A)} d\lambda \) converge in \([D(A)]\) since \( \lambda \mapsto R(\lambda, A) \) is continuous in \( L(X, [D(A)]) \). We thus obtain

\[
(5.6) \quad A \int_{\Gamma_R} e^{\lambda R(\lambda, A)} d\lambda = \int_{\Gamma_R} e^{t\lambda R(\lambda, A)} d\lambda = \int_{\Gamma_R} e^{\lambda R(\lambda, A)} d\lambda - \int_{\Gamma_R} e^{t\lambda d\lambda}.
\]

Let \( C_R = \{\mu = Re^{i\alpha} : \theta \leq \alpha \leq 2\pi - \theta\} \). Using (5.1), one shows as in part a) that

\[
|\int_{\Gamma_R} e^{t\lambda d\lambda}| = |- \int_{C_R} e^{t\lambda d\lambda}| \leq 2\pi R \sup_{\theta \leq \alpha \leq 2\pi - \theta} e^{tR \cos \alpha} \leq 2\pi Re^{t \cos \theta} \longrightarrow 0
\]

as \( R \to \infty \), uniformly for \( t \geq \varepsilon \). Moreover, as in the proof of Lemma 5.11 (with \( r = 1/i \)) we estimate

\[
|\int_{\Gamma_R} \|e^{t\lambda R(\lambda, A)}\| d\lambda| \leq K \left( 2 \int_{\frac{1}{t}}^{\infty} \frac{|s|}{|s|^t} e^{tR \cos \theta} ds + 1 \int_{-\theta}^{\theta} e^{\cos \alpha} d\alpha \right)
\]

\[
\leq \frac{2K}{t|\cos \theta|} + \frac{2eK \theta}{t} =: \frac{C}{t}.
\]

86
Therefore, the right hand side of (5.6) converges to
\[ \int_{\Gamma} \lambda e^{\lambda t} R(\lambda, A) d\lambda \]
as \( R \to \infty \). Since \( A \) is closed, it follows that \( e^{tA} X \subseteq D(A) \) and
\[ Ae^{tA} = \frac{1}{2\pi i} \int_{\Gamma} \lambda e^{\lambda t} R(\lambda, A) d\lambda, \quad \| Ae^{tA} \| \leq \frac{C}{2t} \]
for all \( t > 0 \). In a similar way one sees that
\[ \left| \int_{\Gamma} \lambda e^{\lambda t} R(\lambda, A) d\lambda - \int_{\Gamma_R} \lambda e^{\lambda t} R(\lambda, A) d\lambda \right| \leq 2K \int_{R}^{\infty} e^{ts\cos \theta} ds \leq \frac{2K}{|\cos \theta|} e^{R\epsilon \cos \theta} \to 0 \]
as \( R \to \infty \), uniformly for \( t \geq \epsilon \). As a result,
\[ \int_{\Gamma} \lambda e^{\lambda t} R(\lambda, A) d\lambda \]
converges in \( L(X) \) uniformly for \( t \geq \epsilon \), and so \( t \mapsto e^{tA} \in L(X) \) is continuously differentiable for \( t > 0 \) with
\[ \frac{d}{dt} e^{tA} = Ae^{tA}. \]
For \( x \in D(A) \), it further holds
\[ Ae^{tA} x = \lim_{R \to \infty} \frac{1}{2\pi i} \int_{\Gamma_R} e^{\lambda t} R(\lambda, A) A x d\lambda = e^{tA} A x. \]
c) Let \( x \in D(A), R > r, \) and \( t > 0 \). As in part a), Cauchy’s formula (5.2) yields
\[ \frac{1}{2\pi i} \int_{\Gamma} e^{\lambda t} \frac{d}{d\lambda} = \lim_{R \to \infty} \frac{1}{2\pi i} \int_{\Gamma_R} \frac{e^{\lambda t}}{\lambda - 0} d\lambda = 1. \]
Observing that \( \lambda R(\lambda, A)x - x = R(\lambda, A)Ax \), we conclude that
\[ e^{tA} x - x = \frac{1}{2\pi i} \int_{\Gamma} e^{\lambda t} (R(\lambda, A) - \frac{1}{\lambda}) x d\lambda = \frac{1}{2\pi i} \int_{\Gamma} \frac{e^{\lambda t}}{\lambda} R(\lambda, A) A x d\lambda. \]
Since the integrand is bounded by \( \frac{1}{|\lambda|} \) on \( \Gamma \) for all \( t \in (0, 1] \), Lebesgue’s convergence theorem implies that there exists the limit
\[ \lim_{t \to 0} (e^{tA} x - x) = \frac{1}{2\pi i} \int_{\Gamma} \frac{1}{\lambda} R(\lambda, A) A x d\lambda =: z. \]
Let \( K_R = \{ Re^{i\alpha} : -\theta \leq \alpha \leq \theta \} \). Cauchy’s theorem (5.1) shows that
\[ \int_{\Gamma_R \cup (-K_R)} \frac{1}{\lambda} R(\lambda, A) A x d\lambda = 0. \]
Since also
\[ \| \int_{-K_R}^{\Gamma_R} \frac{1}{\lambda} R(\lambda, A) A x d\lambda \| \leq \frac{2\pi RK}{R^2} \| A x \| \rightarrow 0 \]
as \( R \to \infty \), we arrive at \( z = 0 \). Because of the uniform boundedness of \( e^{tA} \), it follows that \( e^{tA} x \to x \) as \( t \to 0 \) for all \( x \in D(A) \).

Conversely, if \( e^{tA} x \to y \) as \( t \to 0 \), then \( y \in D(A) \) by part b). Moreover, \( R(1, A)e^{tA} x = e^{tA} R(1, A) x \) tends to \( R(1, A) x \) as \( t \to 0 \), since \( R(1, A) x \in D(A) \). We thus obtain \( R(1, A) y = R(1, A) x \), and so \( x \) belongs to \( D(A) \). \qed
Remark 5.13. If \( A - \omega I = A_\omega \) is sectorial of angle \( > \frac{\pi}{2} \) for some \( \omega \in \mathbb{R} \), then

\[
e^{-\omega t} e^{tA_\omega} = \frac{1}{2\pi i} \int_{\Gamma} e^{t(\lambda + \omega)} R(\lambda + \omega, A) d\lambda = \frac{1}{2\pi i} \int_{\omega + \Gamma} e^{\mu t} R(\mu, A) d\lambda =: e^{tA}
\]

holds for all \( t > 0 \). It is easy to see that \( e^{tA} \) has the analogous properties as in the case \( \omega = 0 \).

Corollary 5.14. Let \( A \) be sectorial with constant \((\phi, K, \omega)\), where \( \phi > \frac{\pi}{2} \) and let \( x \in D(A) \). Then \( u(t) = e^{tA}x, \ t \geq 0, \) is the unique solution in \( C^1((0,\infty),X) \cap C((0,\infty),[D(A)] \cap C(\mathbb{R}_+,X) \) of the initial value problem

\[
(5.7) \quad u'(t) = Au(t), \quad t > 0, \quad u(0) = x.
\]

Proof. Existence follows from Theorem 5.12 and Remark 5.13. Let \( v \) be another solution of (5.7). Let \( 0 < \varepsilon \leq s \leq t - \varepsilon \leq t \). Theorem 5.12 then implies that

\[
\frac{d}{ds} e^{(t-s)A}v(s) = -e^{(t-s)A}Av(s) + e^{(t-s)A}v'(s) = 0.
\]

As in Example 4.33, this fact yields \( e^{(t-\varepsilon)A}v(\varepsilon) = e^{\varepsilon A}v(t - \varepsilon) \). Letting \( \varepsilon \to 0 \), one obtains that \( e^{tA}x = v(t) \).

Example 5.15. Consider \( X = C([0,1]), \ A\varphi = \varphi'' \), \( D(A) = \{ \varphi \in C^2([0,1]), \varphi'(0) = \varphi'(1) = 0 \} \). Let \( u_0 \in X \). Then \( u(t) = e^{tA}u_0 \) belongs to

\[
C(\mathbb{R}_+,X) \cap C((0,\infty),C^2([0,1])) \cap C^1((0,\infty),X))
\]

and solves the following partial differential equation

\[
\begin{cases}
\partial_t u(t, x) = \partial_{xx} u(t, x), & t > 0, \ x \in [0,1], \\
\partial_x u(t, 0) = \partial_x u(t, 1), & t > 0, \\
u(0, x) = u_0(x), & x \in [0,1].
\end{cases}
\]
Bibliography


BIBLIOGRAPHY


