Functional Analysis

lecture by

PROF. DR. FELIX FINSTER

during the winter semester 2012/13

revision and layout in LyX by

ANDREAS VÖLKLEIN

Last changed: January 10, 2013

ATTENTION

This script does *not* replace the lecture.
Therefore it is recommended *strongly* to attend the lecture.

Links

The text of the "GNU Free Documentation License" can also be read on the following site:

`https://www.gnu.org/licenses/fdl-1.3.en.html`

A transparent copy of the recent version of this document can be downloaded from:

`https://github.com/andiv/Functional-Analysis`

Literature

- Peter D. Lax: *Functional analysis*; Wiley-Interscience, 2002; ISBN: 0-471-55604-1 (good reference)
- Micheal Reed, Barry Simon: *Methods of Modern Mathematical Physics I - Functional Analysis;* Academic Press, 2010; ISBN: 978-0-12-585050-6
- Friedrich Hirzebruch, Winfried Scharlau: *Einführung in die Funktionalanalysis*; Spektrum Verlag, 1996; ISBN: 3-86025-429-4
(paperback, small)
- Dirk Werner: *Funktionalanalysis*; Springer, 2011; ISBN: 978-3-642-21016-7
- Joachim Weidmann: *Lineare Operatoren in Hilberträumen, Teil I: Grundlagen*; Teubner, 2000; ISBN: 3-519-02236-2
- Walter Rudin: *Functional Analysis*; McGraw-Hill, 1991; ISBN: 7-111-13415-X

Zorn's Lemma:

- Hans-Joachim Kowalsky, Gerhard O. Michler: *Lineare Algebra*; de Gruyter, 2003; ISBN: 3-11-017963-6

Measure theory for the Riesz representation theorem:

- Walter Rudin: *Real and complex analysis*; McGraw-Hill, 2009; ISBN: 0-07-054234-1

Contents

0 Basic Notions **2**
0.1 Definition (metric, ε-ball, Cauchy sequence, complete, Polish space) 2
0.2 Definition (norm, Banach space) 2
0.3 Definition (continuous, bounded) 3
0.4 Lemma (continuous $\Leftrightarrow$ bounded) 3
0.5 Definition (dual space, sup-norm) 3
0.6 Theorem 3

1 The Hahn-Banach Theorem and Applications **4**
1.1 Definition (partial ordering, chain, upper bound, maximal) 4
1.2 Zorn's lemma 4
1.3 Definition (sublinear) 5
1.4 Theorem (Hahn-Banach, real version, 1927/29) 5
1.5 Theorem (Hahn-Banach, complex version) 6
1.6 Theorem 7
1.7 Corollary 7
1.8 Definition (interior point) 8
1.9 Theorem (geometric Hahn-Banach) 9
1.10 Lemma 9
1.11 Lemma 10

2 Normed Spaces **12**
2.0.1 Definition (equivalent norms) 12
2.0.2 Theorem 12
2.0.3 Theorem 12
2.0.4 Constructions (Quotient space, Cartesian product) 12
2.0.5 Definition (separable) 13
2.0.6 Examples 13
2.0.7 Example 14
2.0.8 Example 14
2.1 Non-Compactness of the Unit Ball 14
2.1.1 Theorem 14
2.1.2 Lemma 14
2.2 Spaces of linear Mappings, Dual Spaces 15
2.2.1 Lemma 16
2.2.2 Theorem and Definition (dual pairing) 16
2.2.3 Theorem 16
2.2.4 Definition (reflexive) 17
2.2.5 Example 17
2.3 Weak Convergence (Schwache Konvergenz) 18

2.3.1 Definition (weak convergence, weak Cauchy sequence) 18
2.3.2 Theorem (Uniqueness of weak limit) 18
2.3.3 Theorem (convergence implies weak convergence) 18
2.3.4 Example 19
2.4 The Baire Category Theorem 19
2.4.1 Definition (nowhere dense, set of first/second category) 20
2.4.2 Theorem (René Baire, 1899) 20
2.4.3 Theorem (Uniform boundedness principle, Prinzip der gleichmäßigen Beschränktheit) 21
2.4.4 Corollary 22
2.4.5 Corollary and Definition (Banach-Steinhaus, equicontinuous, uniformly continuous) 23
2.4.6 Definition (open) 24
2.4.7 Theorem (Open mapping theorem, Prinzip der offenen Abbildung) . 24
2.4.8 Corollary 24
2.4.9 Theorem (Closed graph theorem, Satz vom abgeschlossenen Graphen) 27
2.5 Neumann series 28
2.5.1 Lemma and Definition (Neumann series) 28
2.5.2 Theorem 29
2.5.3 Theorem 29

3 Hilbert spaces **30**
3.0.1 Definition (Hilbert space) 30
3.0.2 Lemma (parallelogram equality) 30
3.0.3 Definition (orthogonal, orthonormal) 31
3.0.4 Theorem (Bessel's inequality) 31
3.0.5 Example 32
3.1 Projection on closed convex subsets 32
3.1.1 Theorem (Hilbert) 33
3.1.2 Corollary 34
3.1.3 Theorem (Fréchet-Riesz) 36
3.1.4 Theorem (Lax-Milgram) 37
3.1.5 Corollary 39
3.2 Orthonormal Bases in Separable Hilbert Spaces 40
3.2.1 Example 40
3.2.2 Definition (orthonormal system, Hilbert space basis, cardinality) . . 40
3.2.3 Theorem 42
3.2.4 Theorem (Existence of Hilbert space basis) 43
3.2.5 Theorem 44
3.2.6 Theorem 45
3.3 Weak Compactness of the Closed Unit Ball 46
3.3.1 Definition (weak (sequential) compactness) 46
3.3.2 Proposition 46
3.3.3 Theorem (Weak Compactness of the Closed Unit Ball) 47

4 Operators on Hilbert spaces **50**
4.0.1 Example 50
4.0.2 Definition (linear operator, domain, bounded) 51
4.0.3 Lemma 51

4.1 Isometric and unitary operators . . . 51
4.1.1 Definition (isometric operator) . . . 51
4.1.2 Proposition . . . 52
4.1.3 Definition (unitary operator) . . . 52
4.2 The Closure of an Operator . . . 53
4.2.1 Definition (closable operator) . . . 53
4.2.2 Definition (closed) . . . 53
4.2.3 Theorem (closed graph theorem) . . . 53
4.2.4 Example . . . 53
4.2.5 Theorem (Criterion for closable) . . . 54
4.3 The adjoint of a densely defined operator . . . 54
4.3.1 Theorem . . . 55
4.3.2 Theorem . . . 55
4.4 Symmetric and self-adjoint densely defined operators . . . 55
4.4.1 Definition (symmetric, (essentially) self-adjoint) . . . 55
4.4.2 Example . . . 56
4.4.3 Lemma . . . 56
4.5 Heisenberg's uncertainty principle . . . 57
4.5.1 Theorem (Winter-Wieland) . . . 57
4.5.2 Theorem (Heisenberg's uncertainty principle) . . . 57
4.6 Spectrum and resolvent . . . 58
4.6.1 Definition (continuously invertible, resolvent, spectrum) . . . 58
4.6.2 Lemma . . . 58
4.6.3 Theorem (resolvent equation) . . . 59

5 Compact Operators **60**
5.1 Definition (compact operator) . . . 60
5.2 Example (integral operator) . . . 60
5.3 Theorem . . . 61
5.4 Lemma . . . 61
5.5 Lemma (Fredholm operator) . . . 62
5.6 Theorem (Fredholm Alternative) . . . 63
5.7 Theorem (Riesz-Schauder) . . . 64
5.8 Theorem . . . 65
5.9 Lemma . . . 66
5.10 Theorem (Hilbert-Schmidt) . . . 67
5.11 Definition (spectral radius) . . . 68
5.12 Theorem . . . 68
5.13 Ritz method . . . 72

6 A few (technical) results **76**
6.1 Dini's theorem . . . 76
6.1.1 Definition (point-wise/uniform convergence) . . . 76
6.1.2 Theorem . . . 76
6.1.3 Definition (monotonically increasing/decreasing) . . . 77
6.1.4 Theorem (Dini) . . . 77
6.2 Stone-Weierstraß theorem . . . 78
6.2.1 Definition (polynomials) . . . 78
6.2.2 Lemma . . . 79

6.2.3 Lemma . . . 79
6.2.4 Definition . . . 80
6.2.5 Theorem (Bernstein) . . . 80
6.2.6 Theorem (Weierstraß) . . . 82
6.2.7 Theorem (Stone-Weierstraß) . . . 83
6.2.8 Theorem (Stone-Weierstraß, complex version) . . . 86
6.3 Arzelà-Ascoli theorem . . . 86
6.3.1 Definition (relatively compact) . . . 86
6.3.2 Definition (equicontinuous) . . . 87
6.3.3 Theorem (Arzelà-Ascoli) . . . 87
6.4 The Riesz representation theorem . . . 89
6.4.1 Examples . . . 89
6.4.2 Definition (bounded, positive, regular measure) . . . 90
6.4.3 Theorem (Riesz representation theorem) . . . 90
6.4.4 Example . . . 91
6.4.5 Definition (total variation) . . . 92
6.4.6 Example . . . 94

7 The Spectral Theorem for symmetric bounded operators 98
7.1 The Spectrum of symmetric bounded operators . . . 98
7.1.1 Theorem . . . 99
7.1.2 Theorem . . . 100
7.2 The continuous functional calculus . . . 101
7.2.1 Theorem (continuous functions of operators) . . . 101
7.2.2 Lemma (spectral mapping theorem for polynomials) . . . 102
7.2.3 Definition (normal operator) . . . 103
7.2.4 Theorem . . . 103
7.2.5 Lemma . . . 104
7.3 Spectral Measures . . . 106
7.3.1 Lemma . . . 107
7.3.2 Lemma . . . 108
7.3.3 Theorem . . . 108
7.3.4 Theorem (Spectral theorem in functional calculus form) . . . 109
7.3.5 Remark . . . 112
7.3.6 Definition (projection operator, spectral measure) . . . 112
7.3.7 Theorem . . . 113

Appendix 115
Acknowledgements . . . 115
GNU Free Documentation License . . . 116

Motivation

In linear algebra one mainly considers finite-dimensional vector spaces with additional structures like norm $\|.\|$ or scalar product $\langle .,. \rangle$.

Let $(V, \langle .,. \rangle)$ be a finite-dimensional scalar product space and $A : V \to V$ a linear map, which is self-adjoint, that means for all $u,v \in V$:

$$\langle Au,v \rangle = \langle u,Av \rangle$$

Theorem (orthonormal eigenvector basis)

There exists an orthonormal eigenvector basis $(u_i)_{i \in \{1,\dots,n\}}$, that means with the eigenvalues $\lambda_i \in \mathbb{R}$:

$$\langle u_i,u_j \rangle = \delta_{ij} \qquad\qquad Au_i = \lambda_i u_i$$

In infinite dimensions the generalization is the *spectral theorem.*

First reformulate the result from linear algebra:
Let E_{λ_i} be the orthogonal projection operator on the eigenspace corresponding to λ_i. If this eigenspace is one dimensional, this means:

$$E_{\lambda_i} v = u_i \langle u_i,v \rangle = |u_i\rangle \langle u_i|v\rangle$$

Then one can write A as:

$$A = \sum_{i=1}^{n} \lambda_i E_{\lambda_i}$$

Theorem (spectral theorem)

Let $A \in L(H)$ be a self-adjoint (selbstadjungiert) operator, then it holds:

$$A = \int_{\sigma(A)} \lambda \mathrm{d}E_\lambda$$

$\sigma(A) \subseteq \mathbb{R}$ is the spectrum of A and E_λ the projection-valued measure (Spektralmaß).

Applications typically are differential operators, for example:

$$\Delta_{\mathbb{R}^3} = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}$$

$$\Delta_{\mathbb{R}^3} : C_0^\infty\left(\mathbb{R}^3\right) \to C^\infty\left(\mathbb{R}^3\right) \quad \text{linear operator}$$

Applications in more detail are studied in the lectures on partial differential equations I + II.

0 Basic Notions

Let E be a vector space (Vektorraum), for example the finite-dimensional vector space $E \simeq \mathbb{R}^3$. In the following list the later spaces are special cases of the previous ones:

- topological vector spaces
- metric spaces with a metric $d\left(.,.\right)$ (Polish spaces if complete)
- normed spaces with norm $\|.\|$ (Banach spaces if complete)
- scalar product spaces $\langle .,. \rangle$ (Hilbert spaces if complete)

Let $\mathbb{K}$ be either $\mathbb{R}$ or $\mathbb{C}$.

0.1 Definition (metric, ε-ball, Cauchy sequence, complete, Polish space)

A map $d : E \times E \to \mathbb{R}$ is called *metric*, if for all $x,y,z \in E$ holds:

i) $d\left(x,y\right) = d\left(y,x\right)$ (symmetry)

ii) $d\left(x,y\right) \geq 0$ and $d\left(x,y\right) = 0 \Leftrightarrow x = y$ (positive definiteness)

iii) $d\left(x,y\right) \leq d\left(x,z\right) + d\left(z,y\right)$ (triangle inequality)

$B_\varepsilon\left(x\right) := \left\{ z \in E \middle| d\left(x,z\right) < \varepsilon \right\}$ is called *ε-ball*.
Consider the topology generated by $B_\varepsilon\left(x\right)$: A set $\Omega \subseteq E$ is open if and only if:

$$\underset{x\in\Omega}{\forall} \; \underset{\varepsilon\in\mathbb{R}_{>0}}{\exists} : B_\varepsilon\left(x\right) \subseteq \Omega$$

Completeness:
$\left(x_n\right)_{n\in\mathbb{N}}$ is a *Cauchy sequence* if and only if:

$$\underset{\varepsilon\in\mathbb{R}_{>0}}{\forall} \; \underset{N\in\mathbb{N}}{\exists} \; \underset{n,m\in\mathbb{N}_{>N}}{\forall} : d\left(x_n,x_m\right) < \varepsilon$$

E is *complete* if and only if every Cauchy sequence has a limit.
A complete metric space is also called a *Polish space*.

0.2 Definition (norm, Banach space)

Let $(E, \|.\|)$ be a *normed space*, i.e. a $\mathbb{K}$-vector space with a map $\|.\| : E \to \mathbb{R}_{\geq 0}$ called *norm* with the following properties for $x,y \in E$ and $\lambda \in \mathbb{K}$:

i) $\|x\| \geq 0$ and $\|x\| = 0 \Leftrightarrow x = 0$ (positive definiteness)

ii) $\|\lambda x\| = |\lambda| \cdot \|x\|$ (homogeneity)

iii) $\|u+v\| \leq \|u\| + \|v\|$ (triangle inequality)

Define the metric $d(x,y) := \|x-y\|$. A complete normed spaces is called *Banach space*.

Let $A : E \to F$ be a linear map between the Banach spaces $(E, \|.\|_E)$ and $(F, \|.\|_F)$.

0.3 Definition (continuous, bounded)

A is *continuous* (stetig) if $A^{-1}(\Omega) \subseteq E$ is open for all open $\Omega \subseteq F$.
A is *bounded* (beschränkt) if there exists a $C \in \mathbb{R}_{>0}$ such that for all $u \in E$ holds:

$$\|Au\|_F \leq C\,\|u\|_E$$

0.4 Lemma (continuous ⇔ bounded)

A is continuous ⇔ A is bounded.

(no proof)

0.5 Definition (dual space, sup-norm)

The *dual space* of E is the space of continuous linear mappings from E to $\mathbb{K}$:

$$E^* = L(E,\mathbb{K})$$

$L(E,F)$ is a vector space: For $A,B \in L(E,F)$, $\lambda,\mu \in \mathbb{K}$ and $u \in E$ define:

$$(\lambda A + \mu B)(u) := \lambda A(u) + \mu B(u)$$

Define also a norm on $L(E,F)$, which is called *sup-norm*:

$$\|A\| := \sup_{u \in E, \|u\|_E \leq 1} \|Au\|_F$$

0.6 Theorem

If F is complete, so is $L(E,F)$.
In particular E^* is a Banach space for every E.

(no proof)

1 The Hahn-Banach Theorem and Applications

As a preparation we need Zorn's lemma.

1.1 Definition (partial ordering, chain, upper bound, maximal)

Let A be a set and $\leq$ a *partial ordering* (Halbordnung), i.e. for all a,b,$c \in A$:

i) $a \leq b$ and $b \leq c \Rightarrow a \leq c$ (transitivity)

ii) $a \leq a$ (reflexivity)

iii) $a \leq b \wedge b \leq a \Rightarrow a = b$ (antisymmetry)

Note: We do *not* demand that for all a,$b \in A$ holds:

$$(a \leq b) \vee (b \leq a)$$

This is a property of a ordering relation.
$(A, \leq)$ is called *partially ordered set* (teilweise geordnete Menge).
A subset $K \subseteq A$ is called *chain* (Kette, total geordnete Teilmenge) if for all x,$y \in K$ holds:

$$(x \leq y) \vee (y \leq x)$$

An element $u \in A$ is called *upper bound* (obere Schranke) of $B \subseteq A$ if $x \leq u$ for all $x \in B$.
An element $m \in A$ is called *maximal* if $m \leq a \in A \Rightarrow m = a$.

1.2 Zorn's lemma

Let $(A, \leq)$ be a partially ordered set in which every chain has an upper bound. Then there is a maximal element.

Proof

This follows from the axiom of choice, see e.g. Kowalsky: Linear algebra.

1.3 Definition (sublinear)

Let X be a *real* vector space (without topology) and $l : X \to \mathbb{R}$ linear.
$p : X \to \mathbb{R}$ is called *sublinear* if for all $x,y \in X$ and $a \in \mathbb{R}_{>0}$:

i) $p(ax) = ap(x)$

ii) $p(x+y) \leq p(x) + p(y)$

A typical example is $p(x) = \|x\|$, but p does not need to be positive. Another example is any linear mapping.

1.4 Theorem (Hahn-Banach, real version, 1927/29)

Let X be a real vector space and $Y \subseteq X$ a subspace (Untervektorraum), $p : X \to \mathbb{R}$ sublinear and $l : Y \to \mathbb{R}$ linear with $l(y) \leq p(y)$ for all $y \in Y$.
Then there is a linear extension (Fortsetzung) $\tilde{l} : X \to \mathbb{R}$ of l to X, i.e. $\tilde{l}\big|_Y = l$, such that for all $x \in X$ holds:

$$\tilde{l}(x) \leq p(x)$$

Proof

i) Assume $Y \subsetneq X$, since otherwise there is nothing to prove. Choose a vector $z \in X \setminus Y$. We want to extend l to the span of Y and $\langle z \rangle$. $\tilde{l}(z)$ needs to be prescribed. For all $y \in Y$ and $a \in \mathbb{R}$ holds:

$$\tilde{l}(y + az) \overset{\text{linearity}}{=} l(y) + a\tilde{l}(z) \overset{\text{demand}}{\leq} p(y + az)$$

If $a = 0$, the inequality is clear. By homogeneity assumptions, it is sufficient to consider the case $a = \pm 1$. We thus demand for all $y,y' \in Y$:

$$\begin{aligned} l(y) + \tilde{l}(z) &\leq p(y+z) \\ l(y') - \tilde{l}(z) &\leq p(y'-z) \end{aligned}$$

This is equivalent to:

$$l(y') - p(y'-z) \leq \tilde{l}(z) \leq p(y+z) - l(y)$$

We can choose $\tilde{l}(z)$ if and only if:

$$l(y') - p(y'-z) \leq p(y+z) - l(y)$$

(For example set $\tilde{l}(z) = \sup_{y' \in Y} l(y') - p(y'-z)$.)

$$\Leftrightarrow \quad l(y') + l(y) \overset{\text{lineariy}}{=} l(y'+y) \leq p(y+z) + p(y'-z)$$

Now prove this inequality:
From $y' + y \in Y$ follows that $l(y+y') \leq p(y+y')$ by hypothesis. Moreover, as p is sublinear, it follows:

$$p(y + z - z + y') \leq p(y'+z) + p(y'-z)$$

So the inequality is shown. Thus l can be extended to $Y + \langle z \rangle$.

ii) Consider all extensions:

$$A := \{(Z,l)\,|Y \subseteq Z \subseteq X \text{ subspace, } l : Z \to \mathbb{R} \text{ extension of } l_Y : Y \to \mathbb{R}\}$$

This set has a partial ordering $\leq$ defined by $(Z,l) \leq (Z',l')$ if $Z \subseteq Z'$ and $l'\big|_Z = l$.
For an index set I (possibly infinite, uncountable) let $K = \{(Z_\nu,l_\nu)\,|\nu \in I\}$ be a chain, i.e. for all $(Z,l), (Z',l') \in K$:

$$\left((Z,l) \leq \left(Z',l'\right)\right) \vee \left(\left(Z',l'\right) \leq (Z,l)\right)$$

Set $Z = \bigcup_{\nu \in I} Z_\nu$ and define $l : Z \to \mathbb{R}$ by $l\big|_{Z_\nu} = l_\nu$. (Thus suppose $u \in Z$, so there is a $\nu \in I$ with $u \in Z_\nu$. Set $l(u) := l_\nu(u)$. ν need not be unique. Suppose $u \in Z_{\nu'}$, then we know that either $Z_{\nu'} \subseteq Z_\nu$ and $l_\nu\big|_{Z_{\nu'}} = l_{\nu'}$ or $Z_\nu \subseteq Z_{\nu'}$ and $l_{\nu'}\big|_{Z_\nu} = l_\nu$. In both cases we have $l_\nu(u) = l_{\nu'}(u)$, thus $l(u)$ is well defined.)
This (Z,l) is an upper bound, because for all $\nu \in I$ we have $Z_\nu \subseteq Z = \bigcup_{\lambda \in I} Z_\lambda$ and l is an extension of l_ν.
With Zorn's Lemma follows, that there exists an maximal element $\left(\tilde{Y},\tilde{l}\right)$.

Claim: $\tilde{Y} = X$

Proof: Otherwise there would be a vector $u \in X \setminus \tilde{Y}$, and $\tilde{l}$ could be extended to $\tilde{Y} \oplus \langle u \rangle$, as shown in i), in contradiction to the maximality of $\tilde{l}$. Thus $\left(X = \tilde{Y},\tilde{l}\right)$ is the desired extension. $\square_{\text{Claim}}$

$\square_{1.4}$

1.5 Theorem (Hahn-Banach, complex version)

Let X be a complex vector space and $Y \subseteq X$ a subspace. Before, we had $l(x) \leq p(x)$ as condition, which does not make sense in the complex case, since:

$$l\left(e^{\mathbf{i}\varphi}x\right) = e^{\mathbf{i}\varphi}l(x) \overset{\text{in general}}{\notin} \mathbb{R}$$

Let $p : X \to \mathbb{R}$ be a *seminorm*, i.e.:

i) $p(ax) = |a|\,p(x)$ (homogeneity)

ii) $p(x+y) \leq p(x) + p(y)$ (triangle inequality)

Let $l : Y \to \mathbb{C}$ be a linear functional with $|l(y)| \leq p(y)$ for all $y \in Y$.
Then l can be extended to X such that $|l(x)| \leq p(x)$ holds for all $x \in X$.

Proof

We also consider X as a real vector space. (u and $\mathbf{i}u$ are then linearly independent vectors.) Decompose l into its real and imaginary parts.

$$l(y) = l_1(y) + \mathbf{i}l_2(y)$$
$$l_1 := \operatorname{Re}(l(y))$$
$$l_2 := \operatorname{Im}(l(y))$$

l_1 and l_2 are real-linear and:

$$l_1(\mathbf{i}y) = \operatorname{Re}(l(\mathbf{i}y)) = \operatorname{Re}(\mathbf{i}l(y)) = -\operatorname{Im}(l(y)) = -l_2(y)$$

Conversely, suppose that l_1 is real-linear. Then

$$l(x) := l_1(x) - \mathbf{i} \cdot l_1(\mathbf{i}x)$$

this is indeed a complex-linear function. We know that $|l(y)| \leq p(y)$ holds for all $y \in Y$.

$$\begin{aligned} l_1(y) = \operatorname{Re}(l(y)) &\leq |l(y)| \\ \Rightarrow \qquad l_1(y) &\leq p(y) \end{aligned}$$

Theorem 1.4 yields an real-linear extension $\tilde{l}_1 : X \to \mathbb{R}$ such that $\tilde{l}_1(x) \leq p(x)$ for all $x \in X$. Set $\tilde{l}(x) = \tilde{l}_1(x) - \mathbf{i}\,\tilde{l}_1(\mathbf{i}x)$, so that $\tilde{l} : X \to \mathbb{C}$ is complex-linear.

Claim: $\left|\tilde{l}(x)\right| \leq p(x)\ \forall_{x \in X}$

Proof: Polar decomposition:

$$\begin{aligned} \tilde{l}(x) &= re^{\mathbf{i}\varphi} \\ \left|\tilde{l}(x)\right| &= r = e^{-\mathbf{i}\varphi}\tilde{l}(x) \underset{\text{complex-linear}}{\overset{\tilde{l}\text{ is}}{=}} \tilde{l}\left(e^{-\mathbf{i}\varphi}x\right) = \operatorname{Re}\left(\tilde{l}\left(e^{-\mathbf{i}\varphi}x\right)\right) = \\ &= \tilde{l}_1\left(e^{-\mathbf{i}\varphi}x\right) \leq p\left(e^{-\mathbf{i}\varphi}x\right) \overset{\text{homogeneity}}{=} p(x) \end{aligned}$$

$\square_{\text{Claim}}$

$\square_{1.5}$

Now to applications:

1.6 Theorem

Let $(X, \|.\|)$ be a normed $\mathbb{K}$-space (real or complex), $Y \subseteq X$ a subspace. Let φ be a continuous linear functional from Y to $\mathbb{K}$, i.e. for all $y \in Y$ holds:

$$|\varphi(y)| \leq \|\varphi\| \cdot \|y\|$$

Then φ can be continued to all of X with the same supnorm, i. e.:

$$\|\tilde{\varphi}\| := \sup_{x \in X, \|x\| \leq 1} |\varphi(x)| = \|\varphi\| := \sup_{y \in Y, \|y\| \leq 1} |\varphi(y)|$$

Proof

Apply the Hahn-Banach theorem with $\tilde{\varphi} := \|\varphi\| \cdot \|x\|$. $\square_{1.6}$

1.7 Corollary

Let X be a normed space and $u_0 \in X$ with $\|u_0\| = 1$. Then there exists a linear functional $\varphi : X \to \mathbb{K}$ such that:

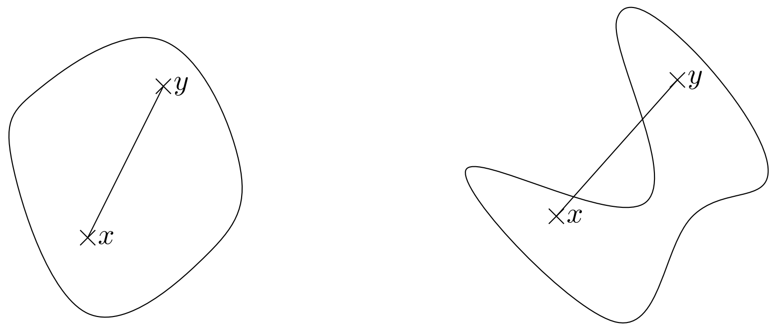
$$\varphi(u_0) = 1 \qquad\qquad \|\varphi\| = 1$$

Proof

Let $Y := \langle u_0 \rangle$ and define $\varphi_0 : \langle u_0 \rangle \to \mathbb{K}$ by $\varphi_0(u_0) = 1$. Extend φ_0 by the Hahn-Banach theorem 1.6. $\square_{1.7}$

The Hahn-Banach theorem also has a geometric formulation. Consider only the real case:
A set $K \subseteq X$ is called *convex* if for all $x,y \in K$ and $\tau \in [0,1]$:

$$\tau x + (1-\tau) y \in K$$



(a) convex set (b) concave set

Figure 1.1: convexity

Geometric question:

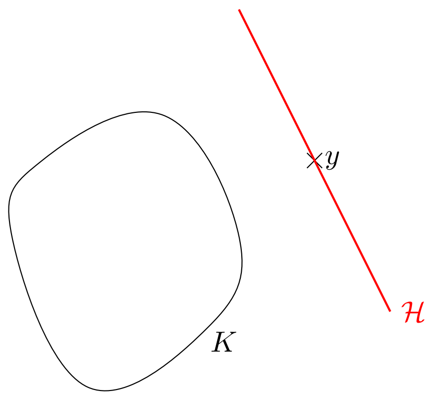


Figure 1.2: not intersecting hyperplane

Is there a hyperplane $\mathcal{H}$, which meets $y \notin K$, but does not intersect K?

1.8 Definition (interior point)

$x_0 \in K$ is an *interior point* (innerer Punkt) *of* K *with respect to* $u \in X$ if there exists an $\varepsilon \in \mathbb{R}_{>0}$ such that $x_0 + tu \in K$ for all $t \in (-\varepsilon,\varepsilon)$.
$x_0 \in K$ is an *interior point* if for all $u \in X$ there is a $\varepsilon = \varepsilon(u) \in \mathbb{R}_{>0}$ such that $x_0 + tu \in X$ for all $t \in (-\varepsilon,\varepsilon)$.

1.9 Theorem (geometric Hahn-Banach)

Let $K \neq \emptyset$ be convex and all points of K be interior points. Let $y \notin K$. Then there is a linear functional $l : X \to \mathbb{R}$ such that $l(x) < 1$ for all $x \in K$ and $l(y) = 1$.
$\mathcal{H} := \left\{x \in X \middle| l(x) = 1\right\}$ defines a hyperplane. Now $y \in \mathcal{H}$ and $l\big|_K < 1$ mean that K lies in one half-space.

First introduce a suitable sublinear functional. Without loss of generality, assume $0 \in K$ (otherwise shift K).

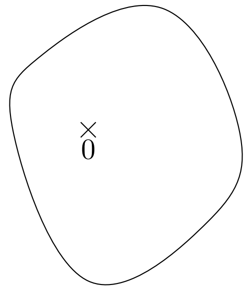


Figure 1.3: $0 \in K$

The functional $p : K \to \mathbb{R}_{\geq 0}$ with

$$p(x) := \inf\left\{a \in \mathbb{R}_{>0} \middle| \frac{x}{a} \in K\right\}$$

is called gauge (Eichung).
Since x is an interior point, we know that $\frac{x}{a} \in K$ if $a > 1 - \varepsilon(x)$.
p is even defined on all of X, because for $x \in X$, now $\tau x \in K$ if $|\tau|$ is sufficiently small, because $0 \in K$ is an interior point.

$$p(x) < 1 \quad \Leftrightarrow \quad x \in K$$

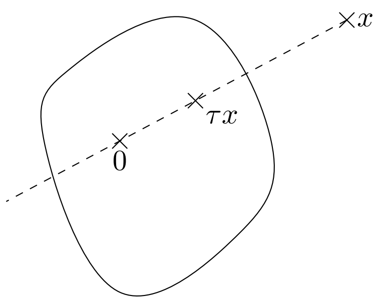


Figure 1.4: $x \notin K$, $\tau x \in K$

1.10 Lemma

p is sublinear.

Proof

The homogeneity is clear from the definition.

sub-additivity (triangle equation):
Take $x{,}y \in K$ and choose $a{,}b \in \mathbb{R}_{>0}$ such that $\frac{x}{a}, \frac{y}{b} \in K$. The convexity of K implies for all $\tau \in [0{,}1]$:

$$\tau \frac{x}{a} + (1-\tau)\frac{y}{b} \in K$$

Choose $\tau = \dfrac{a}{a+b}$, then holds $1-\tau = \dfrac{b}{a+b}$, which gives:

$$\Rightarrow \quad \frac{1}{a+b}(x+y) \in K$$

$$p(x+y) \leq a+b$$

Taking the infimum over a and b gives $p(x+y) \leq p(x) + p(y)$:

$$p(x+y) = \inf \underbrace{\left\{ c \in \mathbb{R}_{>0} \middle| \frac{x+y}{c} \in K \right\}}_{\ni a+b} \leq a+b$$

$$\begin{aligned} p(x) &= \inf \left\{ a \middle| \frac{x}{a} \in K \right\} \quad &\Rightarrow \quad \underset{\varepsilon>0}{\forall} \underset{a \in \mathbb{R}_{>0}}{\exists} : p(x) \geq a - \varepsilon \\ p(y) &= \inf \left\{ b \middle| \frac{x}{b} \in K \right\} \quad &\Rightarrow \quad \underset{\varepsilon>0}{\forall} \underset{b \in \mathbb{R}_{>0}}{\exists} : p(y) \geq b - \varepsilon \end{aligned}$$

$\square_{1.10}$

1.11 Lemma

$p(x) < 1 \Leftrightarrow x \in K$

Proof

If $x \notin K$ then $\frac{1}{a}x \notin K$ for all $0 < a < 1$ and so $p(x) \geq 1$.

For all $x \in K$ exists an $\varepsilon = \varepsilon(x) \in \mathbb{R}_{>0}$ with $(1+t)\,x \in K$ for all $t \in (-\varepsilon{,}\varepsilon)$.

$$\begin{aligned} &\Rightarrow \quad \left(1 + \frac{\varepsilon}{2}\right) x \in K \\ &\Rightarrow \quad p(x) \leq \frac{1}{1+\frac{\varepsilon}{2}} < 1 \end{aligned}$$

$\square_{1.11}$

Proof of Theorem 1.9

Introduce l on $\langle y \rangle$ by $l(y) = 1$. (Assume again that $0 \in K$ and so $y \neq 0$.)
Write $z = ay \in \langle y \rangle$ with $a \in \mathbb{R}$.

- If $a < 0$, then $l(z) = a \cdot l(y) = a < 0$ but $p(z) \geq 0$ and thus the inequality $l(z) \leq p(z)$ is trivially satisfied.
- If $a > 0$ it holds:

$$l(z) = a \underset{\Rightarrow p(y) \geq 1}{\overset{y \notin K}{\leq}} a \cdot p(y) \underset{\text{homogeneity}}{\overset{\text{positive}}{=}} p(ay) = p(z)$$

So for all $z \in \langle y \rangle$ holds $l(z) \leq p(z)$.
The Hahn-Banach Theorem yields an extension $l : X \to \mathbb{R}$ such that $l(x) \leq p(x)$ for all $x \in X$.
Therefore for all $x \in K$ we have:

$$l(x) \leq p(x) < 1$$

$\square_{1.9}$

2 Normed Spaces

Let $(E, \|.\|)$ be a normed space and let the open balls $B_\varepsilon(x) = \left\{y \middle| \, \|x-y\| < \varepsilon\right\}$ generate the topology on E.

2.0.1 Definition (equivalent norms)

Two norms $\|.\|_1$ and $\|.\|_2$ are *equivalent*, if there exists a $C \in \mathbb{R}_{>0}$ such that:

$$\frac{1}{C} \|x\|_1 \leq \|x\|_2 \leq C \|x\|_2$$

2.0.2 Theorem

Equivalent norms give rise to the same topology.

(No proof)

2.0.3 Theorem

If E is finite dimensional, then any two norms on E are equivalent.

(No proof)

2.0.4 Constructions (Quotient space, Cartesian product)

Let $F \subseteq E$ be a *closed* subspace. Define the *quotient space* (Faktorraum) ${}^{E}\!/_{F}$ as follows:

$$x \sim y :\Leftrightarrow x - y \in F$$

defines an equivalence relation on E.

$${}^{E}\!/_{F} := {}^{E}\!/_{\sim}$$

is a vector space.

$$\|u\|_{E/F} := \inf_{\substack{\hat{u} \in E \\ \hat{u} - u \in F}} \|\hat{u}\|_E$$

$\left({}^{E}\!/_{F}, \|.\|_{E/F}\right)$ is a normed space. The closedness of F is essential:
Suppose $F \subseteq E$ is not closed. Then there exists an $x \in \overline{F} \setminus F$, thus there is a $(x_n)_{n \in \mathbb{N}}$, $x_n \in F$

with $x_n \to x$.
Let $[x] \in {}^{E}\!/_{F}$ be the equivalence class. Then $[x] \neq 0$, since $x \notin F$, but:

$$\|[x]\| = \inf_{\substack{\hat{x}\in E \\ \hat{x}-x\in F}} \|\hat{x}\| \overset{x-x_n\sim x}{\leq} \inf \|x - x_n\| = 0$$

If $\|.\|_{E/F}$ was a norm, it would imply $[x] = 0$ and thus $x \in F$ in contradiction to $x \in \overline{F} \setminus F$.
Another construction is the *Cartesian product*: Let E and F be normed spaces.

$$E \times F := \left\{ (u,v) \,\middle|\, u \in E, v \in F \right\}$$

$$\|(u,v)\|_{E\times F} := \|u\|_E + \|v\|_F$$

is a norm on $E \times F$.

2.0.5 Definition (separable)

A normed space is called *separable*, if there is a countable dense subset, i.e. there exists a sequence $(x_n)_{n\in\mathbb{N}}$ such that every nonempty open subset of the space contains at least one element of the sequence.

2.0.6 Examples

The space ℓ^∞ of bounded sequences $(a_n)_{n\in\mathbb{N}}$, $a_n \in \mathbb{K}$ with $\left\|(a_n)_{n\in\mathbb{N}}\right\|_\infty := \sup_n |a_n|$ is a Banach space.

$$A := \left\{ (a_n)_{n\in\mathbb{N}} \,\middle|\, a_{2n} = 0 \underset{n\in\mathbb{N}}{\forall} \right\} \subseteq \ell^\infty$$

is a closed subspace.

$$\ell^\infty /_A \cong \left\{ (a_n) \,\middle|\, a_{2n+1} = 0 \underset{n\in\mathbb{N}}{\forall} \right\}$$

$$d := \left\{ (a_n) \,\middle|\, \underset{N\in\mathbb{N}}{\exists} \underset{n\in\mathbb{N}_{>N}}{\forall} a_n = 0 \right\} \subseteq \ell^\infty$$

is a subspace, but not closed in ℓ^∞. Consider for example $\left(a_n = \frac{1}{n}\right) =: x \in \ell^\infty \setminus d$, $x_n \in d$ with $x_n = (a_{n_l})_{l\in\mathbb{N}}$ and:

$$a_{n_l} = \begin{cases} \frac{1}{l} & \text{if } l \leq n \\ 0 & \text{if } l > n \end{cases}$$

Then converges $x_n \to x \notin d$, and therefore d is not closed. The closure is:

$$\overline{d} = \left\{ (a_n) \,\middle|\, a \xrightarrow{n\to\infty} 0 \right\}$$

ℓ^∞ is not separable.

2.0.7 Example

For $1 \leq p < \infty$ define

$$\ell^p = \left\{ (a_n)_{n\in\mathbb{N}} \,\middle|\, \sum_{n=1}^{\infty} |a_n|^p < \infty \right\}$$

and the ℓ^p-norm:

$$\|(a_n)\|_p := \left(\sum_{n=1}^{\infty} |a_n|^p \right)^{\frac{1}{p}}$$

ℓ^p is a normed space (Hölder's inequality, Minkowski inequality) and also separable (see exercises).

2.0.8 Example

Let (Ω,μ) be a measure space (Maßraum).

$$\begin{array}{ll} L^p(\Omega) \;\; (1 \leq p < \infty) & \|f\|_p = \left(\int_\Omega |f(x)|^p \, \mathrm{d}\mu \right)^{\frac{1}{p}} \\ L^\infty(\Omega) & \|f\|_\infty = \operatorname{supess}_\Omega |f(x)| = \sup \left\{ L \in \mathbb{R} \middle| \mu\left(f^{-1}([L,\infty))\right) > 0 \right\} \end{array}$$

2.1 Non-Compactness of the Unit Ball

Let $(E, \|.\|)$ be a normed vector space.

$$K := \overline{B_1(0)} = \left\{ x \in E \middle| \, \|x\| \leq 1 \right\}$$

If $\dim(E) < \infty$, K is compact by the Heine-Borel theorem.

2.1.1 Theorem

If E is infinite-dimensional, then K is not sequentially compact (folgenkompakt), i.e. it is possible to construct a sequence (y_n), $y_n \in K$, which has no convergent subsequence.

2.1.2 Lemma

Let $Y \subsetneq E$ be a proper (echter) closed subspace. Then there is a $z \in E \setminus Y$ with $\|z\| = 1$ such that holds:

$$\begin{aligned} & \underset{y\in Y}{\forall} : \; \|z - y\| > \frac{1}{2} \\ \Leftrightarrow \quad & \overline{B_{\frac{1}{2}}(z)} \cap Y = \emptyset \end{aligned}$$

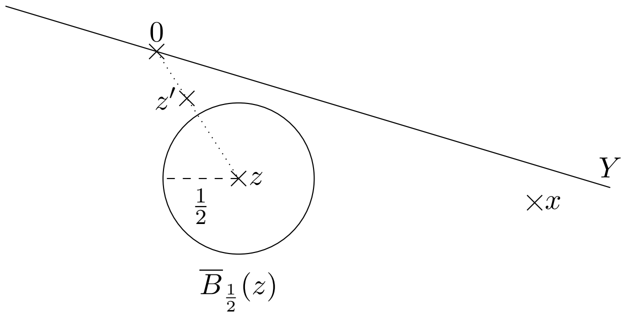


Figure 2.1: $\overline{B_{\frac{1}{2}}(z)} \cap Y = \emptyset$

Proof

Choose $x \in E \setminus Y \neq \emptyset$. As $E \setminus Y$ is open, there is a $\delta \in \mathbb{R}_{>0}$ with $B_\delta(x) \cap Y = \emptyset$. Thus we can define:

$$d := \inf_{y \in Y} \|x - y\| > 0$$

Choose $y_0 \in Y$ such that $\|x - y_0\| < 2d$. Set $z' = x - y_0$. Then $\|z'\| < 2d$ and $\|z' - y\| \geq d$ for all $y \in Y$. Thus $z := \frac{z'}{\|z'\|}$ has the desired properties. $\square_{2.1.2}$

Proof of Theorem 2.1.1

Choose inductively a sequence (y_n): $y_1 \in K$ is arbitrary. $Y_1 := \langle y_1 \rangle$ is a one dimensional subspace, which is closed. Choose $y_2 \in K$ such that $\|y_2 - y\| > \frac{1}{2}$ for all $y \in Y_1$, which is possible according to Lemma 2.1.2.
Suppose $y_1, \ldots, y_n$ are given. $Y_n := \langle y_1, \ldots, y_n \rangle$ is closed. So there exists a $y_{n+1} \in K$ such that for all $y \in Y_n$ holds:

$$\|y_{n+1} - y\| > \frac{1}{2}$$

This sequence has the following properties:

- $y_k \in K$
- For all $k,l \in \mathbb{N}$ with $k < l$ holds $\|y_l - y_k\| > \frac{1}{2}$, since $y_k \in Y_{l-1} = \langle y_1, \ldots, y_{l-1} \rangle$ and we know by construction that $\|y_l - y\| > \frac{1}{2}$ for all $y \in Y_{l-1}$ so especially for $y_k \in Y_{l-1}$.

This implies that (y_k) has no convergent subspace. $\square_{2.1.1}$

2.2 Spaces of linear Mappings, Dual Spaces

Let E,F be normed spaces.
$A : E \to F$ is continuous if and only if it is bounded, i.e. there exists a $C \in \mathbb{R}_{>0}$ such that for all $u \in E$ holds:

$$\|Au\|_F \leq C \|u\|_E$$

Denote by $L(E,F)$ the normed space of all bounded linear maps from E to F and define:

$$\|A\| := \sup_{\|u\| \leq 1} \|Au\| = \sup_{\|u\| = 1} \|Au\|$$

2.2.1 Lemma

If $B \in L(E,F)$ and $A \in L(F,G)$ then Schwarz inequality or Kato inequality holds:

$$\begin{aligned} \|A \cdot B\| &\leq \|A\| \cdot \|B\| \\ \|Au\| &\leq \|A\| \cdot \|u\| \end{aligned}$$

(no proof)

2.2.2 Theorem and Definition (dual pairing)

If F is complete, so is $L(E,F)$.
Special case $F = \mathbb{R}$ and $\|x\|_{\mathbb{R}} = |x|$: $E^* := L(E,\mathbb{R})$ is the dual space.

For $\varphi \in E^*$ and $u \in E$

$$\varphi(u) = (\varphi,u)$$

is called *dual pairing* (duale Paarung).

$$(.,.) : E^* \times E \to \mathbb{R}$$

is a continuous bilinear map. For $u \in E$

$$(.,u) : E^* \to \mathbb{R}$$

defines an element of $E^{**} = L(E^*,\mathbb{R})$. This gives rise to a linear mapping:

$$\iota : E \to E^{**}$$

(no proof)

2.2.3 Theorem

$\iota : E \hookrightarrow E^{**}$ is an isometric embedding of E into E^{**}.

Proof

For $u \in E$ holds:

$$\|\iota(u)\| := \sup_{\varphi \in E^*, \|\varphi\|=1} \|(\iota(u))(\varphi)\| = \sup_{\varphi \in E^*, \|\varphi\|=1} \|\varphi(u)\| \overset{?}{=} \|u\|$$

$$\|\varphi\| = \sup_{v \in E, \|v\|=1} |\varphi(v)|$$

$$\begin{aligned} & \|\varphi(u)\| \leq \|\varphi\| \cdot \|u\| \overset{\|\varphi\|=1}{=} \|u\| \\ \Rightarrow \quad \sup_{\varphi \in E^*, \|\varphi\|=1} & \|\varphi(u)\| \leq \|u\| \end{aligned}$$

To prove $\|\iota(u)\| \geq \|u\|$ apply the Hahn-Banach theorem:
Let $l : \langle u \rangle \to \mathbb{R}$ be the linear map with $l(u) = \|u\|$, thus:

$$\|l\| = \sup_{v \in \langle u \rangle, \|v\|=1} (l(v)) = \sup \left(l \left(\pm \frac{u}{\|u\|} \right) \right) = 1$$

By the Hahn-Banach theorem we can extend l to

$$\tilde{l} : E \to \mathbb{R}$$

with $\left\|\tilde{l}\right\| = 1$ and then holds:

$$\sup_{\varphi \in E^*, \|\varphi\|=1} \varphi(u) \overset{\|\tilde{l}\|=1}{\geq} \tilde{l}(u) = \|u\|$$

Therefore ι is injective, because from $\iota(u) = 0$ follows $\|u\|_E = \|\iota(u)\| = 0$ and therefore $u = 0$. $\square_{2.2.3}$

2.2.4 Definition (reflexive)

A Banach space is called *reflexive* (reflexiv) if ι is bijective, i.e. $E \overset{\sim}{=} E^{**}$.

2.2.5 Example

Let ℓ_1 be the space of absolutely convergent functions with the norm:

$$\|(a_n)\|_1 = \sum_{n=1}^{\infty} |a_n| < \infty$$

Let $(\lambda_n) \in \ell_\infty$ be a bounded sequence and define $\Lambda \in \ell_1^*$:

$$\Lambda : \ell_1 \to \mathbb{R}$$
$$\Lambda((a_n)) = \sum_{n=1}^{\infty} \lambda_n a_n$$

$$|\Lambda((a_n))| = \left| \sum_{n=1}^{\infty} \lambda_n a_n \right| \leq \sum_{n=1}^{\infty} |\lambda_n| \cdot |a_n| \leq \|(\lambda_n)\|_\infty \sum_{n=1}^{\infty} |a_n| = \|(\lambda_n)\|_\infty \cdot \|(a_n)\|_1 < \infty$$

Thus Λ is bounded and:

$$\|\Lambda\| = \sup_{n \in \mathbb{N}} |\lambda_n|$$

Claim: Every bounded linear functional on ℓ_1 is of this form, i.e. $\ell_1^* = \ell_\infty$.

Proof: Let $\Lambda \in \ell_1^*$. Choose $u_l \in \ell_1$ by $u_l = (0, \ldots, 0, 1, 0, \ldots)$ with a one at the l-th position. Setting $\lambda_l := \Lambda(u_l)$ gives:

$$|\lambda_l| = |\Lambda(u_l)| \leq \underbrace{\|\Lambda\|}_{<\infty} \cdot \underbrace{\|u_l\|}_{=1} \leq \|\Lambda\| < \infty$$

So $(\lambda_l) \in \ell_\infty$.
Let (a_k) be a finite sequence, with only zeros for $k > K \in \mathbb{N}$. Then:

$$\varLambda\left((a_k)\right) = \varLambda\left(\sum_{k=1}^{K} a_k u_k\right) = \sum a_k \varLambda\left(u_k\right) = \sum \lambda_k a_k$$

Since the finite sequences are dense in ℓ_1, the claim follows. $\square_{\text{Claim}}$

So $\ell_1^* = \ell_\infty$ and one could assume $\ell_\infty^* = \ell_1$, but this is not the case (see exercises).

Thus $\ell_1^{**} \neq \ell_1$, which means, that ℓ_1 is *not* reflexive.

2.3 Weak Convergence (Schwache Konvergenz)

Let E be a Banach space and (u_n) a sequence in E.

Normal convergence: $u_n \to u$ if and only if $\|u - u_n\| \xrightarrow{n\to\infty} 0$.

2.3.1 Definition (weak convergence, weak Cauchy sequence)

A sequence (u_n) in E *converges weakly* to u, written as $u_n \rightharpoonup u$, if for all $\varphi \in E^*$ the sequence $\varphi\left(u_n\right)$ converges to $\varphi\left(u\right)$, i.e. $\varphi\left(u_n\right) \to \varphi\left(u\right)$.

(u_n) is a *weak Cauchy sequence* if for all $\varphi \in E^*$ the sequence $\varphi\left(u_n\right)$ is a Cauchy sequence.

2.3.2 Theorem (Uniqueness of weak limit)

The weak limit is unique.

Proof

Let (u_n) be a sequence in E, which converges weakly to u and u', i.e. for all $\varphi \in E^*$ holds:

$$\varphi\left(u_n\right) \to \varphi\left(u\right) \qquad\qquad \varphi\left(u_n\right) \to \varphi\left(u'\right)$$

$$\Rightarrow \quad 0 = \varphi\left(u_n - u_n\right) \to \varphi\left(u - u'\right)$$

So $\varphi\left(u - u'\right) = 0$ for all $\varphi \in E^*$.

Claim: $v := u - u' = 0$

Proof: Assume to the contrary that $v \neq 0$.
Choose $\varphi : \langle v\rangle \to \mathbb{R}$ with $\varphi\left(v\right) = 1$. By the Hahn-Banach theorem φ can be extended continuously to E.
Therefore exists a $\varphi \in E^*$ with $\varphi\left(v\right) = 1$, which is a contradiction to $\varphi\left(v\right) = 0$. $\square_{\text{Claim}}$

$\square_{2.3.2}$

2.3.3 Theorem (convergence implies weak convergence)

Every convergent sequence converges weakly.

Proof

Suppose that $u_n \to u$. For $\varphi \in E^*$ follows:

$$|\varphi(u_n) - \varphi(u)| = |\varphi(u_n - u)| \le \underbrace{\|\varphi\|}_{\in\mathbb{R}} \cdot \|u_n - u\| \to 0$$

$$\begin{aligned} \Rightarrow \quad & \varphi(u_n) \to \varphi(u) \\ \Rightarrow \quad & u_n \rightharpoonup u \end{aligned}$$

$\square_{2.3.3}$

2.3.4 Example

$E = \left\{ (a_n) \,\middle|\, a_n \xrightarrow{n\to\infty} 0 \right\} \subsetneq \ell_\infty$ with $\|(a_n)\| = \sup_n |a_n|$ is a Banach space.
Let $u_n = (0,\ldots,0,1,0,\ldots)$ be the sequence with a one at the n-th position and zeros elsewhere.
For $n \neq m$ we have:

$$\|u_n - u_m\| = \sup\{0, |1|, |-1|\} = 1$$

Thus (u_n) is *not* a Cauchy sequence. Every $\varphi \in E^*$ can be represented with $(\lambda_k) \in \ell_1$ as (see exercises):

$$\varphi((a_n)) = \sum_k \lambda_k a_k$$
$$\|\varphi\| = \sum_{k=1}^{\infty} |\lambda_k| < \infty$$

$$\varphi(u_n) = \sum_{k=1}^{\infty} \lambda_k \delta_{kn} = \lambda_n \xrightarrow{n\to\infty} 0$$

From $(\lambda_n) \in \ell_1$ follows $\lambda_n \to 0$. This means that $u_k \rightharpoonup 0$.
This is used in the lectures on partial differential equations.
From $\mathscr{S}(u_n) \to \inf \mathscr{S}$ follows not necessarily $u_n \to u$, but $u_n \rightharpoonup u$.

Consider $A_n \in L(E,F)$.

- *norm convergence*: $A_n \to A$ in $L(E,F)$ means $\|A_n - A\| \to 0$.
- *strong convergence*: $A_n u \to Au$ in F for all $u \in E$.
- *weak convergence*: $A_n u \rightharpoonup Au$ for all $u \in E$, i.e. for all $\varphi \in F^*$ holds $\varphi(A_n u) \to \varphi(Au)$.

2.4 The Baire Category Theorem

Let E be a metric space (e.g. a normed space).

2.4.1 Definition (nowhere dense, set of first/second category)

A subset $A \subseteq E$ is called *nowhere dense* (nirgends dicht) if $\overline{A}^\circ = \emptyset$.
A is called *of first category* (or *meager*) if it can be written as a countable union of nowhere dense sets. Otherwise it is *of second category*.

Example

- $\mathbb{N} \subseteq \mathbb{R}$ is nowhere dense: $\overline{\mathbb{N}} = \mathbb{N}$, $\mathbb{N}^\circ = \emptyset$
- $\mathbb{Q} \subseteq \mathbb{R}$ is dense: $\overline{\mathbb{Q}} = \mathbb{R}$, $\overline{\mathbb{Q}}^\circ = \mathbb{R}^\circ = \mathbb{R}$

2.4.2 Theorem (René Baire, 1899)

Let $E \neq \emptyset$ be a complete metric space (Polish space). Then E is of second category.

Proof

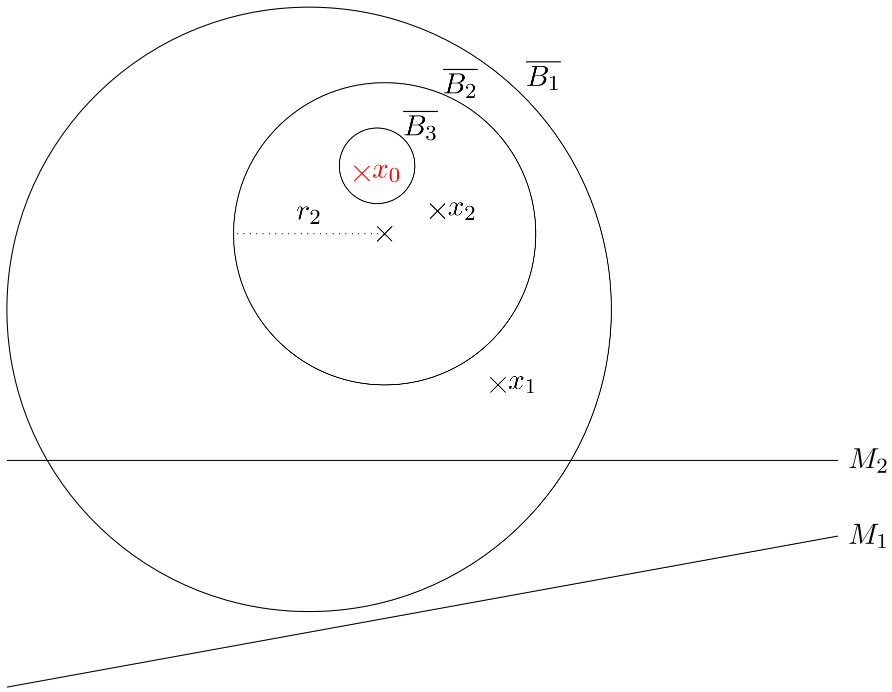


Figure 2.2: $B_n \cap M_n = \emptyset$

Assume in contrast that $E = \bigcup_{n\in\mathbb{N}} M_n$ and the sets M_n are nowhere dense. Without loss of generality assume that the M_n are closed, since otherwise one can replace M_n by $\overline{M_n}$.
We shall construct inductively balls $\overline{B_n} = \overline{B_{r_n}(x_n)}$ such that $\overline{B_{n+1}} \subseteq \overline{B_n}$, $r_n < 2^{-n}$ and $\overline{B_n} \cap M_n = \emptyset$ for all n.

Then the points x_n form a Cauchy sequence, because for all $n < m \in \mathbb{N}$ we have $x_{n+1} \in B_n$

and so $\|x_n - x_{n+1}\| < r_n < 2^{-n}$:

$$\|x_n - x_m\| \leq \|x_n - x_{n+1}\| + \|x_{n+1} - x_m\| \leq \ldots \leq$$
$$\leq 2^{-n} + 2^{-(n+1)} + \ldots + 2^{-(m-1)} \leq 2^{-n}\left(1 + \frac{1}{2} + \frac{1}{4} + \ldots\right) \leq 2 \cdot 2^{-n}$$

Since E is complete, $x_n \to x_0 \in E$ converges. Then $x_0 \in \overline{B_n}$ for all n, which implies $x_0 \notin M_n$ and thus the contradiction $x_0 \notin \bigcup_n M_n = E$ follows.

Construction of the balls $\overline{B_n}$:
M_1 is nowhere dense and therefore $B_1(0) \not\subseteq M_1$. So there exists a $x_1 \in B_1(0) \setminus M_1$. Since M_1 is closed, $B_1(0) \setminus M_1$ is open and therefore there exists a radius r_1 such that $B_{2r_1}(x_1)$ is contained in $B_1(0) \setminus M_1$ and thus $\overline{B_{r_1}(x_1)} \cap M_1 = \emptyset$.

Suppose $\overline{B_n}$ has been constructed. M_{n+1} is nowhere dense and closed and so there exists a $x_{n+1} \in \overline{B_n} \setminus M_{n+1}$ and $r_{n+1} < 2^{-(n+1)}$ such that $B_{2r_{n+1}}(x_{n+1}) \subseteq \overline{B_n} \setminus M_{n+1}$. Then follows $\overline{B_{r_{n+1}}(x_{n+1})} \cap M_{n+1} = \emptyset$. $\square_{2.4.2}$

2.4.3 Theorem (Uniform boundedness principle, Prinzip der gleichmäßigen Beschränktheit)

Let E be a Banach space and F a normed space. Let T_i be a sequence in $L(E,F)$ which is point-wise bounded, i.e. for all $u \in E$:

$$\sup_i \|T_i u\| \leq C(u) < \infty$$

Then sup-norms of T_i are bounded:

$$\sup_i \|T_i\| = \sup_i \sup_{\|u\|=1} \|T_i u\| \leq \tilde{C} < \infty$$

(Thus there exists a constant $C \in \mathbb{R}_{>0}$ such that $\|T_i u\| \leq C$ for all $i \in \mathbb{N}$ and for all $u \in E$ with $\|u\| = 1$.)

Proof

The sets $M_n = \left\{u \in E \,\middle|\, \sup_i \|T_i u\| \leq n\right\}$ are closed by continuity of the $T_i \in L(E,F)$, i.e. for $u_k \to u$ converges $\|T_i u_k\| \xrightarrow{k\to\infty} \|T_i u\|$.
$E = \bigcup_n M_n$, because for any $u \in E$, $\sup_i \|T_i u\| < \infty$ and thus $u \in M_n$ for $n > \sup_i \|T_i u\|$.
If all the sets M_n had empty interior, we would get a contradiction to Baire's theorem.
So there exists an $n_0 \in \mathbb{N}$ such that $\overset{\circ}{M}_{n_0} \neq \emptyset$ and thus there are $u_0 \in E$ and $r \in \mathbb{R}_{>0}$ such that $B_r(u_0) \subseteq M_{n_0}$.

For all $v \in B_r(u_0)$ we know that $\sup_i \|T_i v\| \leq n_0$ which is equivalent to:

$$\sup_{v \in B_r(u_0)} \|T_i v\| \leq n_0 \qquad \forall_{i \in \mathbb{N}}$$

Let $w \in B_r(0)$ be arbitrary. Then $v := u_0 + w \in B_r(u_0)$.

$$T_i w \overset{T_i \text{ linear}}{=} T_i v - T_i u_0$$

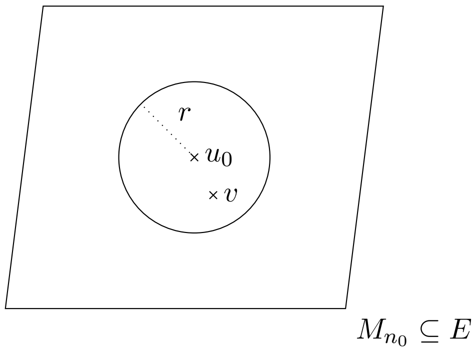


Figure 2.3: $B_r(u_0) \subseteq M_{n_0}$

$$\|T_i w\| \leq \|T_i v\| + \|T_i u_0\| \leq n_0 + \sup_i \|T_i u_0\| < \infty$$

Here $\sup_i \|T_i u_0\| < \infty$, because the T_i are point-wise bounded.

$$\Rightarrow \quad \|T_i w\| \leq C \qquad \forall_{w \in B_r(0)}$$

$$\Rightarrow \quad \|T_i \tilde{w}\| \leq \tilde{C} = \frac{C}{r} \qquad \forall_{\tilde{w} \in \overline{B_1(0)}}$$

So $\|T_i\| \leq \tilde{C}$ for all $i \in \mathbb{N}$ and so $\|T_i\|$ is bounded. $\square_{2.4.3}$

2.4.4 Corollary

Let E be a normed space, not necessarily complete, and (u_n) a weak Cauchy sequence. Then $\|u_n\|$ is a bounded sequence.

Proof

$E^* = L(E,\mathbb{R})$ is a Banach space after theorem 2.2.2, since $\mathbb{R}$ is complete. Now we can view every u_n as operator:

$$u_n : E^* \to \mathbb{R}$$
$$\varphi \mapsto \varphi(u_n)$$

So (u_n) is a sequence in $L(E^*,\mathbb{R})$. For all $\varphi \in E^*$ we know that $\varphi(u_n)$ is a Cauchy sequence and thus bounded:

$$\Rightarrow \quad |\varphi(u_n)| < C(\varphi)$$

Applying theorem 2.4.3 yields:

$$|\varphi(u_n)| < C \qquad \forall_{\varphi \text{ with } \|\varphi\|=1}$$

$$\Leftrightarrow \quad \sup_{n \in \mathbb{N}} \sup_{\varphi \in E^*, \|\varphi\|=1} |\varphi(u_n)| < C$$

For any $v \in E$ we have

$$\sup_{\varphi \in E^*, \|\varphi\|=1} |\varphi(v)| = \|v\|$$

by the Hahn-Banach theorem:

- $|\varphi(v)| \le \|\varphi\| \cdot \|v\| \overset{\|\varphi\|=1}{=} \|v\|$
- Choose $\varphi : \langle v \rangle \to \mathbb{R}$ with $\varphi(v) = \|v\|$ and so $\|\varphi\| = 1$. By the Hahn-Banach theorem we can extend φ to $\tilde{\varphi} : E \to \mathbb{R}$ such that $\|\tilde{\varphi}\| = 1$. Then $\tilde{\varphi}(v) = \|v\|$ and so $\sup_{\|\varphi\|=1} |\varphi(v)| \ge \|v\|$.

Thus we get $\sup_n \|u_n\| < C$. $\square_{2.4.4}$

2.4.5 Corollary and Definition (Banach-Steinhaus, equicontinuous, uniformly continuous)

Let E,F be Banach spaces and $T_i \in L(E,F)$.

If the (T_i) are point-wise bounded, then the T_i are *equicontinuous* (gleichgradig stetig).

Definition (uniformly continuous, equicontinuous)

Let $f : \mathbb{R} \to \mathbb{R}$ be a real-valued function.

Continuity:

$$\underset{x_0 \in \mathbb{R}}{\forall} \ \underset{\varepsilon \in \mathbb{R}_{>0}}{\forall} \ \underset{\delta \in \mathbb{R}_{>0}}{\exists} : \quad |x - x_0| < \delta \quad \Rightarrow \quad |f(x) - f(x_0)| < \varepsilon$$

f is called *uniformly continuous* (gleichmäßig stetig) if:

$$\underset{\varepsilon \in \mathbb{R}_{>0}}{\forall} \ \underset{\delta \in \mathbb{R}_{>0}}{\exists} : \quad \|x - y\| < \delta \quad \Rightarrow \quad \|f(x) - f(y)\| < \varepsilon$$

Let $f_n : \mathbb{R} \to \mathbb{R}$ be a series of real-valued functions. (f_n) is called *equicontinuous* if:

$$\underset{x_0 \in \mathbb{R}}{\forall} \ \underset{\varepsilon \in \mathbb{R}_{>0}}{\forall} \ \underset{\delta \in \mathbb{R}_{>0}}{\exists} \ \underset{n \in \mathbb{N}}{\forall} : \quad \|x - x_0\| < \delta \quad \Rightarrow \quad \|f_n(x) - f_n(x_0)\| < \varepsilon$$

For a linear map $A \in L(E,F)$ holds:

$$\|Au\| \le \|A\| \, \|u\|$$
$$\|Au - Au_0\| \le \|A\| \, \|u - u_0\|$$

Therefore choose $\delta = \frac{\varepsilon}{2\|A\|}$, i.e.:

$$\underset{\varepsilon \in \mathbb{R}_{>0}}{\forall} \ \underset{\delta \in \mathbb{R}_{>0}}{\exists} : \quad \|u\| < \delta \quad \Rightarrow \quad \|Au\| < \varepsilon$$

Proof

Since (T_i) is point-wise bounded there is a $C \in \mathbb{R}_{>0}$ such that for all $i \in \mathbb{N}$ holds $\|T_i\| \le C$ due to the principle of uniform boundedness 2.4.3. So for all $i \in \mathbb{N}$ holds:

$$\|T_i u\| \le \|T_i\| \, \|u\| \le C \, \|u\|$$

Choose $\delta = \frac{\varepsilon}{2C}$ shows that the T_i is equicontinuous. $\square_{2.4.5}$

In the following let E and F be Banach spaces.

2.4.6 Definition (open)

A (not necessarily linear) map $A : E \to F$ is called *open* if the image of every open set is open. (If there exists an inverse A^{-1} then "A open" is equivalent to "A^{-1} continuous".)

Let A be linear and open. $B_1(0) \subseteq E$ is open, so $A(B_1(0)) \subseteq F$ is open.
Since $0 \in A(B_1(0))$, there is a $\varepsilon \in \mathbb{R}_{>0}$ such that $B_\varepsilon(0) \subseteq A(B_1(0))$.

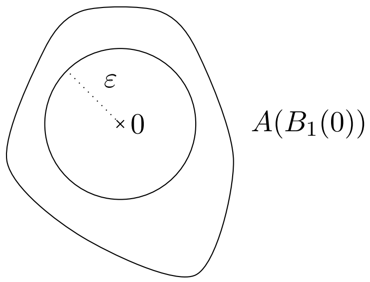


Figure 2.4: $B_\varepsilon(0) \subseteq A(B_1(0))$

Due to the linearity holds in general:

$$B_\lambda(0) \subseteq A\left(B_{\frac{\lambda}{\varepsilon}}(0)\right)$$

In particular, A is surjective.

If A is additionally injective, then A is bijective and the openness means that A^{-1} is continuous.

2.4.7 Theorem (Open mapping theorem, Prinzip der offenen Abbildung)

If $A \in L(E,F)$ is surjective, then A is open.

2.4.8 Corollary

If $A \in L(E,F)$ is bijective, then $A^{-1} \in L(F,E)$ is continuous.

Proof

A is open following 2.4.7, since A is surjective. This means that A^{-1} is continuous. $\square_{2.4.8}$

Proof of 2.4.7

Since A is surjective, $F = A(E)$. Since every element of E has a finite norm, we know:

$$E = \bigcup_{n\in\mathbb{N}} B_n(0)$$

$$\Rightarrow \quad F = A\left(\bigcup_{n\in\mathbb{N}} B_n(0)\right) = \bigcup_{n\in\mathbb{N}} A(B_n(0))$$

According to Baire's theorem there is a $n \in \mathbb{N}$ such that $\overline{A(B_n(0))}^\circ \neq \emptyset$.

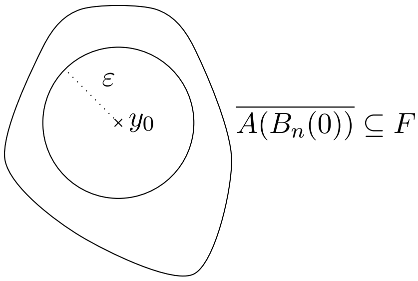


Figure 2.5: $B_\varepsilon(y_0) \subseteq \overline{A(B_n(0))}$

So there exists a $y_0 \in A(B_n(0))$ and a $\varepsilon \in \mathbb{R}_{>0}$ such that $B_\varepsilon(y_0) \subseteq \overline{A(B_n(0))}$. Since A is surjective, there is a $x_0 \in B_n(0)$ with $y_0 = A(x_0)$.

$$\Rightarrow \quad \overline{A(B_n(0) - x_0)} = \overline{A(B_n(0)) - y_0} = \overline{A(B_n(0))} - y_0 \supseteq B_\varepsilon(y_0) - y_0 = B_\varepsilon(0)$$

If n' is large enough, then $B_n(-x_0) \subseteq B_{n'}(0)$ and so $\overline{A(B_{n'}(0))} \supseteq B_\varepsilon(0)$.
Since A is linear, we can rescale, i.e. there is a $c := \frac{\varepsilon}{n'} \in \mathbb{R}_{>0}$ such that for all $r \in \mathbb{R}_{<0}$ holds:

$$\overline{A(B_r(0))} \supseteq B_{cr}(0)$$

Now we show that every $u \in B_c(0)$ is the image of a $x \in B_2(0)$, i.e. $B_c(0) \subseteq A(B_2(0))$:
Ansatz as a series:

$$x = \sum_{j=1}^{\infty} x_j$$

Choose $x_1 \in B_1(0)$ with $\|u - Ax_1\| < \frac{c}{2}$, which is possible since $\overline{A(B_1(0))} \supseteq B_c(0)$.
Choose $x_2 \in B_2(0)$ with $\|u - Ax_1 - Ax_2\| < \frac{c}{4}$, which is possible since $u - Ax_1 \in B_{\frac{c}{2}}(0)$ and $\overline{A\left(B_{\frac{1}{2}}(0)\right)} \subseteq B_{\frac{c}{2}}(0)$.
And so on choose $x_m \in B_{\frac{1}{2^m}}(0)$ with $\|u - \sum_{i=1}^m Ax_i\| < \frac{c}{2^m}$.
The series $\sum_{i=1}^\infty x_i$ converges, since:

$$\left\| \sum_{j=m}^{M} x_j \right\| \leq \sum_{j=m}^{M} \|x_j\| \leq \sum_{j=m}^{M} 2^{-j}$$

So the sequence of partial sums is a Cauchy sequence. Because E is complete, this sequence converges.

The continuity of A yields:

$$Ax = \sum_{j=1}^{\infty} Ax_j = u$$

So there exists a $x \in E$ with $\|x\| < 2$ and $Ax = u$. $\square_{2.4.7}$

$$\begin{aligned} \sum_{j=1}^{n} x_j &\xrightarrow{n\to\infty} x && \|x\| < 2 \\ \sum_{j=1}^{n} Ax_j &\xrightarrow{n\to\infty} u \\ \| \\ A\left(\sum_{j=1}^{n} x_j\right) &\xrightarrow[\text{continuity of } A]{n\to\infty} Ax \end{aligned}$$

Definition (Graph)

For a function $f : \mathbb{R} \to \mathbb{R}$ the *graph* is defined as:

$$\operatorname{graph} f := \left\{(x,f(x)) \,\middle|\, x \in \mathbb{R}\right\} \subseteq \mathbb{R} \times \mathbb{R}$$

For $A : E \to F$ the *graph* is:

$$\operatorname{graph} A := \left\{(u,Au) \,\middle|\, u \in E\right\} \subseteq E \times F$$

Here $E \times F$ is a product of normed spaces which has the norm:

$$\|(u,v)\| := \|u\|_E + \|v\|_F$$

Lemma

If A is continuous, then $\operatorname{graph} A$ is closed.

Proof

Let $(u_n,Au_n) \in \operatorname{graph} A$ be a Cauchy sequence in $E\times F$ for Banach spaces E and F, i.e. $u_n \to u$. Since A is continuous, it follows:

$$Au_n \to v := Au$$

Therefore $(u,v) \in \operatorname{graph}(A)$ and so the graph is closed. $\square_{\text{Lemma}}$

Consider the function:

$$\begin{aligned} f : \mathbb{R} \setminus \{0\} &\to \mathbb{R} \\ x &\mapsto \frac{1}{x} \end{aligned}$$

f is not continuous, but $\operatorname{graph}(f)$ is closed in $(\mathbb{R} \setminus \{0\}) \times \mathbb{R}$.

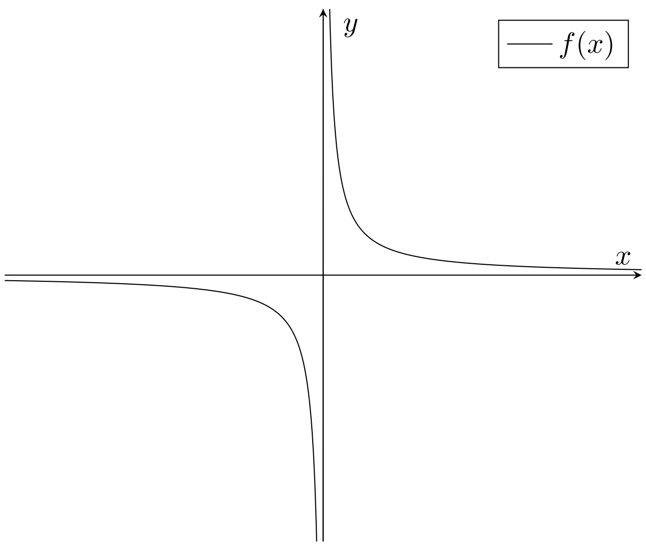


Figure 2.6: f is not continuous, but graph$f$ is closed.

2.4.9 Theorem (Closed graph theorem, Satz vom abgeschlossenen Graphen)

Suppose a linear map $A : E \to F$ between Banach spaces E and F has a closed graph. Then A is continuous.

graph (A) closed means:
For all $u_n \in E$ with $u_n \to u$ *and* $Au_n \to v$, the point $(u,v) \in \operatorname{graph}(A)$, i.e. $Au = v$.

A continuous means:
For all $u_n \in E$ with $u_n \to u$, the sequence $Au_n \to v$ converges and $Au = v$

Proof

On $E \times F$ we have the norm:

$$\|(u,v)\| := \|u\|_E + \|v\|_F$$

The graph

$$G := \left\{(u,Au) \,\middle|\, u \in E\right\} \subseteq E \times F$$

is a subspace of $E \times F$, since for $\lambda \in \mathbb{R}$ and $u,\tilde{u} \in E$ holds:

$$\lambda\,(u,Au) + (\tilde{u},A\tilde{u}) = (\lambda u + \tilde{u},\lambda Au + A\tilde{u}) \overset{A \text{ linear}}{=} (\lambda u + \tilde{u},A\,(\lambda u + \tilde{u})) \in G$$

So G is complete and therefore a Banach space, since we assumed it to be closed.
Define:

$$P : G \to E$$
$$(u,Au) \mapsto u$$

$$\|(u,Au)\| = \|u\| + \|Au\| \geq \|u\| = \|P\,(u,Au)\|$$

So for all $w \in G$ holds $\|Pw\| \leq \|w\|$ and therefore $\|P\| \leq 1$. In particular, P is continuous.
P is obviously surjective and it is also injective, since:

$$P^{-1}(u) = (u,Au)$$

Following the open mapping theorem, P^{-1} is continuous, i.e. there exists a $C \in \mathbb{R}_{>0}$ such that:

$$\|u\| + \|Au\| = \|(u,Au)\| = \left\|P^{-1}(u)\right\| \leq C\,\|u\|$$

Then follows:

$$\|Au\| \leq (C-1)\,\|u\|$$

Therefore A is continuous. $\square_{2.4.9}$

2.5 Neumann series

Let E be a Banach space and $A \in L(E,E) =: L(E)$.

When is A continuously invertible?
Remember that for $x \in \mathbb{K}$ with $|x| < 1$ holds:

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$$

This is the geometric series.
Idea: $A = \mathbb{1} - B$ with $B \in L(E)$

Ansatz: $A^{-1} := \sum_{n=0}^{\infty} B^n$

This works indeed if $\|B\| < 1$.

2.5.1 Lemma and Definition (Neumann series)

The series

$$C := \sum_{n=0}^{\infty} B^n$$

is called Neumann series (Neumannsche Reihe).

If $\|B\| < 1$, then C defines an element of $L(E,E)$, i.e. the Neumann series converges absolutely.

Proof

Consider the partial sums:

$$S_n := \sum_{k=0}^{n} B^k$$

Since $L(E,E)$ is a Banach space, it is enough to show that S_n is a Cauchy series. Without loss of generality assume $m > n$:

$$\|S_n - S_m\| = \left\|\sum_{k=n}^{m} B^k\right\| \overset{\Delta \text{ inequality}}{\leq} \sum_{k=n}^{m} \left\|B^k\right\| \overset{\text{Schwarz}}{\leq} \sum_{k=n}^{m} \|B\|^k < c\,\|B\|^n \to 0$$

$\square_{2.5.1}$

2.5.2 Theorem

$$C = (\mathbb{1} - B)^{-1}$$

Proof

$$(\mathbb{1} - B)\, C = (\mathbb{1} - B) \sum_{n=0}^{\infty} B^n = \left(\mathbb{1} + B + B^2 + \ldots\right) - \left(B + B^2 + \ldots\right) = \mathbb{1}$$

$\square_{2.5.2}$

2.5.3 Theorem

The set of all continuously invertible mappings is open in $L(E)$.

Proof

Assume that $A \in L(E)$ is continuously invertible, i.e. A^{-1} exists and $A^{-1} \in L(E)$. Set:

$$\varepsilon = \frac{1}{2\,\|A^{-1}\|}$$

Let us show, that every element of $B_\varepsilon(A) \subseteq L(E)$ is continuously invertible:
Let $C \in B_\varepsilon(A)$, i.e. $\|A - C\| < \varepsilon$.

$$C = A - (A - C) = A\big(\mathbb{1} - \underbrace{A^{-1}(A - C)}_{=:B}\big)$$

Then holds:

$$\|B\| \leq \left\|A^{-1}\right\| \cdot \|A - C\| < \left\|A^{-1}\right\| \cdot \frac{1}{2\,\|A^{-1}\|} = \frac{1}{2} < 1$$

Hence $\mathbb{1} - B$ is continuously invertible by the Neumann series and therefore

$$C^{-1} = (\mathbb{1} - B)^{-1} \cdot A^{-1}$$

is continuous. $\square_{2.5.3}$

3 Hilbert spaces

Definition (scalar product)

Let H be a real ($\mathbb{K} := \mathbb{R}$) or complex ($\mathbb{K} := \mathbb{C}$) vector space with *scalar product*:

$$\langle .,. \rangle : H \times H \to \mathbb{K}$$

i) Positive definiteness: $\langle u,u \rangle \geq 0$ and $\langle u,u \rangle = 0 \Rightarrow u = 0$.

ii) Linear in the second and anti-linear in the first argument:

$$\langle \lambda u,v \rangle = \overline{\lambda} \langle u,v \rangle$$

iii) Symmetry: $\overline{\langle u,v \rangle} = \langle u,v \rangle$

Define the corresponding norm:

$$\|u\| := \sqrt{\langle u,u \rangle}$$

3.0.1 Definition (Hilbert space)

A complete scalar product space is called *Hilbert space*.

The Schwarz inequality holds:

$$|\langle u,v \rangle| \leq \|u\| \cdot \|v\|$$

3.0.2 Lemma (parallelogram equality)

The parallelogram equality (Parallelogramm-Gleichung) is:

$$\|u+v\|^2 + \|u-v\|^2 = 2\left(\|u\|^2 + \|v\|^2\right)$$

Proof

$$\begin{aligned} \|u+v\|^2 &= \langle u+v,u+v \rangle = \langle u,u \rangle + \langle u,v \rangle + \langle v,u \rangle + \langle v,v \rangle \\ \|u-v\|^2 &= \langle u-v,u-v \rangle = \langle u,u \rangle - \langle u,v \rangle - \langle v,u \rangle + \langle v,v \rangle \\ \Rightarrow \quad \|u+v\|^2 + \|u-v\|^2 &= 2\left(\|u\|^2 + \|v\|^2\right) \end{aligned}$$

$\square_{3.0.2}$

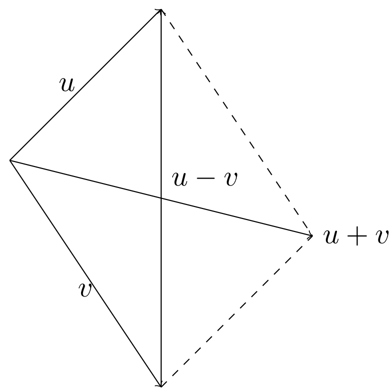


Figure 3.1: parallelogram

3.0.3 Definition (orthogonal, orthonormal)

i) Vectors $u,v \in H$ are called *orthogonal*, symbolically $u \perp v$, if $\langle u,v \rangle = 0$.

ii) Subspaces $M_1,M_2 \subseteq H$ are orthogonal, symbolically $M_1 \perp M_2$, if $\langle u,v \rangle = 0$ for all $u \in M_1$ and $v \in M_2$.

iii) A family $(u_i)_{i \in I}$ of vectors $u_i \in H$ is called *orthonormal* if:

$$\langle u_i,u_j \rangle = \delta_{ij}$$

3.0.4 Theorem (Bessel's inequality)

Let $(u_i)_{1 \le i \le N}$ be an orthonormal family. Then for all $u \in H$ holds:

$$\|u\|^2 = \sum_{i=1}^{N} \langle u_i,u \rangle^2 + \left\| u - \sum_{i=1}^{N} u_i \langle u_i,u \rangle \right\|^2$$

$$\|u\|^2 \ge \sum_{i=1}^{N} \langle u_i,u \rangle^2$$

Proof

$$\begin{aligned}\left\| u - \sum_{i=1}^{N} u_i \langle u_i,u \rangle \right\|^2 &= \left\langle u - \sum_{i=1}^{N} u_i \langle u_i,u \rangle , u - \sum_{j=1}^{N} u_j \langle u_j,u \rangle \right\rangle = \\ &= \langle u,u \rangle - \sum_{j=1}^{N} \langle u,u_j \rangle \langle u_j,u \rangle - \sum_{i=1}^{N} \overline{\langle u_i,u \rangle} \langle u_i,u \rangle + \sum_{i,j=1}^{N} \overline{\langle u_i,u \rangle} \langle u_j,u \rangle \underbrace{\langle u_i,u_j \rangle}_{=\delta_{ij}} = \\ &= \|u\|^2 - 2\sum_{i=1}^{N} |\langle u_i,u \rangle|^2 + \sum_{i=1}^{N} |\langle u_i,u \rangle|^2 = \\ &= \|u\|^2 - \sum_{i=1}^{N} |\langle u_i,u \rangle|^2\end{aligned}$$

$\square_{3.0.4}$

Definition (Hilbert space isomorphism)

Let $(H_1, \langle .,. \rangle_1)$ and $(H_2, \langle .,. \rangle_2)$ be Hilbert spaces.

A *Hilbert space isomorphism* is a mapping $U : H_1 \to H_2$ which is linear, bijective and isometric (isometrisch), i.e. for all $u,v \in H_1$:

$$\langle u,v \rangle_1 = \langle Uu,Uv \rangle_2$$

Definition (Direct sum)

Let $(H_1, \langle .,. \rangle_1)$ and $(H_2, \langle .,. \rangle_2)$ be Hilbert spaces.

Define:

$$H := \left\{ (u,v) \,\middle|\, u \in H_1, v \in H_2 \right\}$$

$$\begin{aligned} (u_1,v_1) + (u_2,v_2) &:= (u_1 + u_2, v_1 + v_2) \\ \lambda\,(u,v) &:= (\lambda u, \lambda v) \\ \langle (u_1,v_1)\,, (u_2,v_2) \rangle &:= \langle u_1,u_2 \rangle + \langle v_1,v_2 \rangle \end{aligned}$$

This makes $H =: H_1 \oplus H_2$ a Hilbert space, called *direct sum* of H_1 and H_2, which is sometimes called orthogonal due to:

$$\langle (u,0)\,, (0,v) \rangle = 0$$

3.0.5 Example

$$\ell_2 = \left\{ (a_n)_{n \in \mathbb{N}} \,\middle|\, a_n \in \mathbb{K}, \sum_{n=1}^{\infty} |a_n|^2 < \infty \right\}$$

Define a scalar product:

$$\langle (a_n)\,, (b_n) \rangle := \sum_{n=1}^{\infty} \overline{a}_n \cdot b_n$$

$$\langle (a_n)\,, (a_n) \rangle = \sum_{n=1}^{\infty} |a_n|^2 = \|a_n\|_2^2$$

$\left(\ell^2, \|.\|_2\right)$ is a Banach space. Thus $\left(\ell^2, \langle .,. \rangle\right)$ is a Hilbert space.

3.1 Projection on closed convex subsets

Let $(H, \langle .,. \rangle)$ be a Hilbert space and $K \subseteq H$ a closed convex subset.

$u,v \in K$ $\qquad$ $w \in H \setminus K$

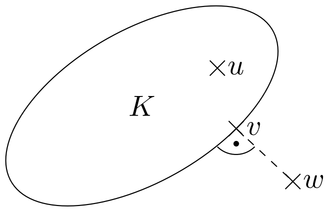


Figure 3.2: $\|v - w\| = \inf_{u \in K} \|u - w\|$

We want to find a vector v such that $\|v - w\| = \inf_{u \in K} \|u - w\|$.

If K were compact, then choose minimizing sequence (Minimalfolge), i.e.:

$$\|u_i - w\| \to \inf_{u \in K} \|u - w\|$$

Choose a convergent subsequence $u_{i_l} \to v$. Then by continuity:

$$\|v - w\| = \lim_{i \to \infty} \|u_i - w\| = \inf_{u \in K} \|u - w\|$$

The main application are closed subspaces $K \subseteq H$.

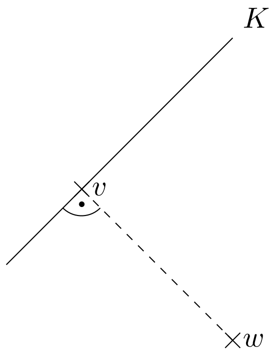


Figure 3.3: $v - w \perp K$

In this case $v - w$ will be called orthogonal to K motivating the name *orthogonal projection.*

3.1.1 Theorem (Hilbert)

There is a unique $v \in K$ with:

$$\|v - w\| = \inf_{u \in K} \|u - w\|$$

Proof

Consider a minimizing sequence u_i:

$$\|u_i - w\| \to \inf_{u \in K} \|u - w\| =: d$$

We show that (u_i) is a Cauchy sequence:

$$\begin{aligned}
\|u_i - u_j\|^2 &= \|(u_i - w) + (w - u_j)\|^2 = \\
&\overset{3.0.2}{=} 2\|u_i - w\|^2 + 2\|w - u_j\|^2 - \|(u_i - w) - (w - u_j)\|^2 = \\
&= 2\|u_i - w\|^2 + 2\|w - u_j\|^2 - \left\|-2\left(w - \frac{u_i + u_j}{2}\right)\right\|^2 = \\
&= 2\left(\underbrace{\|u_i - w\|^2}_{\to d^2} + \underbrace{\|w - u_j\|^2}_{\to d^2} - 2\left\|\frac{u_i + u_j}{2} - w\right\|^2\right)
\end{aligned}$$

$$\|u_i - w\| \xrightarrow{i\to\infty} d = \inf_{u\in K} \|u - w\|$$
$$\|u_j - w\| \xrightarrow{j\to\infty} d = \inf_{u\in K} \|u - w\|$$

Since K is convex and $u_i, u_j \in K$, we know:

$$\frac{u_i + u_j}{2} \in K$$

$$\Rightarrow \quad \left\|\frac{u_i + u_j}{2} - w\right\| \geq d$$

Thus:

$$\|u_i - u_j\|^2 \leq 2\left(\|u_i - w\|^2 + \|w - u_j\|^2 - 2d^2\right) \xrightarrow{i,j\to\infty} 2\left(d^2 + d^2 - 2d^2\right) = 0$$

So there exists a $N \in \mathbb{N}$ such that $\|u_i - u_j\| < \varepsilon$ for all $i, j > N$. Therefore (u_i) is a Cauchy sequence. Since H is complete, we know that $u_i \to u$ converges.
By continuity follows:

$$\|u - w\| = \lim_{i\to\infty} \|u_i - w\| = d$$

Uniqueness follows from the fact, that *every* minimizing sequence converges:
Let $u, \tilde{u}$ be both minimizers, then the sequence $(u, \tilde{u}, u, \tilde{u}, \ldots)$ is a minimizing sequence. Since it converges, $u = \tilde{u}$. $\square_{3.1.1}$

3.1.2 Corollary

Let $M \subseteq H$ be a closed subspace of H. Then a $w \in H$ can be decomposed uniquely in the form

$$w = v + x$$

with $v \in M$ and $x \in M^\perp$. We write $H = M \oplus M^\perp$.

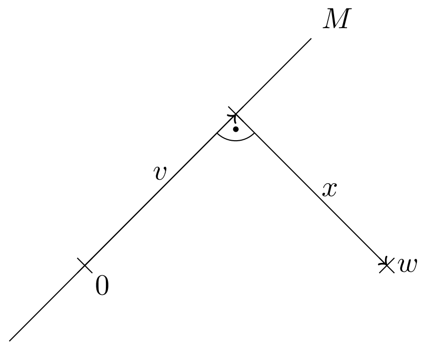


Figure 3.4: $w = v + x$

Proof

Let $v \in M$ be as in Theorem 3.1.1.

$$\|v - w\| = \inf_{u \in M} \|u - w\|$$

Define $x := w - v$.

- H real: For $u \in M$ define $\tilde{u}(\tau) = v + \tau u$ with $\tau \in \mathbb{R}$.

$$\begin{aligned} \|\tilde{u} - w\|^2 = \|x\|^2 + 2\tau \langle u,x \rangle + \tau^2 \|u\|^2 \geq \|x\|^2 \\ 0 \leq 2\tau \langle u,x \rangle + \tau^2 \|u\|^2 =: f(\tau) \end{aligned}$$

  $f(\tau)$ has a minimum at $\tau = 0$ and so $f'(0) = 0$.

$$\begin{aligned} & f'(0) = 2\langle u,x \rangle \\ \Rightarrow \quad & 2\langle u,x \rangle = 0 \quad \forall_{u \in M} \end{aligned}$$

  So $x \in M^\perp$.

- H complex: Define $\tilde{u}(\tau) = v + \tau u$, $\tau = re^{\mathbf{i}\varphi} \in \mathbb{K}$ with $r \geq 0$.

$$\|\tilde{u} - w\|^2 = \|x\|^2 + 2\mathrm{Re}\left(re^{-\mathbf{i}\varphi} \langle u,x \rangle\right) + r^2 \|u\|^2 =: f(r,\varphi)$$

  This has a minimum at $r = 0$.

$$\begin{aligned} \Rightarrow \quad & 0 = \partial_r f(0,\varphi) = 2\mathrm{Re}\left(e^{-\mathbf{i}\varphi} \langle u,x \rangle\right) \\ \overset{\varphi \text{ arbitrary}}{\Rightarrow} \quad & \langle u,x \rangle = 0 \end{aligned}$$

  So $x \in M^\perp$.

Uniqueness: Assume that $w = v_1 + x_1 = v_2 + x_2$ where $v_1,v_2 \in M$, $x_1,x_2 \in M^\perp$.

$$\underbrace{v_1 - v_2}_{\in M} = \underbrace{x_2 - x_1}_{\in M^\perp} \in M \cap M^\perp = \{0\}$$

Because from $u \in M \cap M^\perp$ follows $\langle u,u \rangle = 0$ and so $u = 0$. $\square_{3.1.2}$

For a Banach space E we have E,E^*,E^{**} and a natural injection $\iota : E \hookrightarrow E^{**}$.
For a Hilbert space H, suppose $u \in H$ and define:

$$\varphi : H \to \mathbb{K}$$
$$\varphi(v) := \langle u,v \rangle$$

φ is continuous, because:

$$|\varphi(v)| = |\langle u,v \rangle| \leq \|u\| \cdot \|v\| \leq C \|v\|$$

Now

$$\iota : H \hookrightarrow H^*$$
$$\iota(u) = \varphi$$

is a linear mapping, which is injective.

3.1.3 Theorem (Fréchet-Riesz)

For any $\varphi \in H^*$ there is a unique $v \in H$ such that for all $x \in H$:

$$\varphi(x) = \langle v,x \rangle$$

In other words: $\iota : H \to H^*$ is a Banach space isomorphism.

Proof

Let $\varphi \in H^*$, without loss of generality $\varphi \neq 0$.

$$M := \ker \varphi \subseteq H$$

is a subspace. It is closed by continuity: For $u_n \in \ker \varphi$ with $u_n \to u$ holds:

$$\varphi(u) \overset{\text{continuity}}{=} \lim_{n\to\infty} \varphi(u_n) = 0$$

So $u \in \ker \varphi$.

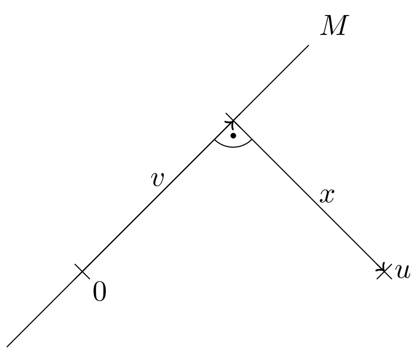


Figure 3.5: $u = v + x$

- $M^\perp$ is a one-dimensional subspace of H:
  $M^\perp \neq \{0\}$:
  Since $\varphi \neq 0$ there exists a $u \in H$ with $\varphi(u) \neq 0$, thus $u \notin M$.
  Now decompose $u = v + x$, $v \in M$, $x \in M^\perp \setminus \{0\}$.
  $M^\perp$ is one-dimensional: Take $u,v \in M^\perp$, $u,v \neq 0$, then $\varphi(u) \neq 0$ and $\varphi(v) \neq 0$.

$$\varphi(\varphi(v)\,u - \varphi(u)\,v) = 0$$

  So $\varphi(v)\,u - \varphi(u)\,v \in M \cap M^\perp = \{0\}$. Thus $\varphi(v)\,u - \varphi(u)\,v = 0$, implying that u and v are linearly dependent.
- Choose $u \in M^\perp$ with $\varphi(u) = 1$, which is always possible by rescaling.

$$\begin{aligned} v &:= \frac{u}{\|u\|^2} \\ \Rightarrow \quad \varphi(v) &= \frac{1}{\|u\|^2}\underbrace{\varphi(u)}_{=1} = \frac{1}{\|u\|^2} \\ \langle v,v\rangle &= \frac{\langle u,u\rangle}{\|u\|^4} = \frac{1}{\|u\|^2} = \varphi(v) \end{aligned}$$

- This v has the desired properties:
  For $x \in H$ decompose:

$$x = \underbrace{m}_{\in M} + \underbrace{\alpha v}_{\in M^\perp = \langle v\rangle}$$

$$\begin{aligned} \Rightarrow \quad \varphi(x) &= \underbrace{\varphi(m)}_{=0} + \alpha\varphi(v) = \alpha\,\langle v,v\rangle = \\ &= \langle v,\alpha v\rangle = \langle v,m+\alpha v\rangle = \langle v,x\rangle \end{aligned}$$

$\square_{3.1.3}$

3.1.4 Theorem (Lax-Milgram)

Let H be a Hilbert space and $B : H \times H \to \mathbb{K}$ be a mapping with the following properties:

i) $B(x,y)$ is linear in the second an anti-linear in the first argument.

ii) $|B(x,y)| \leq C\,\|x\| \cdot \|y\|$ (continuity)

iii) B is symmetric ($\overline{B(x,y)} = B(y,x)$) and positive definite, i.e. $B(x,x) \geq b\,\|x\|^2$ with $b \in \mathbb{R}_{>0}$.

iii') $|B(x,x)| \geq b\,\|x\|^2$ with $b \in \mathbb{R}_{>0}$.

Then every $l \in H^*$ can be represented uniquely as:

$$l(y) = B(x,y) \qquad \forall_{y \in H}$$

Proof

First the easy case iii):
We introduce a new scalar product $\langle .,. \rangle_B$ by:

$$\langle x,y \rangle_B := B(x,y)$$

Using ii) and iii) one sees that $\|.\|_B$ is equivalent to $\|.\|$, i.e. there exists a $C \in \mathbb{R}_{>0}$ such that:

$$\frac{1}{C}\|x\| \leq \|x\|_B \leq C\|x\|$$

According to the Fréchet-Riesz theorem, there exists a unique $v \in H$ with

$$\varphi(x) = \langle v,x \rangle_B = B(v,x)$$

for all $x \in H$.

More difficult case iii'): Given $x \in H$,

$$B(x,.) : H \to \mathbb{K}$$

is a linear bounded functional according to i) and ii), i.e. $B(x,.) \in H^*$.
According to the Fréchet-Riesz theorem there exists a unique $z \in H$ such that $B(x,y) = \langle z,y \rangle$ for all $y \in H$. This yields a mapping:

$$\begin{aligned} \varphi : H &\to H \\ x &\mapsto z \end{aligned}$$

$$B(x,y) = \langle \varphi(x), y \rangle$$

- φ is linear, because both B and $\langle .,. \rangle$ are anti-linear in their first arguments.
- $\varphi(H) \subseteq H$ is closed:

$$b\|x\|^2 \overset{\text{iii')}}{\leq} |B(x,x)| = |\langle z,x \rangle| \leq \|z\| \cdot \|x\|$$
$$b\|x\| \leq \|z\| \tag{3.1}$$

  Let $z_n \in \varphi(H)$ be a sequence with $z_n \to z \in H$. Choose x_n such that $\varphi(x_n) = z_n$, i.e. $B(x_n,y) = \langle z_n,y \rangle$ for all $y \in H$.
  Due to the anti-linearity in the first argument follows that:

$$B(x_n - x_m, y) = \langle z_n - z_m, y \rangle$$

  (3.1) yields that $\|x_n - x_m\| \leq \|z_n - z_m\|$.
  Hence (x_n) is a Cauchy sequence and so $x_n \to x \in H$ converges. Since B is continuous according to ii), we get:

$$\underbrace{B(x_n,y)}_{\to B(x,y)} = \underbrace{\langle z_n,y \rangle}_{\to \langle z,y \rangle}$$

  This gives:

$$\begin{aligned} B(x,y) &= \langle z,y \rangle \\ \varphi(x) &= z \end{aligned}$$

  Thus z is in $\varphi(H)$.

- $\varphi(H) = H$: Otherwise there would be a vector $y \in \varphi(H)^{\perp} \setminus \{0\}$ and thus for all $x \in H$ holds.

$$B(x,y) = \langle \varphi(x), y \rangle = 0$$

In particular for $x = y$ this gives:

$$\begin{aligned} & 0 = |B(y,y)| \geq b \|y\|^2 \\ \Rightarrow \quad & y = 0 \end{aligned}$$

This is a contradiction and so $\varphi(H) = H$.

- φ is injective: Suppose there are $x, x' \in H$ with $\varphi(x) = \varphi(x')$. Then follows:

$$B\left(x - x', y\right) = \left\langle \underbrace{\varphi(x) - \varphi\left(x'\right)}_{=0}, y \right\rangle = 0$$

Choose $y = x - x'$ so we get:

$$B\left(x - x', x - x'\right) = 0$$

Since B is positive definite, it follows $x = x'$.

- Let $l \in H^*$. According to Fréchet-Riesz there exists a unique $z \in H$ with $l(y) = \langle z, y \rangle$ for all $y \in H$ and we have

$$\langle z, y \rangle = B(x,y)$$

for $x = \varphi^{-1}(z)$. So $l(y) = B(x,y)$.

$\square_{3.1.4}$

3.1.5 Corollary

Every Hilbert space is reflexive.

Proof

Recall $\iota : H \hookrightarrow H^{**}$. H is *reflexive* if and only if ι is surjective, i.e. a Banach space isomorphism.

$$\begin{gathered} \tilde{\iota} : H \to H^* \\ (\tilde{\iota}(u))(v) = \langle u, v \rangle \end{gathered}$$

is bijective by Fréchet-Riesz. This holds also for $\bar{\iota} : H^* \to H^{**}$.

$$H \xrightarrow{\tilde{\iota}} H^* \xrightarrow{\bar{\iota}} H^{**}$$

So $\iota = \bar{\iota} \circ \tilde{\iota}$ is bijective as composition of bijective maps. $\square_{3.1.5}$

3.2 Orthonormal Bases in Separable Hilbert Spaces

3.2.1 Example

$$\ell_2 = \left\{ (a_n)_{n\in\mathbb{N}} \,\middle|\, \sum_{n\in\mathbb{N}} |a_n|^2 < \infty \right\}$$

with the scalar product

$$\langle (a_n), (b_n) \rangle := \sum_n \overline{a}_n b_n$$

is a Hilbert space.

Idea: Let H be an abstract Hilbert space. Choose an "orthonormal basis" (e_i).

$$H \ni u = \sum_{i=1}^{\infty} \lambda_i e_i$$
$$v = \sum_{i=1}^{\infty} \nu_i e_i$$

$$\langle u,v \rangle = \sum_{i,j=1}^{\infty} \langle \lambda_i e_i, \nu_j e_j \rangle = \sum_{i,j=1}^{\infty} \overline{\lambda_i} \nu_j \delta_{ij} = \sum_i \overline{\lambda_i} \nu_i$$

3.2.2 Definition (orthonormal system, Hilbert space basis, cardinality)

A system $(e_i)_{i\in J}$ is an *orthonormal system*, if $\langle e_i, e_j \rangle = \delta_{ij}$. The algebraic span is the vector space of *finite* linear combinations:

$$\langle (e_i) \rangle = \left\{ \sum_{i=1}^{N} \lambda_i e_i \middle| N \in \mathbb{N}, \lambda_i \in \mathbb{K} \right\}$$

This is a subspace of H. Now the subspace $\overline{\langle (e_i) \rangle} \subseteq H$ is called *Hilbert space span* (Hilbertraumerzeugnis).

An orthonormal system (e_i) is called a *orthonormal Hilbert space basis* if $\overline{\langle (e_i) \rangle} = H$.

Two sets A and B have the same cardinality if there exists an bijective map $\varphi : A \to B$.

Theorem (Bernstein-Schröder)

A and B have the same cardinality if and only if there exists an injective map from A to B and an injective map from $B \to A$.

(no proof)

A typical application of the Lax-Milgram theorem is for $x \in \mathbb{R}^n$, given real-valued functions $V(x)$, $f(x)$ and looking for $u(x)$ that solves:

$$-\Delta u(x) + V(x)u(x) = f(x)$$

Question: Is there a solution which "decays at infinity"?

1. Weak formulation:
   Suppose we have a solution $u \in \mathcal{C}^2(\mathbb{R}^n)$

   $$-\Delta u + Vu - f = 0$$

   Let $\eta \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ be a test function.

   $$0 = \int_{\mathbb{R}^n} (-\Delta u + Vu - f)\,\eta \mathrm{d}^n x \overset{\text{integration}}{\underset{\text{by parts}}{=}} \underbrace{\int_{\mathbb{R}^n} (\langle \nabla u, \nabla \eta\rangle + Vu\eta)\,\mathrm{d}^n x}_{=:B(u,\eta)} - \underbrace{\int_{\mathbb{R}^n} f\eta \mathrm{d}^n x}_{=l(\eta)}$$

   So for all $\eta \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ holds:

   $$B(u,\eta) = l(\eta)$$

   Definition: u is a *weak solution* of the equation $-\Delta u + Vu = f$ if for all $\eta \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ holds:

   $$B(u,\eta) = l(\eta)$$

2. Choose the correct Hilbert space. The first idea is $L^2(\mathbb{R}^n)$ with the scalar product:

   $$\langle u,v\rangle = \int_{\mathbb{R}^n} uv \mathrm{d}^n x$$

   $$u_n(x) := e^{-|x|^2} \sin(nx_1)$$

   Then for all $n \in \mathbb{N}$ holds:

   $$\|u_n\|_{L^2} \le C$$

   But $B(u_n,u_n) \xrightarrow{n\to\infty} \infty$ diverges. Thus B is *not* continuous.
   Better choose instead:

   $$\langle u,v\rangle = \int_{\mathbb{R}^n} (uv + \langle \nabla u, \nabla v\rangle)\,\mathrm{d}^n x$$

   The corresponding Hilbert space $H^{1,2}(\mathbb{R}^n)$ is a Sobolev space.

   $$L^2\left(\mathbb{R}^3\right) \supseteq H^{1,2}\left(\mathbb{R}^3\right) \ni u$$

   Assume for simplicity that $0 < \varepsilon \le V \le C < \infty$, then we get:

   $$B(u,u) = \int_{\mathbb{R}^n} \left(|\nabla u|^2 + Vu^2\right) \mathrm{d}^n x \le \int_{\mathbb{R}^n} \left(|\nabla u|^2 + Cu^2\right) \mathrm{d}^n x \le (1 + C)\,\|u\|^2_{H^{1,2}}$$

   $$|B(u,u)| \ge \int \left(|\nabla u|^2 + \varepsilon u^2\right) \ge \min\{1,\varepsilon\}\,\|u\|^2_{H^{1,2}}$$

   Thus the Lax-Milgram theorem applies and yields a unique weak solution and then a regularity theorem says that u is smooth.

Consider a matrix equation

$$Au = f$$

with $A \in \mathrm{Symm}\left(\mathbb{R}^n\right)$ and $f \in \mathbb{R}^n$.
For a general existence and uniqueness result one needs that A is invertible or equivalently:

$$\underset{u\in\mathbb{R}^n\setminus\{0\}}{\forall} :\ Au \neq 0$$

This follows from the condition:

$$\underset{u\in\mathbb{R}^n\setminus\{0\}}{\forall} :\ \underbrace{\langle u,Au\rangle}_{=B(u,u)} \neq 0$$

In finite dimension this is equivalent to:

$$\underset{u\in\mathbb{R}^n}{\forall} :\ |B\left(u,u\right)| > b\left\|u\right\|^2$$

$\left(e_i\right)_{i\in I}$ is an orthonormal Hilbert space basis of H if

$$\langle e_i,e_j\rangle = \delta_{ij}$$

and:

$$\overline{\langle e_i\rangle} = H$$

3.2.3 Theorem

Let $\left(e_i\right)_{i\in\mathbb{N}}$ be an orthonormal system. Then the mapping

$$\begin{aligned} \ell_2 &\to \overline{\langle e_i\rangle} \overset{\text{closed}}{\subseteq} H \\ (\lambda_i) &\mapsto \sum_{i\in\mathbb{N}} \lambda_i e_i \end{aligned}$$

is a Hilbert space isomorphism.

Proof

The mapping is well-defined and isometric:
For $(\lambda_i) \in \ell_2$, i.e. $\sum_{i\in\mathbb{N}} |\lambda_i|^2 < \infty$ we construct:

$$u_N := \sum_{i=1}^{N} \lambda_i e_i \in H$$

Without loss of generality take $M < N$, then follows:

$$\left\|u_N - u_M\right\|^2 = \left\|\sum_{i=M}^{N} \lambda_i e_i\right\|^2 = \left\langle \sum_{i=M}^{N} \lambda_i e_i, \sum_{i=M}^{N} \lambda_i e_i \right\rangle = \sum_{i,j=M}^{N} \overline{\lambda_i}\lambda_j \underbrace{\langle e_i,e_j\rangle}_{=\delta_{ij}} = \sum_{i=M}^{N} |\lambda_i|^2$$

Thus u_N is a Cauchy sequence and converges since $\overline{\langle e_i \rangle}$ is complete as a closed subset of a complete space.

$$u := \lim_{N\to\infty} u_N = \sum_{i=1}^{N} \lambda_i e_i$$

$$\|u\|^2 = \lim_{N\to\infty} \|u_N\|^2 = \lim_{N\to\infty} \sum_{i=1}^{N} |\lambda_i|^2 = \|(\lambda_i)\|_{\ell_2}$$

The mapping is also surjective:
Let $u \in \overline{\langle e_i \rangle}$ and $\varepsilon > 0$. So there exists a $v = \sum_{i=1}^{N} \lambda_i e_i \in \langle e_i \rangle$ with $\|v - u\| < \varepsilon$.
In other words there exists a finite $J \subseteq \mathbb{N}$ such that $d\left(\left\langle (e_i)_{i\in J} \right\rangle, u\right) < \varepsilon$. The vector which minimizes this distance is the orthogonal projection of u on $\left\langle (e_i)_{i\in J} \right\rangle$ since this is a finite-dimensional subspace, which is automatically closed.

$$u_J = \sum_{i\in J} e_i \left\langle e_i, u \right\rangle$$

Choose an increasing sequence $J_1 \subsetneq J_2 \subsetneq \ldots$ of finite sets such that:

$$\|u_{J_k} - u\| \to 0 \qquad\qquad \Rightarrow \quad u_{J_k} \to u$$

Thus u_{J_k} is bounded by a $C \in \mathbb{R}_{>0}$.

$$\begin{aligned} u_{J_k} &= \sum_{i\in J_k} e_i \underbrace{\left\langle e_i, u \right\rangle}_{=\lambda_i} \\ C > \|u_{J_k}\| &= \sum_{i\in J_k} |\lambda_i|^2 \end{aligned}$$

This gives:

$$\sum_{i\in\mathbb{N}} |\lambda_i|^2 < \infty$$

And so we get:

$$u = \sum_{i\in\mathbb{N}} \lambda_i e_i$$

$\square_{3.2.3}$

3.2.4 Theorem (Existence of Hilbert space basis)

In every Hilbert space H exists an orthonormal Hilbert space basis.

Proof

Consider $(u_i)_{i\in I}$ with $I = H$ and $u_h = h$ for all $h \in H$. $(u_i)_{i\in I}$ is obviously a generating system of H. On the set

$$X := \left\{ \tilde{I} \subseteq I \,\middle|\, (u_i)_{i\in\tilde{I}} \text{ is an orthonormal system} \right\}$$

defines „$\subseteq$“ a partial ordering.
Let $U \subseteq X$ be a totally ordered subset and define:

$$I_U := \bigcup_{\tilde{I} \in U} \tilde{I} \subseteq I$$

I_U is an upper bound of U in X if $I_U \in X$. Assume $(u_i)_{i \in I_U}$ would not be orthonormal. Then there would exist $j,k \in I_U$ with $\langle u_j,u_k \rangle \neq \delta_{jk}$.
For $j = k$ would hold $\langle u_j,u_j \rangle \neq 1$, but j lies in $\tilde{I} \in U \subseteq X$ and therefor has to hold$\langle u_j,u_j \rangle = 1$.
For $j \neq k$ we would get $\langle u_j,u_k \rangle \neq 0$. But j lies in $\tilde{I}_j \in U$ and k in $\tilde{I}_k \subseteq U$ and U is totally ordered, i.e. either holds $\tilde{I}_j \subseteq \tilde{I}_k$ or $\tilde{I}_k \subseteq \tilde{I}_j$.
Without loss of generality assume $\tilde{I}_j \subseteq \tilde{I}_k$ (otherwise exchange j and k). Then $j,k \in \tilde{I}_k \in U \subseteq X$ and hence $(u_i)_{i \in \tilde{I}_j}$ is an orthonormal system in contradiction to $\langle u_j,u_k \rangle \neq 0$. Therefore holds $I_U \in X$ and thus I_U is an upper bound of U.
Using Zorn's lemma we get a maximal element $I_{\max}$ in X. Because $(u_i)_{i \in I_{\max}}$ is an orthonormal system and thus especially linearly independent, it suffices to show that this is an generating system of H.
Assume there exists a $i_0 \in I$ with $u_{i_0} \notin K := \overline{\left\langle (u_i)_{i \in I_{\max}} \right\rangle_{\text{alg.}}}$. Since $K \subseteq H$ is closed and convex, there is an unique projection v of u_{i_0} on K and thus $h := u_{i_0} - v \in K^\perp$. It holds $h = u_h$ with $h \in H = I$.
Because $I_{\max}$ is maximal, holds then $I_{\max} \cup \{h\} \notin X$ and hence there is a $j \in I_{\max}$ with$\langle h,u_j \rangle \neq 0$, because $h = j$ cannot hold due to $h \notin I_{\max}$. This is a contradiction to $h \in K^\perp$ and thus holds $K = H$.
Therefore $(u_i)_{i \in I_{\max}}$ is an orthonormal Hilbert space basis of H. $\Box_{3.2.4}$

3.2.5 Theorem

Let H be a Hilbert space.

i) For any $v \in H$ and for any orthonormal system $\{e_j \,|\, j \in J\}$, the set of elements $j \in J$ for which $\langle e_j,v \rangle = 0$ is finite or countable.

ii) Any two Hilbert space bases of H have the same cardinality (Mächtigkeit).

Proof

i) Consider $v \in J$. First we show that every $n \in \mathbb{N}$, the set $J_n := \left\{ j \in J \,\middle|\, \langle e_j,v \rangle > \frac{1}{n} \right\}$ is finite. Indeed, by Bessel's inequality, for every finite number of elements $e_{j_1}, \ldots, e_{j_N}$ of the given orthonormal system, we have:

$$\sum_{k=1}^{N} |\langle e_{j_k},v \rangle|^2 \leq \|v\|^2$$

Now suppose that for some $n \in \mathbb{N}$, the set J_n were not finite. Then for any $N \in \mathbb{N}$ we could find elements $e_{j_1}, \ldots, e_{j_N}$ such that $\langle e_{j_k},v \rangle > \frac{1}{n}$ for all $k \in \{1, \ldots, N\}$. Hence, for these elements holds:

$$\sum_{k=1}^{N} |\langle e_{j_k},v \rangle|^2 > N \cdot \frac{1}{n}$$

Clearly these becomes larger than $\|v\|$ if we make N sufficiently large. Hence all the sets J_n must be finite. But then, we see that the set

$$\left\{j \in J \,\middle|\, \langle e_j,v\rangle \neq 0\right\} = \bigcup_{n\in\mathbb{N}} J_n$$

is a countable union of finite sets, and as such can be at most countable. $\square_{\text{i)}}$

ii) If H has is finite-dimensional, every Hilbert basis is a Hamel basis of H and thus the claim follows from linear algebra.
If H is infinite-dimensional, let $(e_i)_{i\in I}$ and $(b_j)_{j\in J}$ be two Hilbert bases of H. (I and J have infinitely many elements.)
For $x \in H = \overline{\langle (e_i)_{i\in I}\rangle} = \overline{\left\langle (b_j)_{j\in J}\right\rangle}$ define:

$$B_x := \left\{j \in J \,\middle|\, \langle x,b_j\rangle \neq 0\right\}$$

By i), the set B_x is at most countable for any $x \in H$. Next, let $j \in J$ be given. Since $\overline{\langle (e_i)_{i\in I}\rangle} = H$, we must have $\langle b_j,e_i\rangle \neq 0$ for some $i \in I$. Otherwise, $b_j \in \overline{\langle (e_i)_{i\in I}\rangle}^{\perp} = \{0\}$, which is not possible since $b_j \neq 0$. Therefore, we have $j \in B_{e_i}$ for some $i \in I$, and since $j \in J$ was arbitrary, it follows that $J \subseteq \bigcup_{i\in I} B_{e_i} \subseteq I \times \mathbb{N}$. Here the second inclusion uses that all the sets B_{e_j} are at most countable. It follows:

$$|J| \leq |I| \cdot |\mathbb{N}| = |I|$$

If we exchange the roles of I and J above, we also obtain $|I| \leq |J|$. By the Schröder-Bernstein theorem, we can combine both estimates to obtain that $|I| = |J|$. $\square_{\text{ii)}}$

$\square_{3.2.5}$

3.2.6 Theorem

If H is separable, then there exists a countable orthonormal Hilbert space basis $(e_i)_{i\in\mathbb{N}}$.
Thus H is Hilbert space isomorphic to ℓ_2.

Proof

Since H is separable, there is a countable dense subset $(x_i)_{i\in\mathbb{N}}$.

1. Arrange that the x_i are linearly independent:
   Start with $n = 1$ and $k = 1$ set:

   $$y_1 = x_1$$

   If the $y_1,\ldots,y_{n-1},x_k$ are linearly independent, we set $y_n = x_k$ and increase n and k by one.
   If the $y_1,\ldots,y_{n-1},x_k$ are linearly dependent, we only increase k by one.
   Then the y_i are linearly independent and $\langle (y_i)\rangle = \langle (x_i)\rangle$.

2. Gram-Schmidt procedure for orthonormalization:

$$\begin{aligned} e_1 &:= y_1 \\ e_2 &:= \frac{y_2 - e_1 \langle u_1, y_2 \rangle}{\| y_2 - e_1 \langle u_1, y_2 \rangle \|} \\ e_n &:= \frac{y_n - \mathrm{Pr}_{\langle e_1, \ldots, e_{n-1} \rangle} y_n}{\left\| y_n - \mathrm{Pr}_{\langle e_1, \ldots, e_{n-1} \rangle} y_n \right\|} \end{aligned}$$

Since the y_i are linearly independent, $y_n - \mathrm{Pr}_{\langle e_1, \ldots, e_{n-1} \rangle} y_n$ is never zero.
Then by construction the e_i are orthonormal and $\langle e_i \rangle = \langle x_i \rangle \subseteq H$ is dense and so $(e_i)_{i \in \mathbb{N}}$ is a Hilbert space basis. $\square_{3.2.6}$

3.3 Weak Compactness of the Closed Unit Ball

For a Banach space E *weak convergence* for $(u_i)_{i \in \mathbb{N}}$ with $u_i \in E$ means:

$$u_n \rightharpoonup u \quad \Leftrightarrow \quad \underset{\varphi \in E^*}{\forall} : \varphi(u_n) \to \varphi(u)$$

In Hilbert spaces, we can identify H^* with H via the Fréchet-Riesz theorem.

3.3.1 Definition (weak (sequential) compactness)

$x_n \rightharpoonup x$ *converges weakly* if $\langle y, x_n \rangle \to \langle y, x \rangle$ converges for all $y \in H$.

Weak compactness is for us by definition the same as *weak sequential compactness* (schwache Folgenkompaktheit):
$K \subseteq H$ is *weakly compact* if every sequence (x_n) with $x_n \in K$ has a weakly convergent subsequence.

3.3.2 Proposition

Let H be *separable* and infinite-dimensional and let $(e_i)_{i \in \mathbb{N}}$ be an orthonormal Hilbert space basis.

Then $e_n \rightharpoonup 0$ converges weakly.

Proof

Take $y \in H$ and expand it in the basis:

$$\begin{aligned} y &= \sum_{i=1}^{\infty} y_i e_i \\ y_i &= \langle e_i, y \rangle \end{aligned}$$

We know $(y_i)_{i \in \mathbb{N}} \in \ell_2$ and in particular $y_i \xrightarrow{i \to \infty} 0$, since the elements of an absolutely convergent series converge to zero. Therefore holds:

$$\langle y, e_n \rangle = \overline{y_n} \xrightarrow{n \to \infty} 0$$

Thus $e_n \rightharpoonup 0$ converges weakly. $\square_{3.3.2}$

3.3.3 Theorem (Weak Compactness of the Closed Unit Ball)

If H is *separable*, then the closed unit ball $\overline{B_1(0)} = \left\{u \middle| \|u\| \leq 1\right\}$ is weakly compact.

Proof

Let (u_l) be a sequence with $u_l \in \overline{B_1(0)}$. Choose an orthonormal Hilbert space basis $(e_n)_{n\in\mathbb{N}}$.

$$u_l = \sum_{n=1}^{\infty} u_{ln} e_n \qquad u_{ln} = \langle e_n, u_l \rangle \qquad (u_{l,n})_{n\in\mathbb{N}} \in \ell_2$$

$$|u_{ln}| = |\langle e_n, u_l \rangle| \leq \underbrace{\|e_n\|}_{=1} \cdot \|u_l\| \leq 1$$

For $n = 1$: $(u_{l,1})_{l\in\mathbb{N}}$ is a bounded sequence of complex or real numbers. Therefore there exists a convergent subsequence of u_l, which we denote by $u_l^{(1)} \in H$. Then follows:

$$u_{l,1}^{(1)} = \left\langle e_1, u_l^{(1)} \right\rangle \xrightarrow{l\to\infty} v_1$$

For $n = 2$: Next we choose a subsequence $u_l^{(2)}$ of $u_l^{(1)}$ such that:

$$\left\langle e_2, u_l^{(2)} \right\rangle \xrightarrow{l\to\infty} v_2$$

Proceed inductively to obtain:

$$\left\langle e_n, u_l^{(n)} \right\rangle \to v_n$$

Then $w_l = u_l^{(l)} \in \overline{B_1(0)}$ for a sequence (w_l) in $\overline{B_1(0)}$.

Claim: $w_l \overset{l\to\infty}{\rightharpoonup} v := \sum_n v_n e_n$

Proof: We proceed as follows:

$$v_n = \lim_{l\to\infty} \left\langle e_n, u_l^{(n)} \right\rangle = \lim_{l\to\infty} \left\langle e_n, u_l^{(l)} \right\rangle = \lim_{l\to\infty} \langle e_n, w_l \rangle$$

This is because $u_l^{(l)} = u_{l'}^{(n)}$ for $l' \geq l$.

1. $(v_n) \in \ell_2$:

$$\sum_{n=1}^{N} |v_n|^2 = \sum_{n=1}^{N} \left| \lim_{l\to\infty} \langle e_n, w_l \rangle \right|^2 \overset{\text{finite sum}}{=} \lim_{l\to\infty} \underbrace{\sum_{n=1}^{N} |\langle e_n, w_l \rangle|^2}_{\overset{\text{Bessel's}}{\underset{\text{inequality}}{\leq}} \|w_l\|^2 \leq 1}$$

So we get for all $N \in \mathbb{N}$:

$$\sum_{n=1}^{N} |v_n|^2 \leq 1$$

And thus $(v_n) \in \ell_2$ and $v := \sum_{n=1}^{\infty} v_n e_n$ is well-defined and has $\|v\| \leq 1$.

2. $w_l \rightharpoonup v$, i.e. $\langle y, w_l - v\rangle \xrightarrow{l\to\infty} 0$ for all $y \in H$:

$$y = \sum_{n=1}^{\infty} y_n e_n$$

$$\begin{aligned} y_n &= \langle e_n, y\rangle \\ y_< &:= \sum_{n\leq N} y_n e_n \\ y_> &:= \sum_{n>N} y_n e_n \\ \|y\|^2 &= \|y_<\|^2 + \|y_>\|^2 \end{aligned}$$

$$\langle y, w_l - v\rangle = \sum_{n=1}^{\infty} y_n \langle e_n, w_l - v\rangle$$

Choose $N \in \mathbb{N}$ so large that

$$\|y_>\| = \left(\sum_{n>N} |y_n|^2\right)^{\frac{1}{2}} < \frac{\varepsilon}{4}$$

to get:

$$\begin{aligned} &|\langle y, w_l - v\rangle| \leq |\langle y_<, w_l - v\rangle| + |\langle y_>, w_l - v\rangle| \leq \\ &\qquad \leq \sum_{n=1}^{N} |y_n| \, |\langle e_n, w_l - v\rangle| + \underbrace{\|y_>\|}_{<\frac{\varepsilon}{4}} \cdot \underbrace{\|w_l - v\|}_{\leq 2} < \sum_{n=1}^{N} |y_n| \, |\langle e_n, w_l - v\rangle| + \frac{\varepsilon}{2} \end{aligned}$$

We know $|\langle e_n, w_l - v\rangle| \xrightarrow{l\to\infty} 0$ for each n. So we can choose $|\langle e_n, w_l - v\rangle| \leq \frac{\varepsilon}{2}$ for $n \leq N$ and for all $l > L(\varepsilon)$ for a sufficiently large $L(\varepsilon)$ and therefore:

$$|\langle y, w_l - v\rangle| \leq \varepsilon \qquad \underset{l>L(\varepsilon)}{\forall}$$

Therefore $\langle y, w_l\rangle \to \langle y, v\rangle$ converges, which means $w_l \rightharpoonup v$. $\square_{\text{Claim}}$

$\square_{3.3.3}$

The corresponding statement in Banach spaces is the *Banach-Alaoglu theorem*:

Banach proved it in 1932 for separable Banach spaces using diagonal sequences.

Alaoglu proved it in 1938 for any Banach space. The proof is based on Tychonov's theorem.

We have E, E^*, E^{**} and an injection $\iota : E \to E^{**}$.

Theorem (Banach-Alaoglu)

The closed unit ball in E^* is *weak-∗-sequentially compact.*

I.e. in simple terms:

If $\varphi_n \in \overline{B_1(0)} \subseteq E^*$, then there exists a subsequence φ_{n_l} such that $\varphi_{n_l}(u)$ converges for all $u \in E$.

Application: Consider

$$E = C^0(\mathbb{R}^n)$$

with the sup-norm:

$$\|f\| = \sup_{x \in \mathbb{R}^n} |f(x)|$$

$$E^* = \{\text{regular Borel measures}\}$$

Suppose μ_n is a sequence of measures with $\|\mu_n\| \leq C$ for all $n \in \mathbb{N}$. Then there exists a measure μ such that $\mu_{n_l} \to \mu$ converges as a measure.

4 Operators on Hilbert spaces

Let H be a Hilbert space.

$$L(H) := L(H,H)$$

is the Banach space of bounded linear operators. (An linear map on an infinite dimensional space is usually called *linear operator*.) For $A \in L(H)$ define the norm:

$$|||A||| := \sup_{\|u\|=1} \|Au\|$$

4.0.1 Example

$H = L^2(\mathbb{R},\mathrm{d}x)$ with the Lebesgue measure $\mathrm{d}x$.

$$\langle f,g\rangle = \int_{\mathbb{R}} \overline{f} g \mathrm{d}x$$

$$A := \frac{\mathrm{d}}{\mathrm{d}x}$$

We would like to introduce this as an operator on H.
The inequality $\|Au\| \leq C\|u\|$ is violated even for $u \in C_0^\infty(\mathbb{R})$ for any constant $C \in \mathbb{R}$.
Namely consider

$$u_n(x) = \eta(x)\sin(nx)$$

with $\eta \in C_0^\infty(\mathbb{R})$ and $\eta\big|_{[-1,1]} = 1$. Then $\|u_n\| < \infty$ and $\|Au_n\| \xrightarrow{n\to\infty} \infty$.

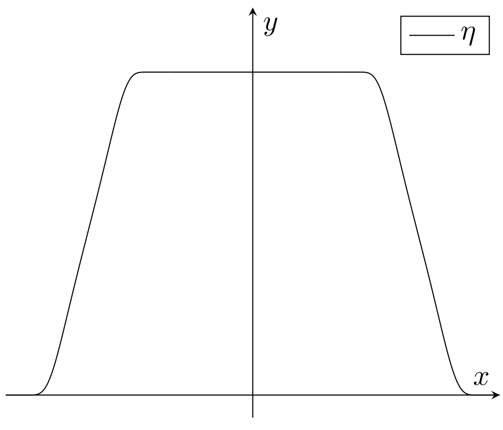


Figure 4.1: $\eta \in C_0^\infty(\mathbb{R})$ with $\eta\big|_{[-1,1]} = 1$

Moreover $\frac{\mathrm{d}}{\mathrm{d}x}f$ makes no sense for every vector f in H, because f does not need to be differentiable.

Way out: Define A only on a suitable subspace $\mathcal{D}(A)$ of H, called *domain* of definition.
For example: Choose $\mathcal{D}(A) = C_0^\infty(\mathbb{R}) \subseteq H$ and:

$$A : \mathcal{D}(A) \xrightarrow{\text{linear}} H$$

$\mathcal{D}(A)$ is dense in H, i.e. $\overline{\mathcal{D}(A)} = H$.

4.0.2 Definition (linear operator, domain, bounded)

i) Let $\mathcal{D} \subseteq H$ be a dense subspace. A linear map $A : \mathcal{D} \to H$ is called a *linear operator* on H with domain (of definition) $\mathcal{D}$.

ii) A is called *bounded*, if there exists a $C \in \mathbb{R}_{>0}$ such that for all $u \in \mathcal{D}$ holds:

$$\|Au\| \leq C \|u\|$$

Otherwise A is called unbounded.

4.0.3 Lemma

If A is a bounded operator with dense domain $\mathcal{D} \subseteq H$, then it can be extended by continuity to a unique operator $A \in L(H)$.

Proof

Let $u \in H$, not necessarily in $\mathcal{D}$. Since $\overline{\mathcal{D}} = H$, there is a sequence (u_l) in $\mathcal{D}$ with $u_l \to u$.

$$\|Au_i - Au_j\| = \|A(u_i - u_j)\| \leq C \cdot \|u_i - u_j\| \xrightarrow{i,j\to\infty} 0$$

Therefore we can set:

$$Au := \lim_{l\to\infty} Au_l$$

Since Au_l converges for any sequence $u_l \to u$, this is well-defined.

$$\|Au\| \leftarrow \|Au_i\| \leq C \|u_i\| \to C \|u\|$$

So there exists a C such that $\|Au\| \leq C \|u\|$ for all $u \in H$ and therefore $A \in L(H)$. $\square_{4.0.3}$

4.1 Isometric and unitary operators

4.1.1 Definition (isometric operator)

A operator $V : \mathcal{D}(V) \to H$ with dense domain $\mathcal{D}(V) \subseteq H$ is called *isometric* if for all $u \in \mathcal{D}(V)$ holds:

$$\langle Vu, Vu \rangle = \langle u, u \rangle$$

This operator is bounded, because:

$$\|Vu\| = \sqrt{\langle Vu,Vu\rangle} = \sqrt{\langle u,u\rangle} = \|u\| \overset{C:=1}{\leq} C\,\|u\|$$

Therefore we can extend it by continuity to H and

$$V : H \to H$$

is again isometric.

The "Hilbert hotel"

Consider $H = \ell_2$ and $(a_i) = (a_1,a_2,\ldots) \in \ell_2$.

$$A\,(u_1,u_2,\ldots) := (0,u_1,u_2,\ldots)$$

A is isometric, but it is no bijection.

Suppose you have a hotel with an infinite number of rooms and an infinite number of guest, in every room one guest.
If a new guest arrives, just move the guest from room n to room $n+1$ and the first room gets unoccupied, so the new guest can use it.

4.1.2 Proposition

For an isometric operator V the subspace $V\,(H) \subseteq H$ is closed.

Proof

Consider $y \in \overline{V\,(H)}$ and show $y \in V\,(H)$:
There exists a (y_n) with $y_n \in V\,(H)$ and $y_n \to y$ and a (x_n) with $V\,(x_n) = y_n$. Then holds:

$$\|x_i - x_j\| \overset{V \text{ isometric}}{=} \|V\,(x_i - x_j)\| = \|y_i - y_j\| \xrightarrow{i,j\to\infty} 0$$

Thus $x_i \to x$ converges. By continuity we get:

$$V\,(x) = \lim_{i\to\infty} V\,(x_i) = \lim_{i\to\infty} y_i = y$$

$\square_{4.1.2}$

4.1.3 Definition (unitary operator)

If $V : H \to H$ is an isometric operator and $V\,(H) = H$, then V is called *unitary* (unitär).

4.2 The Closure of an Operator

Let E and F be Banach spaces and $A : \mathcal{D}(A) \subseteq E \to F$ be a densely defined linear operator.

$$\begin{aligned} &\operatorname{graph}(A) := \left\{(u,Au) \,\middle|\, u \in \mathcal{D}(A)\right\} \subseteq E \times F \\ &\overline{\operatorname{graph}(A)} \subseteq E \times F \end{aligned}$$

Try to realize this as the graph of a new operator $\overline{A}$.

$$\mathcal{D}\left(\overline{A}\right) := \mathrm{pr}_1\left(\overline{\operatorname{graph} A}\right) = \left\{u \,\middle|\, \underset{v \in F}{\exists} : (u,v) \in \overline{\operatorname{graph} A}\right\}$$

For $u \in \mathcal{D}\left(\overline{A}\right)$ and $(u,v) \in \overline{\operatorname{graph} A}$ define:

$$\overline{A}u := v$$

v exists by definition of $\mathcal{D}\left(\overline{A}\right)$. Is v unique?
Suppose $(u,v) \in \overline{\operatorname{graph} A}$. Then there exists a sequence $(u_n,v_n) \in \operatorname{graph}(A)$, with $(u_n,v_n) \to (u,v)$. Equivalently:

$$\underset{n \in \mathbb{N}}{\forall} \; \underset{u_n \in \mathcal{D}(A)}{\exists} : (u_n \to u) \wedge (Au_n \to v)$$

Then we set $\overline{A}u := v$.

Problem: There might be two different series (u_n) and $(\tilde{u}_n)$ with $u_n \to u$, $\tilde{u}_n \to u$, $Au_n \to v$ and $A\tilde{u}_n \to \tilde{v} \neq v$.

4.2.1 Definition (closable operator)

A densely defined operator A is called closable (abschließbar) if $\overline{\operatorname{graph} A}$ is the graph of an operator B.
B is called the *closure* of A, symbolically $B = \overline{A}$.

4.2.2 Definition (closed)

A is called *closed* if $\operatorname{graph} A$ is a closed subset of $E \times F$.

4.2.3 Theorem (closed graph theorem)

Reformulation of 2.4.9:

If $\mathcal{D}(A) = E$, then A is closed if and only if A is bounded.

4.2.4 Example

Consider $E = C^0([0,1])$ with the norm $\|f\| = \sup_{x \in [0,1]} |f(x)|$.

$$\mathcal{D}(A) = C^1([0,1]) \subseteq E$$

$$\begin{aligned} A : \mathcal{D}(A) &\to E \\ f &\mapsto f' \end{aligned}$$

A is a densely defined, unbounded operator. Is A closed?

Consider $(u,v) \in \overline{\text{graph}A}$, i.e. there exists a sequence $(u_n) \subseteq \mathcal{D}(A)$ with $u_n \to u$ and $Au_n \to v$.
$u_n \to u$ means uniform convergence of $u_n \rightrightarrows u$, so u is continuous as a uniform limit of continuous functions.
$Au_n \to u$ means uniform convergence of $Au_n \rightrightarrows v$, so v is also continuous.
It follows that $u \in C^1$ and $u' = v$.
So $(u,v) \in \text{graph}A$ and therefore A is closed.

Consider $F := C^1([0,1])$ with $\|u\| = \sup_{[0,1]} |u| + \sup_{[0,1]} |u'|$. This is a Banach space.

Remark

The closure of a closable operator is always closed.
This is obvious, because $\text{graph}\overline{A} \overset{\text{def.}}{=} \overline{\text{graph}A}$, which is closed.

4.2.5 Theorem (Criterion for closable)

A is closable if and only if:

$$(u_n \in \mathcal{D}(A)) \wedge (u_n \to 0) \wedge (Au_n \to v) \quad \Rightarrow \quad v = 0$$

Proof

"$\Rightarrow$": Suppose A is closable. Thus there is an operator $\overline{A}$ such that $\text{graph}\overline{A} = \overline{\text{graph}A}$.
Suppose that $u_n \in \mathcal{D}(A)$, $u_n \to 0$ and $Au_n \to v$. Then $(u_n,Au_n) \to (0,v) \in \overline{\text{graph}A} = \text{graph}\overline{A}$ and thus $v = \overline{A}(0) = 0$.

"$\Leftarrow$": Suppose that the implication

$$(u_n \in \mathcal{D}(A)) \wedge (u_n \to 0) \wedge (Au_n \to v) \quad \Rightarrow \quad v = 0$$

holds.
Define $\mathcal{D}\left(\overline{A}\right)$ by: $u_n \in \mathcal{D}(A)$ with $u_n \to u$ and $Au_n \to v$. Then for $u \in \mathcal{D}\left(\overline{A}\right)$ set $\overline{A}(u) = v$.
This is well-defined: Suppose $u_n,\tilde{u}_n \to u$, $Au_n \to v$ and $A\tilde{u}_n \to \tilde{v}$. Then $u_n - \tilde{u}_n \to 0$ and $A(u_n - \tilde{u}_n) \to v - \tilde{v}$. By assumption follows $v - \tilde{v} = 0$. $\square_{4.2.5}$

4.3 The adjoint of a densely defined operator

Let $A : \mathcal{D}(A) \to H$ be a linear operator with $\overline{\mathcal{D}(A)} = H$.
In finite-dimensional linear algebra the definition of the adjoint A^* is:

$$\langle u,Av \rangle =: \langle A^*u,v \rangle \qquad \underset{u,v\in H}{\forall}$$

Here it is more complicated, since in general $\mathcal{D}(A) \neq H$.

$$M := \left\{ (u,w) \in H \times H \,\middle|\, \underset{v\in\mathcal{D}(A)}{\forall} : \langle u,Av \rangle = \langle w,v \rangle \right\}$$

Claim: M is the graph of a linear map A^*.

Proof: $M \neq \emptyset$ since $(0,0) \in M$.

- The image is unique: $u \mapsto w$ is well-defined, as from $(u,w), (u,w') \in M$ follows for all $v \in \mathcal{D}(A)$:

  $$\langle w - w', v \rangle = \langle u - u, Av \rangle = 0$$

  Since $\mathcal{D}(A)$ is dense, $w - w' = 0$ follows.
- A^* is linear: For $(u,w), (u',w') \in M$ and $\lambda \in \mathbb{K}$ follows $(u + \lambda u', w + \lambda w') \in M$, which is obvious from the definition of M. $\square_{\text{Claim}}$

4.3.1 Theorem

A^* is closed.

Proof

TODO: Proof from exercises

4.3.2 Theorem

A^* is the maximal, i.e. not extensible, operator S with the property that for all $u \in \mathcal{D}(A)$ and $v \in \mathcal{D}(S)$:

$$\langle Au, v \rangle = \langle u, Sv \rangle$$

Proof

TODO: Proof from exercises

4.4 Symmetric and self-adjoint densely defined operators

4.4.1 Definition (symmetric, (essentially) self-adjoint)

i) A is *symmetric* $:\Leftrightarrow \forall_{u,v \in \mathcal{D}(A)} : \langle Au, v \rangle = \langle u, Av \rangle$

ii) A is *self-adjoint* $:\Leftrightarrow A^* = A$ (in particular, $\mathcal{D}(A^*) = \mathcal{D}(A)$)

iii) A is *essentially self-adjoint* $:\Leftrightarrow \overline{A}$ is self-adjoint

For bounded A with $\mathcal{D}(A) = H$ all these notions coincide.

4.4.2 Example

Consider the operator $A := \Delta = \sum_{i=1}^{n} \partial_i^2$ on $L^2(\Omega)$ for a bounded open region $\Omega \subseteq \mathbb{R}^n$ with $\mathcal{D}(A) = C_0^\infty(\Omega) \overset{\text{dense}}{\subseteq} L^2(\Omega)$.

- A is symmetric:

$$\langle Af, g\rangle \overset{\text{integration by parts}}{=} \langle f, Ag\rangle$$

- Adjoint of Δ on L^2:

$$\int \mathrm{d}^n r\,(\Delta f)\cdot g = \int \mathrm{d}^n r f \cdot \underbrace{h}_{\in L^2}$$

  Here $h := A^* g$. It is sufficient to consider $g \in H^{2,2}(\Omega)$ (Sobolev space). $\mathcal{D}(A^*) \supsetneq \mathcal{D}(A)$

4.4.3 Lemma

Let A be a symmetric operator. Then A is closable and $\overline{A}$ and A^* are extensions of A and $\mathcal{D}(A) \overset{\text{i)}}{\subseteq} \mathcal{D}\left(\overline{A}\right) \overset{\text{ii)}}{\subseteq} \mathcal{D}(A^*)$.

Proof

Let $u_n \in \mathcal{D}(A)$ with $u_n \to 0$ and $Au_n \to w$.

$$\langle Au, v\rangle = \langle u, Av\rangle \qquad \underset{u,v\in\mathcal{D}(A)}{\forall}$$

$$\langle w, v\rangle \leftarrow \langle Au_n, v\rangle = \langle u_n, Av\rangle \to \langle 0, Av\rangle = 0$$

Since this holds for all $v \in \mathcal{D}(A)$ now $w = 0$ follows. From the criterion 4.2.5 follows that A is closable.

i) is obvious from the definition of $\overline{A}$.

ii) Take $u \in \mathcal{D}\left(\overline{A}\right)$. Then there is a sequence $u_n \in \mathcal{D}(A)$ with $u_n \to u$ and $Au_n \to \overline{A}u$.
For all $v \in \mathcal{D}(A)$ holds:

$$\left\langle \overline{A}u, v\right\rangle \leftarrow \langle Au_n, v\rangle = \langle u_n, Av\rangle \to \langle u, Av\rangle$$

So $u \in \mathcal{D}(A^*)$ and $A^* u = \overline{A}u$.

$\square_{4.4.3}$

„The smaller one chooses $\mathcal{D}(A)$, the larger becomes $\mathcal{D}(A^*)$."

$$B \subseteq \mathcal{D}(A) \quad \Rightarrow \quad \mathcal{D}\left(\left(A\big|_B\right)^*\right) \supseteq \mathcal{D}(A^*)$$

Difficulty: Construct $\mathcal{D}(A)$ such that $\mathcal{D}(A) = \mathcal{D}(A^*)$. (More on this later in the lecture.)

4.5 Heisenberg's uncertainty principle

In quantum mechanics:
The Hilbert space for one dimensional problems is usually $H = L^2(\mathbb{R})$.
The position operator is $x =: B$ and the momentum operator is $\frac{\hbar}{\mathbf{i}}\frac{\mathrm{d}}{\mathrm{d}x} =: A$.

$$[A,B] := AB - BA = \frac{\hbar}{\mathbf{i}}\mathbb{1}$$

4.5.1 Theorem (Winter-Wieland)

For two continuous operators A and B with $[A,B] = c \cdot \mathbb{1}$ and $B^n = B$ for all $n \in \mathbb{N}_{\geq 1}$, i.e. B is idempotent, follows $c = 0$.

Proof

Consider:

$$B^k A B^{n-k} = B^k (AB) B^{n-k-1} = B^k (BA + c\mathbb{1}) B^{n-k-1} = B^{k+1} A B^{n-k-1} + cB^{n-1}$$

$$\Rightarrow \quad cB^{n-1} = B^k A B^{n-k} - B^{k+1} A B^{n-k-1}$$

Sum this from $k = 0$ to $k = n-1$:

$$ncB^{n-1} = \sum_{k=0}^{n-1} B^k A B^{n-k} - B^{k+1} A B^{n-k-1} \overset{\text{telescope}}{\underset{\text{sum}}{=}} AB^n - B^n A$$

$$n\,|c|\,\left\|\left|B^{n-1}\right\|\right| = \||AB^n - B^n A\|| \overset{\Delta\text{-inequality}}{\leq} \|AB^n\| + \|B^n A\| \leq (\|AB\| + \|BA\|) \cdot \left\|B^{n-1}\right\|$$

Since this must hold for all n either $c = 0$ or there exists a $n \in \mathbb{N}_{>1}$ with $\left\|B^{n-1}\right\| = 0$, i.e. $B^{n-1} = 0$. Since B is idempotent follows $B = 0$ and therefore $[A,B] = 0$ and also $c = 0$. $\square_{4.5.1}$

Consider $u \in \mathcal{D}(A)$ with $\|u\| = 1$, which represents a quantum mechanical state.
The expectation value of A in u is after the probabilistic interpretation:

$$E_u(A) := \langle u, Au \rangle$$

The "uncertainty", i.e. the variance, is:

$$\Delta_u(A) := \|(A - E_u(A)\,\mathbb{1})\,u\|$$

4.5.2 Theorem (Heisenberg's uncertainty principle)

Let H be a $\mathbb{C}$-Hilbert space and $A : \mathcal{D}(A) \to H$, $B : \mathcal{D}(B) \to H$ be two symmetric operators with $\overline{\mathcal{D}(A)} = H = \overline{\mathcal{D}(B)}$. Assume for the image domains $\mathcal{R}$:

$$\mathcal{R}(A) \subseteq \mathcal{D}(B) \qquad\qquad \mathcal{R}(B) \subseteq \mathcal{D}(A)$$

So $[A,B]$ is well-defined on $\mathcal{D}(A) \cap \mathcal{D}(B)$.
Assume furthermore that $[A,B] = \frac{\hbar}{\mathbf{i}}\mathbb{1}$ with $\hbar > 0$.

Then for all $u \in \mathcal{D}(A) \cap \mathcal{D}(B)$ with $\|u\| = 1$ holds:

$$\Delta_u(A) \cdot \Delta_u(B) \geq \frac{\hbar}{2}$$

Proof

Replace A by $\tilde{A} := A - E_u(A) \cdot \mathbb{1}$ and $\tilde{B} := B - E_u(B) \cdot \mathbb{1}$. Then holds:

$$\left[\tilde{A},\tilde{B}\right] = \frac{\hbar}{\mathbf{i}}\mathbb{1}$$

$$\begin{aligned}\varDelta_u(A) &= \left\|\tilde{A}u\right\| \\ \varDelta_u(B) &= \left\|\tilde{B}u\right\|\end{aligned}$$

We have to show:

$$\varDelta_u(A) \cdot \varDelta_u(B) = \left\|\tilde{A}u\right\| \cdot \left\|\tilde{B}u\right\| \geq \frac{\hbar}{2}$$

$$\begin{aligned}\frac{\hbar}{2} &= \frac{\hbar}{2}\langle u,u\rangle = \frac{\mathbf{i}}{2}\left\langle u, \left(\tilde{A}\tilde{B} - \tilde{B}\tilde{A}\right)u\right\rangle \overset{\text{symmetry}}{=} \frac{\mathbf{i}}{2}\left(\left\langle \tilde{A}u,\tilde{B}u\right\rangle - \left\langle \tilde{B}u,\tilde{A}u\right\rangle\right) = \\ &= -\operatorname{Im}\left(\left\langle \tilde{A}u,\tilde{B}u\right\rangle\right) \overset{\text{Cauchy-Schwarz}}{\leq} \left\|\tilde{A}u\right\| \cdot \left\|\tilde{B}u\right\|\end{aligned}$$

$\square_{4.5.2}$

4.6 Spectrum and resolvent

Let $A : \mathcal{D}(A) \to H$ be a closed, densely defined operator.

4.6.1 Definition (continuously invertible, resolvent, spectrum)

A is *continuously invertible* if and only if $A : \mathcal{D}(A) \to H$ is bijective and $A^{-1} : H \to \mathcal{D}(A)$ is continuous.

$$\varrho(A) := \left\{\lambda \in \mathbb{K} \middle| (\lambda\mathbb{1} - A) \text{ is continously invertible}\right\}$$

The *resolvent* (Resolvente) is defined for $\lambda \in \varrho(A)$ as

$$\mathcal{R}_\lambda(A) = (\lambda\mathbb{1} - A)^{-1} \in L(H)$$

and the *spectrum* of A as:

$$\sigma(A) = \mathbb{K} \setminus \varrho(A)$$

4.6.2 Lemma

$\varrho(A)$ is open and $\sigma(A)$ is closed.

Proof

For bounded operators cf. Theorem 2.5.3.

It's method works even for unbounded operators:

Take $\lambda,\mu \in \varrho(A)$.

$$\begin{aligned}(A-\mu) &= (A-\lambda)+(\lambda-\mu) = \\ &= \underbrace{(A-\lambda)}_{\text{continuously invertible}} \cdot \left(\mathbb{1} + (A-\lambda)^{-1}(\lambda-\mu)\right)\end{aligned}$$

$\mathbb{1} + (A-\lambda)^{-1}(\lambda-\mu)$ is continously invertible using the Neumann series if:

$$|\lambda-\mu| < \frac{1}{\left\|(A-\lambda)^{-1}\right\|}$$

So $\varrho(A)$ is open and therefore the complement $\sigma(A)$ is closed. $\square_{4.6.2}$

4.6.3 Theorem (resolvent equation)

The map $\lambda \mapsto \mathcal{R}_\lambda(A)$ is complex analytic on $\varrho(A)$.

We have the *resolvent equation* (Resolventengleichung):

$$\mathcal{R}_\lambda - \mathcal{R}_\mu = -(\lambda-\mu)\,\mathcal{R}_\lambda \cdot \mathcal{R}_\mu$$

Proof

Analogy with $\mathbb{C}$-numbers:

$$\begin{aligned}\frac{1}{\lambda-x} - \frac{1}{\mu-x} &= \frac{\mu-\lambda}{(\lambda-x)(\mu-x)} \\ (\mu-x)-(\lambda-x) &= \mu-\lambda\end{aligned}$$

Same thing for operators:

$$\begin{aligned}(\mu-A)-(\lambda-A) &= \mu-\lambda \\ \mathcal{R}_\mu^{-1} - \mathcal{R}_\lambda^{-1} &= \mu-\lambda \qquad /\mathcal{R}_\mu\cdot \quad /\cdot\mathcal{R}_\lambda \\ \mathcal{R}_\lambda - \mathcal{R}_\mu &= (\mu-\lambda)\,\mathcal{R}_\mu\mathcal{R}_\lambda \\ \mathcal{R}_\lambda &= \mathcal{R}_\mu + (\mu-\lambda)\,\mathcal{R}_\mu\mathcal{R}_\lambda\end{aligned}$$

Assume $|\mu-\lambda| < \frac{1}{|||\mathcal{R}_\lambda|||}$.

$$\mathcal{R}_\mu = \mathcal{R}_\lambda\left(1+(\mu-\lambda)\,\mathcal{R}_\lambda\right)^{-1} = \mathcal{R}_\lambda \sum_{n=0}^{\infty} (-1)^n (\mu-\lambda)^n\,\mathcal{R}_\lambda$$

This series converges absolutely and so the map is analytic in $L(H)$. $\square_{4.6.3}$

5 Compact Operators

Let E and F be Banach spaces and $A \in L(E,F)$.

Remember: There exists a $C \in \mathbb{R}_{>0}$ such that for all $u \in E$ holds:

$$\|Au\| \leq C\|u\|$$

A maps bounded sets in E to bounded sets in F.

But: Bounded sets are not precompact in general.

5.1 Definition (compact operator)

A is called *compact* operator if and only if A maps bounded sets to relatively compact sets, i.e. the closure is compact.
(In complete spaces relatively compact is equivalent to precompact.)

5.2 Example (integral operator)

Let $E = \left(C^0([0,1]), \|.\|_\infty\right)$ and consider an integral kernel $K \in C^0([0,1] \times [0,1])$, $K : E \to E$.

$$(K\varphi)(x) := \int_0^1 K(x,y)\,\varphi(y)\,\mathrm{d}y$$

$$\begin{aligned} |(K\varphi)(x)| &\leq \sup_y |K(x,y)|\,\|\varphi\| \qquad / \sup_x \\ \|K\varphi\| &\leq C\|\varphi\| \end{aligned}$$

So $K \in L(E)$. Furthermore the integral kernel K is continuous and defined on a compact set. Therefore K is uniformly continuous after the Heine-Cantor theorem.

$$\underset{\varepsilon \in \mathbb{R}_{>0}}{\forall}\ \underset{\delta \in \mathbb{R}_{>0}}{\exists} : \left|K(x,y) - K\left(x',y\right)\right| < \varepsilon \qquad \underset{|x-x'|<\delta,\ y \in [0,1]}{\forall}$$

$$\left|(K\varphi)(x) - (K\varphi)\left(x'\right)\right| = \left|\int_0^1 \left(K(x,y) - K\left(x',y\right)\right)\varphi(y)\,\mathrm{d}y\right| \leq \varepsilon\|\varphi\|_\infty$$

Let now $B := B_M(0)$ with $M \in \mathbb{R}_{>0}$. Then $K(B) \subseteq E$.

- uniformly bounded ($\|\varphi\| < CM$)
- uniformly continuous

The Arzelà-Ascoli theorem yields, that $K(B)$ is precompact and so K is a compact operator.

5.3 Theorem

Let H be a Hilbert space.
A compact operator $A : H \to H$ maps weakly convergent sequences to convergent sequences.

Proof

Let $x_n \rightharpoonup x$, then (x_n) is bounded, i.e. there is a $C \in \mathbb{R}_{>0}$ such that $\|x_n\| < C$ for all $n \in \mathbb{N}$. Define $y_n := Ax_n$. For all $z \in H$ holds:

$$\langle z, y_n - y \rangle = \langle z, A (x_n - x) \rangle = \langle A^* z, x_n - x \rangle \to 0$$

Therefore $y_n \rightharpoonup y$ converges weakly. Because A is compact, every subsequence of y_n contains a convergent subsequence with limes $\tilde{y}$. For $z = \tilde{y} - y$ converges:

$$0 \leftarrow \langle z, y_n - y \rangle \to \langle \tilde{y} - y, \tilde{y} - y \rangle = \|\tilde{y} - y\|$$

Therefore $\tilde{y} = y$.
Since this holds for every subsequence of y_n follows $y_n \to y$. $\square_{5.3}$

5.4 Lemma

Consider operators $A, B : E \to F$.

i) If A and B are compact, so are $A + B$ and λA for all $\lambda \in \mathbb{K}$.

ii) If $A : E \to F$ is compact (continuous) and $B : F \to E$ continuous (compact), than $B \circ A$ is compact.
(In particular A^n is compact for $A : E \to E$.)

iii) The compact operators form a closed subspace of $L(E, F)$.

Proof

i) is obvious. $\square_{\text{i)}}$

ii) follows, since a continuous operator is bounded. $\square_{\text{ii)}}$

iii) Let (x_n) be bounded and T_k a convergent sequence of compact operators. By diagonal choice get a subsequence, also written x_n, such that $T_k x_n$ converges for all $k \in \mathbb{N}$.

$$\begin{aligned}\|Tx_n - Tx_m\| &\leq \underbrace{\|Tx_n - T_k x_n\|}_{\leq \|T - T_k\| \cdot \|x_n\|} + \|T_k x_n - T_k x_m\| + \underbrace{\|T_k x_m - Tx_m\|}_{\leq \|T - T_k\| \cdot \|x_m\|} \leq \\ &\leq \|T - T_k\| \cdot \|x_n\| + \|T_k x_n - T_k x_m\| + \|T - T_k\| \cdot \|x_m\| \xrightarrow{n,m,k \to \infty} 0\end{aligned}$$

$\square_{5.4}$

5.5 Lemma (Fredholm operator)

Let $A : E \to E$ be compact and define $T := \mathbb{1} - A$. T is called *Fredholm operator*.

i) $\ker(T)$ is finite-dimensional.

ii) There exists a $i \in \mathbb{N}$ such that $\ker\left(T^k\right) = \ker\left(T^i\right)$ for all $k \in \mathbb{N}_{>i}$.

iii) The image of T is closed.

Proof

i) $\ker(T) =: Z = \left\{u \middle| u = Au\right\}$. Since $Z \cap B_1(0)$ is bounded

$$A\left(Z \cap B_1(0)\right) = Z \cap B_1(0)$$

is precompact and therefore Z is finite-dimensional. $\square_{\text{i)}}$

ii) Define $N_i := \ker\left(T^i\right)$, which are closed subspaces of E, since the T^i are continuous. Suppose the claim is wrong, then $N_j \subsetneq N_{j+1} \subsetneq \ldots$, so in particular all N_j are proper subspaces. Choose $y_j \in N_j$ with:

$$\|y_j\| = 1 \qquad\qquad d\left(y_j, N_{j-1}\right) > \frac{1}{2}$$

This is possible after Lemma 2.1.2.
For all $m < n$ holds:

$$Ay_n - Ay_m = y_n - \underbrace{T_{y_n} - y_m + T_{y_m}}_{\in N_{n-1}}$$

Therefore follows:

$$\|Ay_n - Ay_m\| > \frac{1}{2}$$

So (Ay_n) has no accumulation value in contradiction to the compactness of A. $\square_{\text{ii)}}$

iii) Let $y_k \in \operatorname{im}(T)$ with $y_k \to y$ and $y_k = Tx_k$. We want to show $y \in \operatorname{im}(T)$. Define:

$$d_k := d\left(x_k, \ker(T)\right) = \inf_{z \in \ker(T)} \|x_k - z\|$$

Claim: (d_k) is bounded. Equivalently $(D_k) = |\max\{1, d_k\}|$ is bounded.

Proof: Choose $z_k \in \ker(T)$, $w_k := x_k - z_k$ with $\|w_k\| < 2d_k$ and $Tw_k = y_k$.
Assume D_k is unbounded. Since y_k is convergent and thus bounded, follows:

$$T\left(\frac{w_k}{D_k}\right) = \frac{y_k}{D_k} \xrightarrow{k\to\infty} 0$$

Now consider $u_k := \frac{w_k}{D_k}$. We know $\|u_k\| < 2$ and $T(u_k) \to 0$.
Thus $u_k - Au_k \to 0$. Since A is compact, every subsequence of Au_k has a convergent subsequence, and therefore $u_k \to 0$ converges.
The continuity of T gives:

$$T(u) = \lim_{k\to\infty} T(u_k) = 0$$

So $u \in \ker(T)$.
On the other hand we have for all $z \in \ker(T)$:

$$\|w_k - z\| \geq D_k$$
$$\Rightarrow \quad \left\| u_k - \frac{z}{D_k} \right\| \geq 1$$

Since T is a subspace this means, that for all $z \in \ker(T)$ holds:

$$\|u_k - z\| \geq 1$$

This is a contradiction to $u \in \ker(T)$. $\square_{\text{Claim}}$

So u_k is bounded and $T(w_k) = T(x_k) = y_k \to y$. So we get:

$$w_k - Aw_k \to y$$

Since A is compact Aw_k converges and with this follows, that $w_k \to w$ also converges. By continuity we get:

$$T(w) = \lim_{k\to\infty} T(w_k) = y$$

So $w \in \operatorname{im}(T)$. $\square_{5.5}$

5.6 Theorem (Fredholm Alternative)

Let $A : E \to E$ be compact and define $T := \mathbb{1} - A$.
If the kernel $\ker(T) = \{0\}$ is trivial, then T is continuously invertible.

Proof

$\ker(T) = \{0\}$ means, that T is injective. We only need to show, that T is surjective, because then T is invertible and 2.4.7 yields then, that T is open and therefore T^{-1} continuous.

$\operatorname{im}(T)$ is closed following 5.5 iii).
$\operatorname{im}(T) = E$, since otherwise $T(E) \subsetneq E$. Then the injectivity implies for all $k \in \mathbb{N}$:

$$T^{k+1}(E) \subsetneq \underbrace{T^k(E)}_{=E_k}$$

E_k is closed for all $k \in \mathbb{N}$:

$$E_k = (\mathbb{1} - A)^k (E) = \Bigg(\mathbb{1} + \underbrace{\sum_{l=1}^{k} (-1)^l \begin{pmatrix} k \\ l \end{pmatrix} A^l}_{A:=A_k}\Bigg)(E)$$

Now A_k is compact, as the compact operators form a (closed) ideal subalgebra $\mathrm{CP}(E)$.
Choose $x_k \in E_k$ with $\|x_k\| = 1$ and $d(x_k, E_k) > \frac{1}{2}$, which is possible after Lemma 2.1.2. Then holds for all $m < n$:

$$Ax_m - Ax_n = x_m - \underbrace{Tx_m - x_n + Tx_n}_{\in H_{m+1}}$$

$$\Rightarrow \quad \|Ax_m - Ax_n\| > \frac{1}{2}$$

This is a contradiction to the compactness of A.
Therefore T is surjective and the theorem follows. $\square_{5.6}$

5.7 Theorem (Riesz-Schauder)

Let $A \in L(H)$ be compact.

i) $\sigma(A)$ consists of a a finite or countable set of complex numbers and 0 is the only possible accumulation point.

ii) Every $0 \neq \lambda \in \sigma(A)$ is an eigenvalue of finite multiplicity, i.e. $\ker(A-\lambda)$ is finite-dimensional. That means, there exists a $i \in \mathbb{N}$ such that for all $k > i$ holds:

$$\ker(A-\lambda)^k = \ker(A-\lambda)^i$$

One says also that the Jordan chains are finite.

Proof

ii) is an immediate consequence of the Lemmas 5.5 and 5.6. (Divide A by λ.)

i) Assume $\lambda_n \neq 0$ are pairwise different eigenvalues. Choose eigenvectors $x_n \in H$ such that:

$$Ax_n = \lambda_n x_n$$

$$Y_n := \langle x_1, \dots, x_n \rangle$$

Since the eigenvalues are pairwise different $Y_n \subsetneq Y_{n+1}$ must hold, because the x_k are linearly independent.
Assume $Y_n \subsetneq H$, since otherwise H would be finite-dimensional and therefore $\sigma(A)$ a finite set.
So following Lemma 2.1.2 we can choose $y_n \in Y_n$ with $\|y_n\| = 1$ and:

$$d(y_n, Y_{n+1}) > \frac{1}{2}$$

Since $y_n \in Y_n$ one can find $\alpha_j \in \mathbb{K}$ such that:

$$y_n = \sum_j \alpha_j x_j$$

Then follows:

$$(A - \lambda_n)\, y_n = \sum_{i=1}^{n-1} (\lambda_j - \lambda_n)\, \alpha_j x_j =: \tilde{y}_n \in Y_{n-1}$$

For all $n > m$ holds:

$$Ay_n - Ay_m = \lambda_n y_n - \underbrace{\tilde{y}_n - Ay_m}_{\in Y_{n-1}}$$

So we get:

$$\|Ay_n - Ay_m\| \geq \frac{|\lambda_n|}{2}$$

But (Ay_n) is precompact and thus for all $\delta \in \mathbb{R}_{>0}$ exist only finitely many λ_n with $|\lambda_n| > \delta$. Therefore 0 is the only accumulation point and $\sigma(A)$ is a countable union of finite sets and thus countable. $\square_{5.7}$

Jordan decomposition:

$$A = \begin{pmatrix} \lambda_1 & & & & & & 0 \\ 1 & \ddots & & & & & \\ & 1 & \lambda_1 & & & & \\ & & & \lambda_2 & & & \\ & & & 1 & \ddots & & \\ & & & & 1 & \lambda_2 & \\ 0 & & & & & & \ddots \end{pmatrix}$$

$$\lambda_1 - A = \begin{pmatrix} 0 & & & & & & 0 \\ -1 & \ddots & & & & & \\ & -1 & 0 & & & & \\ & & & -\lambda_2 & & & \\ & & & -1 & \ddots & & \\ & & & & -1 & -\lambda_2 & \\ 0 & & & & & & \ddots \end{pmatrix}$$

So the first block is nilpotent. If it has k dimensions this means:

$$(\lambda_1 - A)^k = \begin{pmatrix} 0 & & 0 \\ & * & \\ 0 & & * \end{pmatrix}$$

So k is the length of the Jordan chain.

5.8 Theorem

Let $A \in L(H)$ be compact and H be a separable Hilbert space.
Then A can be approximated in $L(H)$ by operators of finite rank.

Proof

Choose a countable orthonormal Hilbert basis $(\varphi_j)_{j\in\mathbb{N}}$ of H, which is possible, since H is separable. Define:

$$\lambda_n := \sup_{\psi \in \langle \varphi_1, \ldots, \varphi_n \rangle^\perp, \|\psi\|=1} \|A\psi\|$$

Since A is bounded, this supremum exists. Obviously $\lambda_1 \geq \lambda_2 \geq \ldots$. Thus $\lambda_n \searrow \lambda \geq 0$.

Claim: $\lambda = 0$

Proof: Choose $\psi_n \in \langle \varphi_1, \ldots, \varphi_n \rangle^\perp$ with $\|\psi_n\| = 1$ and $\|A\psi_n\| \geq \frac{\lambda}{2}$ which is possible after Lemma 2.1.2, since $\langle \varphi_1, \ldots, \varphi_n \rangle$ is a proper closed subspace of H. Write:

$$\psi_n = \sum_{j=1}^{\infty} \nu_j \varphi_j = (\nu_1, \nu_2, \ldots)$$

Due to $\psi_n \in \langle \varphi_1, \ldots, \varphi_n \rangle^\perp$ follows:

$$\psi_n = (0, \ldots 0, \nu_{n+1}, \nu_{n+2}, \ldots)$$

For $u \in H$ holds:

$$\langle u, \psi_n \rangle = \sum_{j=n+1}^{\infty} \nu_j \cdot \overline{u}_j \underset{\text{inequality}}{\overset{\text{Schwarz}}{\leq}} \underbrace{\left(\sum_{j=n+1}^{\infty} |\nu_j|^2 \right)^{\frac{1}{2}}}_{=\|\psi_n\|} \cdot \left(\sum_{j=n+1}^{\infty} |u_j|^2 \right)^{\frac{1}{2}} \xrightarrow{n \to \infty} 0$$

So by construction $\psi_n \rightharpoonup 0$. Therefore $A\psi_n \to 0$ and thus $\|A\lambda_n\| \to 0$.
On the other hand we have $\|A\psi_n\| \geq \frac{\lambda}{2}$ and so $\lambda = 0$. $\square_{\text{Claim}}$

Let P_n be the orthogonal projection on $\langle \varphi_1, \ldots, \varphi_n \rangle$.

$$P_n u = \sum_{j=1}^{n} \varphi_j \langle \varphi_j, u \rangle$$

AP_n is an operator of finite rank $r \leq n$, since $\operatorname{rank}(P_n) = n$.

Claim: $AP_n \xrightarrow{n \to \infty} A$ in $L(H)$.

Proof: Consider:

$$|||A - AP_n||| = \sup_{u \in H, \|u\|=1} \|A(\mathbb{1} - P_n) u\|$$

$(\mathbb{1} - P_n) u \in \langle \varphi_1, \ldots, \varphi_n \rangle^\perp$ and $\|(\mathbb{1} - P_n) u\| \leq \|u\| = 1$. $(\mathbb{1} - P_n = P_{\langle \varphi_1, \ldots, \varphi_n \rangle^\perp})$
Thus we get:

$$|||A - AP_n||| \leq \sup_{v \in \langle \varphi_1, \ldots, \varphi_n \rangle^\perp, \|v\| \leq 1} \|Av\| = \lambda_n \xrightarrow{n \to \infty} 0$$

$\square_{\text{Claim}}$

$\square_{5.8}$

5.9 Lemma

Let $A \in L(H)$ be compact and symmetric. (This implies that A is bounded and self-adjoint.) Then $\sigma(A) \subseteq \mathbb{R}$ and if u is an eigenvector, $Au = \lambda u$, then its orthogonal is invariant under A.

Proof

For $\lambda \in \sigma(A)$ holds $\ker(A-\lambda) \neq \{0\}$. Thus there exists a $u \in \ker(\lambda - A) \setminus \{0\}$.

$$\lambda \langle u,u \rangle = \langle u,Au \rangle = \langle Au,u \rangle = \overline{\lambda} \langle u,u \rangle$$

Since $\|u\| \neq 0$ follows $\lambda = \overline{\lambda}$, which means that $\lambda \in \mathbb{R}$.
For $v \in \langle u \rangle^{\perp}$ holds:

$$\langle Av,u \rangle = \langle v,Au \rangle = \lambda \langle v,u \rangle = 0$$

Therefore $Av \in \langle u \rangle^{\perp}$. $\square_{5.9}$

5.10 Theorem (Hilbert-Schmidt)

Let $A \in L(H)$ be a symmetric compact operator on the separable Hilbert space H.
Then there exists an orthonormal Hilbert space basis of eigenvectors $(u_n)_{n\in\mathbb{N}}$, so with the eigenvalues $\lambda_n \in \mathbb{R}$ holds:

$$Au_n = \lambda_n u_n$$

Proof

$\sigma(A)$ is countable and therefore we can write $\sigma(A) \setminus \{0\} = \{\lambda_1,\lambda_2,\ldots\} \subseteq \mathbb{R}$ with $\lambda_i \neq \lambda_j$ for $i \neq j$. $\ker(\lambda_j - A)$ is finite-dimensional. So we choose a (finite) orthonormal basis of the eigenspace. Taking these eigenvectors for all eigenvalues, we obtain a countable orthonormal system $(u_n)_{n\in\mathbb{N}}$.

$$M := \overline{\langle u_n \rangle} \overset{\text{closed}}{\subseteq} H$$

$M^{\perp}$ is an invariant subspace of H under A, i.e.:

$$\tilde{A} := A\big|_{M^{\perp}} : M^{\perp} \to M^{\perp}$$

This is again symmetric and compact. We know that $\sigma\left(\tilde{A}\right) = \{0\}$.

Question: Why is $\tilde{A} = 0$?
This is not true for a general operator, e.g.:

$$A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \qquad\qquad \sigma(A) = \{0\}$$

Answer: If A is symmetric and $\sigma(A) = \{0\}$, then one can show $A = 0$ using the following theorem 5.12:
From $\sigma\left(\tilde{A}\right) = 0$ follows $r\left(\tilde{A}\right) = 0$ and since $\tilde{A}$ is self-adjoint theorem 5.12 gives $\left\|\tilde{A}\right\| = 0$ and thus $\tilde{A} = 0$. In other words $A\big|_{M^{\perp}} = 0$.
Now choose an orthonormal Hilbert basis $(v_n)_{n\in\mathbb{N}_{\leq N}}$ of $M^{\perp}$ for an $N \in \mathbb{N}\cup\{\infty\}$. Therefore $\{u_n\} \cup \{v_n\}$ is the desired orthonormal Hilbert basis of H. $\square_{5.10}$

5.11 Definition (spectral radius)

Let $A : \mathcal{D}(A) \subset H \to H$ be a densely defined operator. Then the *spectral radius* $r(A)$ of A is defined by:

$$r(A) = \sup_{\lambda \in \sigma(A)} |\lambda| \in \mathbb{R}_{\geq 0} \cup \{\infty\}$$

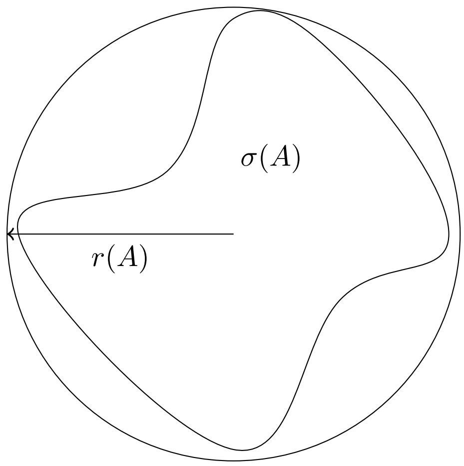


Figure 5.1: $\sigma(A) \subseteq \overline{B_{r(A)}(0)}$

5.12 Theorem

For $A \in L(H)$ holds:

$$r(A) = \limsup_{n \to \infty} |||A^n|||^{\frac{1}{n}}$$

If A is symmetric, then:

$$r(A) = |||A|||$$

Proof

Recall for a power series

$$\sum_{n=0}^{\infty} a_n z^n$$

with a_n,$z \in \mathbb{K}$ the root test (Wurzelkriterium):

- If

$$\limsup_{n \to \infty} |a_n z^n|^{\frac{1}{n}} =: c < 1$$

then $|a_n z^n| < c^n$ and therefore is

$$\sum_{n=0}^{\infty} c^n$$

a convergent dominating sequence. Thus $\sum_{n=0}^{\infty} a_n z^n$ converges as well.

– If

$$\limsup_{n\to\infty} |a_n z^n|^{\frac{1}{n}} =: c > 1$$

then $|a_n z^n| > c^n > 1$ for an infinite number of n. Therefore $a_n z^n$ does *not* converge to zero, which implies that $\sum_{n=0}^{\infty} a_n z^n$ does not converge as well.

– If

$$\limsup_{n\to\infty} |a_n z^n|^{\frac{1}{n}} = 1$$

no conclusion is possible.

$$\limsup_{n\to\infty} |a_n z^n|^{\frac{1}{n}} = |z| \cdot \limsup_{n\to\infty} |a_n|^{\frac{1}{n}}$$

The Radius of convergence (Konvergenzradius) is thus defined by:

$$R := \frac{1}{\limsup_{n\to\infty} |a_n|^{\frac{1}{n}}}$$

If $|z| < R$ the sum converges absolutely and if $|z| > R$ the sum diverges.

In our setting for $A = 0$ is nothing to prove. For $\lambda \in \varrho(A) \setminus \{0\}$ we make a formal expansion:

$$\mathcal{R}_\lambda = (\lambda - A)^{-1} = \frac{1}{\lambda}\left(\mathbb{1} - \frac{A}{\lambda}\right)^{-1} = \frac{1}{\lambda}\sum_{n=0}^{\infty} A^n \cdot \left(\frac{1}{\lambda}\right)^n$$

This is a power series in $\frac{1}{\lambda}$, but the coefficients are operators.

$$R := \frac{1}{\limsup_{n\to\infty} |||A^n|||^{\frac{1}{n}}}$$

For $\frac{1}{|\lambda|} < R$

$$\left|\left|\left|\sum_{n=0}^{\infty} A^n \left(\frac{1}{\lambda}\right)^n\right|\right|\right| \leq \sum_{n=0}^{\infty} |||A^n||| \frac{1}{\lambda^n}$$

converges absolutely and so

$$\sum_{n=0}^{\infty} A^n \left(\frac{1}{\lambda}\right)^n$$

converges in $L(H)$. Thus the resolvent

$$\mathcal{R}_\lambda = (\lambda - A)^{-1}$$

exists and $\sigma(A) \subseteq \overline{B_{\frac{1}{R}}(0)}$, i.e.:

$$r(A) \leq \frac{1}{R} = \limsup_{n\to\infty} |||A^n|||^{\frac{1}{n}}$$

If $\frac{1}{|\lambda|} > R$

$$\left|\left|\left|\sum_{n=0}^{\infty} A^n \left(\frac{1}{\lambda}\right)^n\right|\right|\right|$$

diverges.

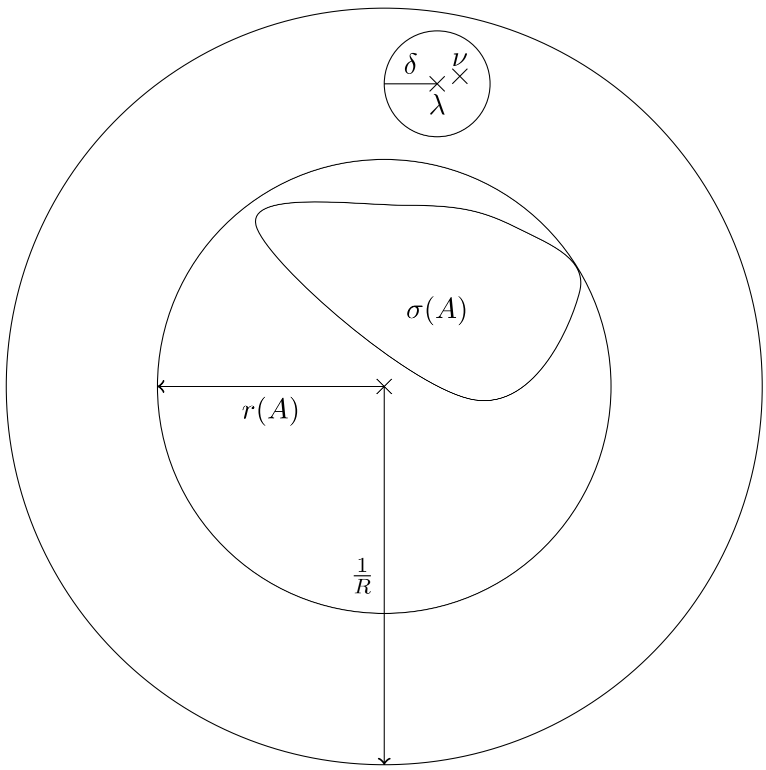


Figure 5.2: $\frac{1}{R} > r(A)$?

Why is r not smaller than $\frac{1}{R}$?
Assume that $r < \frac{1}{R}$ and choose λ with $r < |\lambda| < \frac{1}{R}$. Then $\mathcal{R}_\lambda$ exists and is analytic. Consider a $\nu \in B_\delta(\lambda)$.

$$\begin{aligned}
\mathcal{R}_\nu &= (\nu - A)^{-1} = ((\nu - \lambda) + (\lambda - A))^{-1} = \\
&= (((\nu - \lambda)\mathcal{R}_\lambda + \mathbb{1})(\lambda - A))^{-1} = \\
&= \mathcal{R}_\lambda (\mathbb{1} + (\nu - \lambda)\mathcal{R}_\lambda)^{-1} = \\
&= \mathcal{R}_\lambda \sum_{n=0}^{\infty} (-(\nu - \lambda))^n \mathcal{R}_\lambda^n
\end{aligned}$$

For $|\nu - \lambda| < \delta := \frac{1}{|||\mathcal{R}_\lambda|||}$ the Neumann series converges.
Thus $\mathcal{R}_\lambda$ can be expanded locally in a power series, i.e. $\mathcal{R}_\lambda$ is complex analytic or holomorphic.

Furthermore for $|\lambda| > \frac{1}{R}$ holds:

$$\mathcal{R}_\lambda = \sum_{n=0}^{\infty} A^n \frac{1}{\lambda^{n+1}}$$

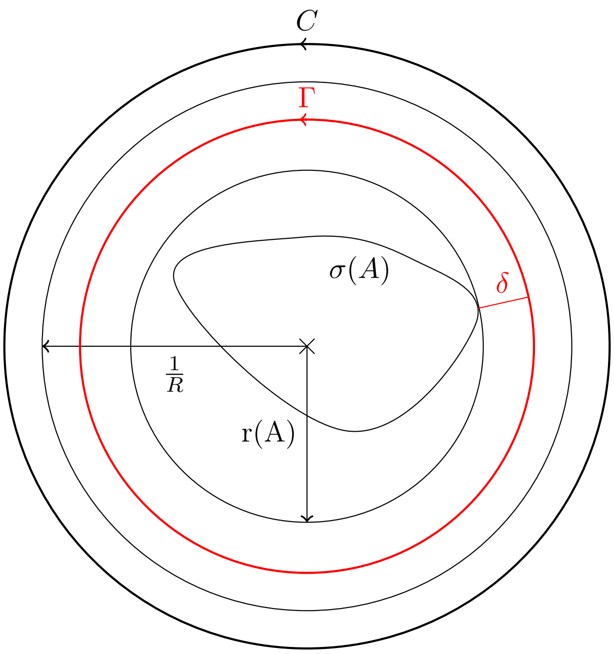


Figure 5.3: Contours Γ and C for integration

Integrate along the contour C:

$$\frac{1}{2\pi\mathbf{i}} \oint_C \lambda^n \mathcal{R}_\lambda \mathrm{d}\lambda = \sum_{k=0}^{\infty} A^k \underbrace{\frac{1}{2\pi\mathbf{i}} \oint_C \frac{\lambda^n}{\lambda^{k+1}} \mathrm{d}\lambda}_{=:I}$$

Since the geometric series converges absolutely, the summation and the integration can be interchanged. The residue theorem gives:

$$I = \begin{cases} 1 & \text{if } k = n \\ 0 & \text{otherwise} \end{cases}$$

Therefore we get:

$$\frac{1}{2\pi\mathbf{i}} \oint_C \lambda^n \mathcal{R}_\lambda \mathrm{d}\lambda = A^n$$

Choose $\Gamma = \partial B_{r+\delta}(0)$. We know, that $\mathcal{R}_\lambda$ is holomorphic outside Γ. Thus we may continuously deform the contour to obtain:

$$\frac{1}{2\pi\mathbf{i}} \oint_\Gamma \lambda^n \mathcal{R}_\lambda \mathrm{d}\lambda = A^n$$

Thus we have:

$$|||A^n||| = \left|\left|\left| \frac{1}{2\pi\mathbf{i}} \oint_\Gamma \lambda^n \mathcal{R}_\lambda \mathrm{d}\lambda \right|\right|\right| \leq C\,(r+\delta)^n\,(r+\delta)$$

$$C := \frac{1}{2\pi} \sup_{\lambda \in \Gamma} |||\mathcal{R}_\lambda|||$$

$$\Rightarrow \quad |||A^n|||^{\frac{1}{n}} \leq (r+\delta) \left(C^{\frac{1}{n}} (r+\delta)^{\frac{1}{n}} \right) \xrightarrow{n\to\infty} r+\delta$$

Therefore:

$$\limsup_{n\to\infty} |||A^n|||^{\frac{1}{n}} \leq r+\delta$$

Since δ is arbitrary, it follows that:

$$\frac{1}{R} = \limsup_{n\to\infty} |||A^n|||^{\frac{1}{n}} = r$$

We even conclude:

$$|||A^n|||^{\frac{1}{n}} \xrightarrow{n\to\infty} r(A)$$

Assume that A is *symmetric* (to show $|||A^n|||^{\frac{1}{n}} = |||A|||$). The Schwarz inequality gives:

$$\left|\left|\left|A^2\right|\right|\right| \leq |||A||| \cdot |||A||| = |||A|||^2$$

$$|||A|||^2 = \sup_{\|u\|=1} \langle Au, Au \rangle = \sup_{\|u\|=1} \left\langle u, Au^2 \right\rangle \leq \sup_{\|u\|=1} \underbrace{\|u\|}_{=1} \cdot \left\|A^2 u\right\|$$

Iteratively for $n \in \mathbb{N}$:

$$\left|\left|\left|A^{2^n}\right|\right|\right| = |||A|||^{2^n}$$

For arbitrary $m \in \mathbb{N}$ the Schwarz inequality gives:

$$|||A^m||| \leq |||A|||^m$$

Choose n such that $2^n > m$. THen:

$$\begin{aligned} & |||A|||^{2^n} = \left|\left|\left|A^{2^n}\right|\right|\right| = \left|\left|\left|A^m \cdot A^{2^n-m}\right|\right|\right| \leq |||A^m||| \cdot |||A|||^{2^n-m} \\ \Rightarrow \quad & |||A|||^m \leq |||A|||^m \end{aligned}$$

$\square_{5.12}$

5.13 Ritz method

Let $A \in L(H)$ be a symmetric compact operator on the separable Hilbert space H.
From the Hilbert-Schmidt theorem 5.10 we know that there exists an orthonormal eigenvalue basis (u_n) of H.

$$Au_n = \lambda_n u_n$$

We now want to construct the u_n:

Consider the "expectation value" functional:

$$\begin{aligned} S : H &\to \mathbb{R} \\ u &\mapsto \langle u, Au \rangle \end{aligned}$$

This is well defined, since:

$$\overline{S(u)} = \overline{\langle u, Au \rangle} = \langle Au, u \rangle = \langle u, Au \rangle = S(u)$$

S is bounded, because:

$$|S(u)| = |\langle u, Au \rangle| \leq |||A||| \cdot \|u\|^2 \overset{\|u\| \leq 1}{\leq} |||A|||$$

Maximize $|S(u)|$ on $\left\{ u \in H \,\middle|\, \|u\| = 1 \right\}$:
Choose a maximizing sequence (u_n) with $\|u_n\| = 1$ and:

$$|S(u_n)| \xrightarrow{n \to \infty} \sup_{\|u\|=1} |S(u)|$$

Since $\overline{B_1(0)}$ is weakly compact, there is a subsequence u_{k_l}, which converges weakly $u_{k_l} \rightharpoonup u$.
Since A is compact, the sequence

$$v_{k_l} := Au_{k_l} \to v$$

converges and $Au = v$. As a consequence:

$$S(u_{k_l}) = \langle u_{k_l}, Au_{k_l} \rangle = \langle u_{k_l}, v_{k_l} \rangle = \underbrace{\langle u_{k_l}, v \rangle}_{\to \langle u, v \rangle} + \langle u_{k_l}, v_{k_l} - v \rangle \xrightarrow{l \to \infty} \langle u, v \rangle = \langle u, Au \rangle = S(u)$$

This follows, because:

$$|\langle u_{k_l}, v_{k_l} - v \rangle| \leq \underbrace{\|u_{k_l}\|}_{=1} \cdot \underbrace{\|v_{k_l} - v\|}_{\to 0} \xrightarrow{l \to \infty} 0$$

Thus S is weakly continuous, i.e. for any $u_k \rightharpoonup u$ converges $S(u_k) \to S(u)$.
Because (u_n) is a maximizing sequence, we get:

$$|S(u)| = \sup_{\|\tilde{u}\|=1} |S(\tilde{u})|$$

Therefore u is the desired maximizer.

- u is on the unit sphere:
  The simple approach

  $$\|u\|^2 \neq \lim_{l \to \infty} \|u_{k_l}\|^2$$

  does not work, because u_{k_l} only converges weakly.
  Example:
  If (e_l) is an orthonormal Hilbert basis in a separable Hilbert space, then $e_l \rightharpoonup 0$, but:

  $$\lim_{l \to \infty} \|e_l\| = 1 \neq 0 = \|0\|$$

But it holds:

$$\begin{aligned}\|u\|^2 &= \lim_{l\to\infty} |\langle u,u_{k_l}\rangle| \le \lim_{l\to\infty} \|u_{k_l}\| \cdot \|u\| = \|u\| \\ \Rightarrow\quad \|u\| &\le 1\end{aligned}$$

Assume $\|u\| < 1$, then the vector $\hat{u} := \frac{u}{\|u\|}$ would satisfy the equation:

$$|S(\hat{u})| = |\langle \hat{u},A\hat{u}\rangle| = \frac{1}{\|u\|^2} |\langle u,Au\rangle| = \frac{1}{\|u\|^2} \sup_{\|v\|=1} |S(v)| \overset{\|u\|<1}{>} \sup_{\|v\|=1} |S(v)|$$

This is a contradiction. Therefore u is in fact a unit vector.

– u is an eigenvector corresponding to the eigenvalue $\lambda = \langle u,Au\rangle \in \mathbb{R}$: Consider the variation for $v \in H$:

$$\tilde{u}(\tau) = u + \tau v$$

$$S\left(\frac{\tilde{u}}{\|\tilde{u}\|}\right) = \frac{\langle \tilde{u},A\tilde{u}\rangle}{\langle \tilde{u},\tilde{u}\rangle} = \frac{\langle u+\tau v,A(u+\tau v)\rangle}{\langle u+\tau v,u+\tau v\rangle}$$

This is called *Rayleigh quotient*. We know that $S(\tilde{u}(\tau))$ is extremal at $\tau = 0$:

$$\begin{aligned}0 &= \frac{\mathrm{d}}{\mathrm{d}\tau} S(\tilde{u}(\tau))\bigg|_{\tau=0} = \\ &= \frac{\langle u,Av\rangle + \langle v,Au\rangle + 2\tau\langle v,v\rangle}{\langle u+\tau v,u+\tau v\rangle} - \frac{\langle u+\tau v,A(u+\tau v)\rangle}{\langle u+\tau v,u+\tau v\rangle^2} \cdot (\langle v,u\rangle + \langle u,v\rangle + \tau\langle v,v\rangle)\bigg|_{\tau=0} = \\ &\overset{A\text{ symmetric}}{=} 2\frac{\mathrm{Re}(\langle v,Au\rangle)}{\langle u,u\rangle} - 2\mathrm{Re}(\langle v,u\rangle)\frac{\langle u,Au\rangle}{\langle u,u\rangle^2} = \\ &\underset{\langle u,u\rangle=1}{\overset{\lambda=\langle u,Au\rangle}{=}} 2(\mathrm{Re}(\langle v,Au\rangle) - \lambda\mathrm{Re}(\langle v,u\rangle)) = 2\mathrm{Re}(\langle v,(A-\lambda)u\rangle)\end{aligned}$$

Set $v = e^{\mathbf{i}\varphi}w$ for any $\varphi \in \mathbb{R}$ and $w \in H$. So:

$$0 = \mathrm{Re}(\langle v,(A-\lambda)u\rangle) = \mathrm{Re}\left(e^{-\mathbf{i}\varphi}\langle w,(A-\lambda)u\rangle\right) \qquad \underset{\varphi\in\mathbb{R}}{\forall}$$

$$\begin{aligned}\Rightarrow\quad \langle w,(A-\lambda)u\rangle &= 0 \qquad \underset{w\in H}{\forall} \\ (A-\lambda)u &= 0 \\ Au &= \lambda u\end{aligned}$$

– It holds $|\lambda| = |||A|||$:
There is no point ν in the spectrum of A with $|\nu| > |\lambda|$, because otherwise for all $v \in H$ with $Av = \nu v$ follows:

$$\frac{|\langle v,Av\rangle|}{\langle v,v\rangle} = |\nu| > |\lambda| = |\langle u,Au\rangle| = \sup_{w\in H} \frac{|\langle w,Aw\rangle|}{\langle w,w\rangle}$$

This is a contradiction. Thus we get:

$$|\lambda| = \sup_{\nu\in\sigma(A)} |\nu| \overset{\text{by definition}}{=} r(A) \overset{5.12}{=} |||A|||$$

Thus we have *constructed* a $u \in H$ with $\|u\| = 1$, $Au = \lambda u$ and $|\lambda| = |||A|||$. Now one can proceed inductively:

$$H_1 := \langle u \rangle^{\perp}$$

$$A\big|_{H_1} : H_1 \to H_1$$

(We saw that H_1 is invariant under A.)
Repeat the above procedure to maximize $|\langle v,Av\rangle|$ on $H_1 \cap \left\{ v \in H \big| \, \|v\| = 1 \right\}$. This gives u_1 with $\|u\|_1 = 1$, $Au_1 = \lambda_1 u_1$ and:

$$|\lambda_1| = \left|\left|\left| A\big|_{H_1} \right|\right|\right| \leq \left|\left|\left| A\big|_{H} \right|\right|\right| = |\lambda|$$

Now set $H_2 = \langle u,u_1 \rangle^{\perp}$ and proceed inductively.
This gives a sequence $u_0 := u,\ u_1,\ u_2,\ \ldots$ of orthonormal eigenvectors, i.e. $Au_j = \lambda_j u_j$, with decreasing eigenvalues$|\lambda_j|$.
These (u_j) are an orthonormal basis. (Proof as in Theorem 5.10) $\square_{5.13}$

Ritz, Galerkin: Finite element method
Example: Helium molecule wave function in $H = L^2\left(\mathbb{R}^3,\mathbb{C}\right)$

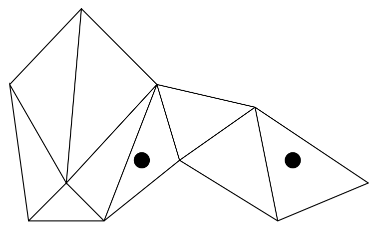


Figure 5.4: finite lattice for numerical approximation

$$A = -\frac{\hbar^2}{2m}\Delta - \frac{ze^2}{\|x - x_1\|} - \frac{ze^2}{\|x - x_2\|}$$

Now minimize

$$\frac{\langle u,Au \rangle}{\langle u,u \rangle}$$

on a finite subspace of H.

6 A few (technical) results

6.1 Dini's theorem

Let E be a metric space and $f_n : E \to \mathbb{R}$ a sequence of real valued functions.

6.1.1 Definition (point-wise/uniform convergence)

f_n *converges point-wise* to f if $f_n(x) \to f(x)$ converges for all $x \in E$, i.e.:

$$\underset{x\in E}{\forall}\ \underset{\varepsilon\in\mathbb{R}_{>0}}{\forall}\ \underset{N(\varepsilon,x)}{\exists}\ \underset{n\in\mathbb{N}_{\geq N}}{\forall} : |f_n(x) - f(x)| < \varepsilon$$

f_n *converges uniformly* to f, in symbols $f_n \rightrightarrows f$, if for all $\varepsilon \in \mathbb{R}_{>0}$ exists a $N(\varepsilon)$ such that for all $n \geq N$ and all $x \in E$ holds:

$$|f_n(x) - f(x)| < \varepsilon$$

With quantifiers this is:

$$\underset{\varepsilon\in\mathbb{R}_{>0}}{\forall}\ \underset{N(\varepsilon)}{\exists}\ \underset{n\in\mathbb{N}_{\geq N}}{\forall}\ \underset{x\in E}{\forall} : |f_n(x) - f(x)| < \varepsilon$$

6.1.2 Theorem

If (f_n) is a sequence of continuous functions with $f_n \rightrightarrows f$, then f is also continuous.
This is not true in general for point wise convergence:

$$f_n(x) = \begin{cases} 1 & \text{for } 0 \leq x \leq \frac{1}{2}\left(1 - \frac{1}{n}\right) \\ 0 & \text{for } x \geq \frac{1}{2} \\ n(1-2x) & \text{for } \frac{1}{2}\left(1 - \frac{1}{n}\right) < x < \frac{1}{2} \end{cases}$$

$f_n \to f$ converges pointwise to:

$$f(x) = \begin{cases} 1 & x < \frac{1}{2} \\ 0 & x \geq \frac{1}{2} \end{cases}$$

This f is *not* continuous.

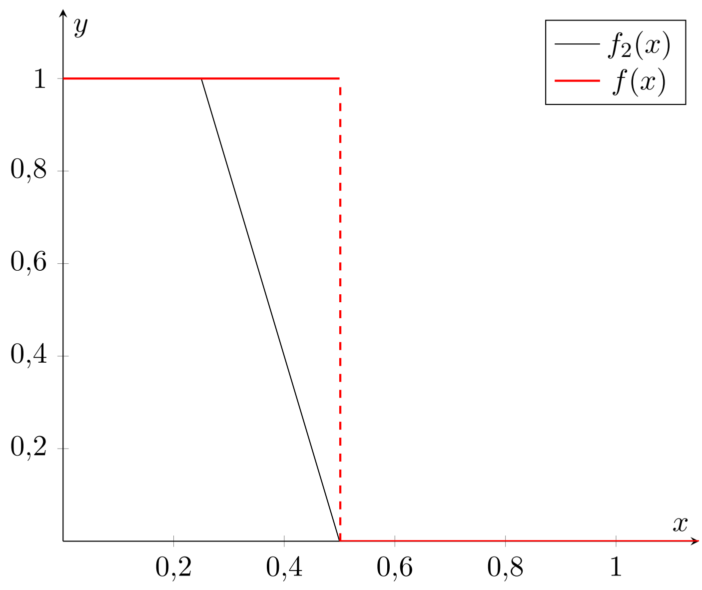


Figure 6.1: $f_n(x)$ is continuous, but not $f(x)$

Proof

Show that for all $x \in E$ the ε-δ-criterion is satisfied:
Since $f_n \rightrightarrows f$ converges uniformly, there is a $N \in \mathbb{N}$ such that for all $n \in \mathbb{N}_{\geq N}$ and all $x \in E$ holds:

$$|f_n(x) - f(x)| < \frac{\varepsilon}{3}$$

Because the f_n are continuous, there exists a $\delta \in \mathbb{R}_{>0}$ such that for all $y \in B_\delta(x)$ holds:

$$|f_N(x) - f_N(y)| < \frac{\varepsilon}{3}$$

Now follows for all $y \in B_\delta(x)$:

$$|f(y) - f(x)| \leq \underbrace{|f(y) - f_N(y)|}_{<\frac{\varepsilon}{3}} + \underbrace{|f_N(y) - f_N(x)|}_{<\frac{\varepsilon}{3}} + \underbrace{|f_N(x) - f(x)|}_{<\frac{\varepsilon}{3}} < \varepsilon$$

Therefore f is continuous. $\square_{6.1.2}$

6.1.3 Definition (monotonically increasing/decreasing)

The sequence of functions (f_n), $f_n : E \to \mathbb{R}$ is called *monotonically increasing (decreasing)* if for all $x \in E$ the real sequence $f_n(x)$ is monotonically increasing (decreasing).

6.1.4 Theorem (Dini)

Let E be a *compact* metric space, (f_n) monotone and $f_n \to f$.
If f_n and f are continuous, then the convergence $f_n \rightrightarrows f$ is uniform.

Proof

Without loss of generality we assume (f_n) is a monotonically increasing sequence (otherwise consider $-f_n$), i.e. $f_n(x) \leq f_{n+1}(x)$ for all $x \in E$ and all $n \in \mathbb{N}$.
Given $\varepsilon > 0$ we want to show:

$$\underset{N\in\mathbb{N}}{\exists}\ \underset{x\in E}{\forall}\ \underset{\in\mathbb{N}_{\geq N}}{\forall} : |f(x) - f_n(x)| < \varepsilon$$

For any $x \in E$ there exists an $N(x)$ such that $|f_n(x) - f(x)| < \frac{\varepsilon}{2}$ for all $n \in \mathbb{N}_{\geq N}$ (point-wise convergence). Since both $f_{N(x)}$ and f are continuous functions, there exists a neighborhood $U(x) = B_{\delta(x)}(x)$ of x such that for all $z \in U(x)$ holds:

$$\left|f_{N(x)}(z) - f_{N(x)}(x)\right| \leq \frac{\varepsilon}{4}$$
$$|f(z) - f(x)| \leq \frac{\varepsilon}{4}$$

Then follows:

$$\left|f_{N(x)}(z) - f(z)\right| \leq \underbrace{\left|f_{N(x)}(z) - f_{N(x)}(x)\right|}_{\leq\frac{\varepsilon}{4}} + \underbrace{\left|f_{N(x)}(x) - f(x)\right|}_{<\frac{\varepsilon}{2}} + \underbrace{|f(x) - f(z)|}_{\leq\frac{\varepsilon}{4}} < \varepsilon$$

Since $f_n(z)$ is monotonically increasing, it follows that $|f_n(z) - f(z)| < \varepsilon$ for all $z \in B_{\delta(x)}(x)$.
Now use a standard compactness argument: Since E is compact, it can be covered by a finite number of these balls $B_{\delta(x_1)}(x_1), \ldots, B_{\delta(x_n)}(x_n)$. Define:

$$N = \max\{N(x_1), \ldots, N(x_n)\}$$

So for all $n \in \mathbb{N}_{\geq N}$ holds:

$$|f_n(x) - f(x)| < \varepsilon$$

$\square_{6.1.4}$

6.2 Stone-Weierstraß theorem

We follow the nice (since constructive) proof by Bernstein.

6.2.1 Definition (polynomials)

Let $E = C^0([0,1])$ be the Banach space of real valued functions with norm:

$$\|f\| = \sup_{x\in[0,1]} |f(x)|$$

$\mathcal{P}([0,1])$ are the *real polynomials*, i.e. for $f \in \mathcal{P}([0,1])$ there are $a_j \in \mathbb{R}$ such that:

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0$$

Clearly $\mathcal{P}([0,1]) \subseteq C^0([0,1])$ forms a subspace.

We want to show that $\mathcal{P}([0,1])$ is dense in $C^0([0,1])$.

6.2.2 Lemma

For $x \in [0,1]$ holds:

$$\sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = 1$$

Proof

$$\sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = (x+1-x)^n = 1$$

$\square_{6.2.2}$

6.2.3 Lemma

For $x \in [0,1]$ holds:

$$\sum_{k=0}^{n} (nx-k)^2 \binom{n}{k} x^k (1-x)^{n-k} = nx(1-x) \leq \frac{n}{4}$$

Obviously holds

$$(nx-k)^2 \leq 4n^2$$

and therefore:

$$\sum_{k=0}^{n} (nx-k)^2 \binom{n}{k} x^k (1-x)^{n-k} \leq 4n^2 \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = 4n^2$$

Proof

It holds:

$$\begin{aligned}
\sum_{k=0}^{n} k \binom{n}{k} x^k (1-x)^{n-k} &= \sum_{k=0}^{n} k \frac{n!}{k!\,(n-k)!} x^k (1-x)^{n-k} = \\
&= 0 + \sum_{k=1}^{n} \frac{n \cdot (n-1)!}{(k-1)!\,(n-k)!} x^k (1-x)^{n-k} = \\
&= n \sum_{k=1}^{n} \binom{n-1}{k-1} x^k (1-x)^{n-k} = \\
&\overset{j:=k-1}{=} n \sum_{j=0}^{n-1} \binom{n-1}{j} x^{j+1} (1-x)^{n-j-1} = \\
&= nx \sum_{j=0}^{n-1} \binom{n-1}{j} x^j (1-x)^{(n-1)-j} = nx(x+1-x)^{n-1} = nx
\end{aligned}$$

Similarly one gets:

$$\sum_{k=0}^{n} k\,(k-1)\begin{pmatrix} n \\ k \end{pmatrix} x^k\,(1-x)^{n-k} = n\,(n-1)\sum_{k=2}\begin{pmatrix} n-2 \\ k-2 \end{pmatrix} x^k\,(1-x)^{n-k} = n\,(n-1)\,x^2$$

Together this gives:

$$\begin{aligned}\sum_{k=0}^{n}(nx-k)^2\begin{pmatrix} n \\ k \end{pmatrix} x^k\,(1-x)^{n-k} &= \sum_{k=0}^{n}\left(n^2x^2-2nxk+k^2\right)\begin{pmatrix} n \\ k \end{pmatrix} x^k\,(1-x)^{n-k} = \\ &= \sum_{k=0}^{n}\left(n^2x^2-2nxk+k\,(k-1)+k\right)\begin{pmatrix} n \\ k \end{pmatrix} x^k\,(1-x)^{n-k} = \\ &= n^2x^2-2nx\cdot nx+n\,(n-1)\,x^2+nx = \\ &= -n^2x^2+n^2x^2-nx^2+nx = nx\,(1-x)\end{aligned}$$

$\square_{6.2.3}$

A more elegant method is to use derivatives:

$$\begin{aligned}\sum_{k=0}^{n}\begin{pmatrix} n \\ k \end{pmatrix} x^k y^{n-k} &= (x+y)^n \\ \sum_{k=0}^{n} k\begin{pmatrix} n \\ k \end{pmatrix} x^k y^{n-k} &= x\cdot\frac{\mathrm{d}}{\mathrm{d}x}\left(\sum_{k=0}^{n}\begin{pmatrix} n \\ k \end{pmatrix} x^k y^{n-k}\right) \\ \sum_{k=0}^{n} k^2\begin{pmatrix} n \\ k \end{pmatrix} x^k y^{n-k} &= \left(x\cdot\frac{\mathrm{d}}{\mathrm{d}x}\right)^2\left(\sum_{k=0}^{n}\begin{pmatrix} n \\ k \end{pmatrix} x^k y^{n-k}\right)\end{aligned}$$

6.2.4 Definition

For $f\in C^0\left([0{,}1]\right)$ define:

$$B_n f\,(x) := \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\begin{pmatrix} n \\ k \end{pmatrix} x^k\,(1-x)^{n-k}$$

6.2.5 Theorem (Bernstein)

For any $f\in C^0\left([0{,}1]\,,\mathbb{R}\right)$, $B_n f\rightrightarrows f$ converges uniformly.

Example: $f\,(x)=10x\cdot e^{-3x}+\frac{1}{5}\cos\left((4x)^2\right)$

$$\begin{aligned}B_4 f\,(x) &\approx 0{,}2\cdot(1-x)^4+5{,}2\cdot x\cdot(1-x)^3+5{,}9\cdot x^2\cdot(1-x)^2+2{,}4\cdot x^3\cdot(1-x)+0{,}3\cdot x^4 \\ B_{10} f\,(x) &\approx 0{,}2\cdot(1-x)^{10}+9{,}4\cdot x\cdot(1-x)^9+56{,}6\cdot x^2\cdot(1-x)^8+ \\ &\quad +149{,}5\cdot x^3\cdot(1-x)^7+217{,}9\cdot x^4\cdot(1-x)^6+248{,}2\cdot x^5\cdot(1-x)^5+ \\ &\quad +244{,}7\cdot x^6\cdot(1-x)^4+103{,}2\cdot x^7\cdot(1-x)^3+26{,}5\cdot x^8\cdot(1-x)^2+ \\ &\quad +7{,}9\cdot x^9\cdot(1-x)+0{,}3\cdot x^{10}\end{aligned}$$

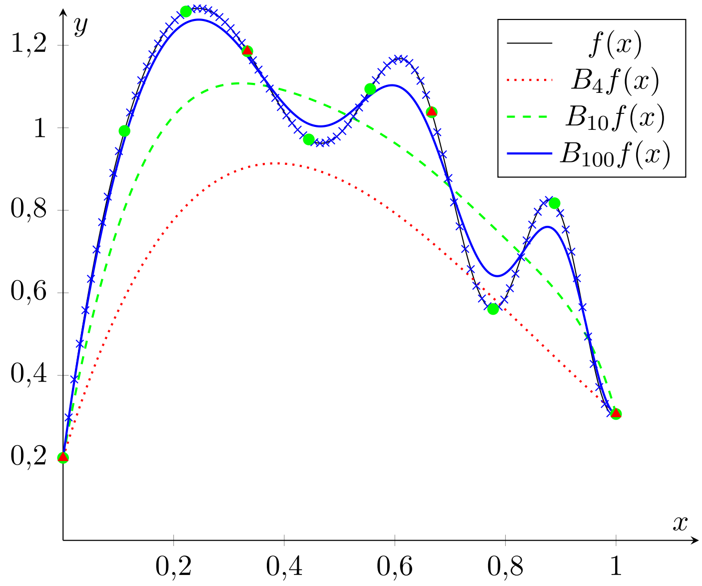


Figure 6.2: Approximation of $f(x)$ by $B_nf(x)$

Proof

Without loss of generality assume $f \neq 0$ (otherwise $B_nf = 0 = f$).

$$M := \|f\| > 0$$

Consider an arbitrary $\varepsilon \in \mathbb{R}_{>0}$. f is continuous on the compact set [0,1] and thus uniformly continuous, i.e. there exists a $\delta \in \mathbb{R}_{>0}$ such that:

$$|x-y| < \delta \quad \Rightarrow \quad |f(x) - f(y)| < \frac{\varepsilon}{2}$$

Choose $\mathbb{N} \ni N \geq \frac{M}{\varepsilon\delta^2}$.

Claim: $|B_nf(x) - f(x)| < \varepsilon$ for all $x \in [0,1]$ and all $n \geq N$.

Proof: It holds:

$$f(x) = \sum_{k=0}^{n} f(x) \begin{pmatrix} n \\ k \end{pmatrix} x^k (1-x)^{n-k}$$
$$B_nf(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \begin{pmatrix} n \\ k \end{pmatrix} x^k (1-x)^{n-k}$$

$$(B_nf - f)(x) = \sum_{k=0}^{\infty} \left(f\left(\frac{k}{n}\right) - f(x)\right) \begin{pmatrix} n \\ k \end{pmatrix} x^k (1-x)^{n-k}$$

Define:

$$A := \left\{k \,\middle|\, \left|\frac{k}{n} - x\right| < \delta\right\} \qquad B := \left\{k \,\middle|\, \left|\frac{k}{n} - x\right| \geq \delta\right\}$$

We have:

$$\sum_{k\in A}\underbrace{\left|f\left(\frac{k}{n}\right)-f\left(x\right)\right|}_{<\frac{\varepsilon}{2}}\left(\begin{array}{c}n\\k\end{array}\right)x^{k}\left(1-x\right)^{n-k}<\frac{\varepsilon}{2}\sum_{k\in A}\left(\begin{array}{c}n\\k\end{array}\right)x^{k}\left(1-x\right)^{n-k}\leq\frac{\varepsilon}{2}$$

$$\begin{aligned}&\sum_{k\in B}\underbrace{\left|f\left(\frac{k}{n}\right)-f\left(x\right)\right|}_{\leq2\|f\|=2M}\left(\begin{array}{c}n\\k\end{array}\right)x^{k}\left(1-x\right)^{n-k}\leq\\&\leq 2M\sum_{k\in B}\left(\begin{array}{c}n\\k\end{array}\right)x^{k}\left(1-x\right)^{n-k}\leq\\&\underset{\Rightarrow n^{2}\delta^{2}\leq(k-nx)^{2}}{\overset{k\in B}{\leq}}\frac{2M}{n^{2}\delta^{2}}\sum_{k=0}^{n}\underbrace{\left(k-nx\right)^{2}\left(\begin{array}{c}n\\k\end{array}\right)x^{k}\left(1-x\right)^{n-k}}_{\leq\frac{n}{4}}\leq\\&\overset{n\geq N}{\leq}\frac{M}{2n\delta^{2}}\leq\frac{M}{2\frac{M}{\varepsilon\delta^{2}}\delta^{2}}=\frac{\varepsilon}{2}\end{aligned}$$

Therefore holds for all $x\in[0,1]$.

$$\left|B_{n}f\left(x\right)-f\left(x\right)\right|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon$$

$\square_{\text{Claim}}$

Therefore $B_{n}f\rightrightarrows f$ converges uniformly. $\square_{6.2.5}$

Now generalize: Let E be a compact metric space. $C^{0}\left(E,\mathbb{R}\right)$ with

$$\|f\|=\sup_{x\in E}\left|f\left(x\right)\right|$$

is a Banach space. Moreover, it is an algebra with the point-wise multiplication:

$$\left(f\cdot g\right)\left(x\right):=f\left(x\right)\cdot g\left(x\right)$$

The multiplication is continuous:

$$\|f\cdot g\|\leq\|f\|\cdot\|g\|$$

In summary $\left(C^{0}\left(E,\mathbb{R}\right),\|.\|,+,\cdot\right)$ is a *Banach algebra.*

6.2.6 Theorem (Weierstraß)

The polynomials are dense in $C^{0}\left([0,1],\mathbb{R}\right)$.

Proof

For any $f\in C^{0}\left([0,1],\mathbb{R}\right)$, $B_{n}f\rightrightarrows f$ converges uniformly and since the $B_{n}f$ are polynomials, these are dense. $\square_{6.2.6}$

6.2.7 Theorem (Stone-Weierstraß)

Let $\mathcal{A} \subseteq C^0(E,\mathbb{R})$ be a subalgebra with the following properties:

1. $\mathcal{A}$ contains $f = 1$ and so by scalar multiplication all the constant functions.
2. $\mathcal{A}$ separates the points of E, i.e. for all $x,y \in E$ with $x \neq y$ there exists a $f \in \mathcal{A}$ such that $f(x) \neq f(y)$.

Then $\mathcal{A}$ is dense in $C^0(E,\mathbb{R})$.

Proof

i) There is a sequence of polynomials u_n on $[0,1]$ such that $u_n \rightrightarrows f$ with $f(t) = \sqrt{t}$. This follows immediately from theorem 6.2.6.

ii) If $f \in \mathcal{A}$, then $|f|$ defined by $|f|(x) := |f(x)|$ is in the closure $\overline{\mathcal{A}}$ of $\mathcal{A}$:
For $f \in \mathcal{A}$ define:

$$a := \|f\| = \max_{x \in E} |f(x)|$$

$$\Rightarrow \qquad \frac{f^2(x)}{a^2} \in [0,1]$$

Then converges:

$$u_n\left(\frac{f^2(x)}{a^2}\right) \xrightarrow{n\to\infty} \sqrt{\frac{f^2(x)}{a^2}} = \frac{|f(x)|}{a}$$

The functions $u_n\left(\frac{f^2}{a^2}\right)$ lie in $\mathcal{A}$, since these are a polynomials of f and thus again elements of the algebra $\mathcal{A}$. Moreover $u_n\left(\frac{f^2}{a^2}\right)$ converges uniformly to $\frac{|f|}{a}$, because for a given $\varepsilon \in \mathbb{R}_{>0}$ exists a $N \in \mathbb{N}$ such that for all $n \in \mathbb{N}_{\geq N}$ and all $t \in [0,1]$ holds:

$$\left|u_n(t) - \sqrt{t}\right| < \varepsilon$$

Then follows with $t = \frac{f^2(x)}{a^2}$:

$$\left|u_n\left(\frac{f^2(x)}{a^2}\right) - \frac{|f|}{a}\right| < \varepsilon$$

Thus $\frac{|f|}{a} \in \overline{\mathcal{A}}$ and therefore also $|f| \in \overline{\mathcal{A}}$.

iii) For $f,g \in \overline{\mathcal{A}}$ also $\min(f,g)$ and $\max(f,g)$ (defined point-wise) are again in $\overline{\mathcal{A}}$:

$$\min(f,g) = \frac{1}{2}(f + g - |f - g|)$$
$$\max(f,g) = \frac{1}{2}(f + g + |f - g|)$$

Choose $f_n,g_n \in \mathcal{A}$ such that $f_n \rightrightarrows f$ and $g_n \rightrightarrows g$. By ii) follows $|f_n - g_n| \in \overline{\mathcal{A}}$ and $|f_n - g_n| \rightrightarrows |f - g|$. Therefore holds:

$$\overline{\mathcal{A}} \ni \min(f_n,g_n) \rightrightarrows \min(f,g) \in \overline{\mathcal{A}}$$

Similarly the claim follows for max.

iv) For all $x,y \in E$ with $x \neq y$ and $\alpha,\beta \in \mathbb{R}$ exists a $f \in \mathcal{A}$ such that $f(x) = \alpha$ and $f(y) = \beta$:
For $\alpha = \beta$ we choose $f = \alpha$ as constant function.
For $\alpha \neq \beta$ there exists, since $\mathcal{A}$ separates points of E, a $g \in \mathcal{A}$ with $g(x) \neq g(y)$. Set $f = c_0 + c_1 g$ and choose:

$$\begin{aligned}
& \alpha = c_0 + c_1 g(x) \\
& \beta = c_0 + c_1 g(y) \\
\Rightarrow \quad & c_1 = \frac{\alpha - \beta}{g(x) - g(y)} \\
\Rightarrow \quad & c_0 = \alpha - \frac{\alpha - \beta}{g(x) - g(y)} g(x) = \frac{\alpha g(x) - \alpha g(y) - \alpha g(x) + \beta g(x)}{g(x) - g(y)} = \\
& = \frac{\beta g(x) - \alpha g(y)}{g(x) - g(y)}
\end{aligned}$$

v) For all $f \in C^0$, $x \in E$ and $\varepsilon \in \mathbb{R}_{>0}$ there is a $g \in \overline{\mathcal{A}}$ such that

$$g(x) = f(x)$$

and for all $y \in \overline{\mathcal{A}}$ holds:

$$g(y) \leq f(y) + \varepsilon$$

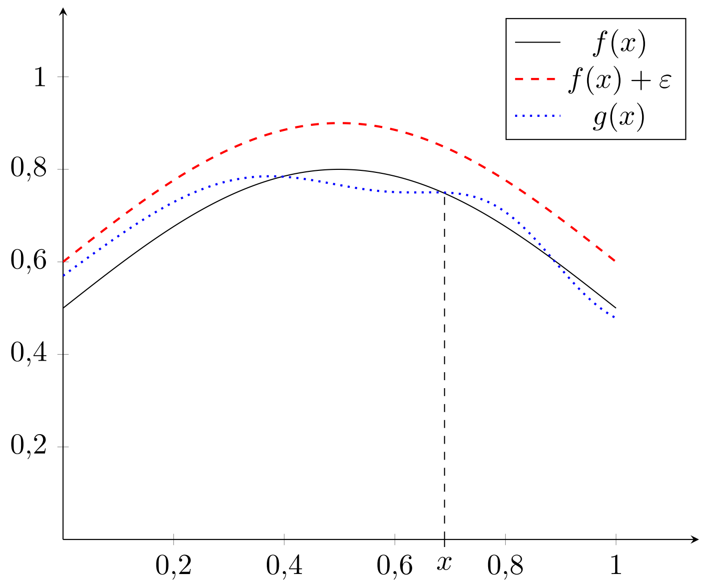


Figure 6.3: $g(x) \leq f(x) + \varepsilon$

To show this, choose for any $z \in E$ a $h_z \in \overline{\mathcal{A}}$ with $h_z(x) = f(x)$ and $h_z(z) \leq f(z) + \frac{\varepsilon}{2}$, which is possible after iv).
Since h_z is continuous, there is a neighborhood U_z of z such that $h_z \leq f + \varepsilon$ on U_z.

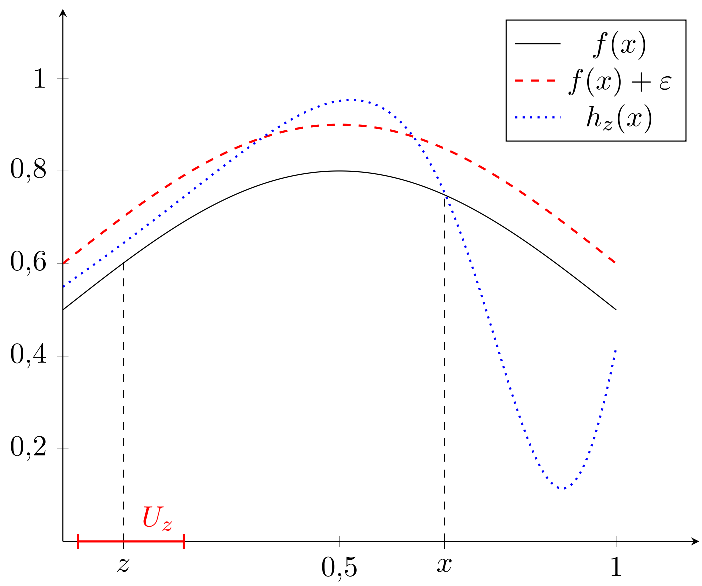


Figure 6.4: $h_z \leq f + \varepsilon$ on U_z

Since E is compact, we can cover it by a finite number of such neighborhoods $U_{z_1}, \ldots, U_{z_N}$. Define:

$$g := \min \{h_{z_1}, \ldots, h_{z_N}\} \in \overline{\mathcal{A}}$$

It holds $g(x) = f(x)$, because $h_{z_i}(x) = f(x)$. We also know:

$$g\big|_{U_j} \leq h_{z_j}\big|_{U_j} \leq f + \varepsilon$$

vi) $\overline{\mathcal{A}} = C^0$: Denote the function g constructed in step v) by g_x.

$$\begin{aligned} g_x(x) &= f(x) \\ g_x &\leq f + \varepsilon \end{aligned}$$

By continuity of g_x there exists a neighborhood U_x of x such that $g_x \geq f - \varepsilon$ on U_x. By compactness we can cover E by a finite number of such neighborhoods $U_{x_1}, \ldots, U_{x_k}$ and define:

$$g := \max \{g_{x_1}, \ldots, g_{x_k}\}$$

Then follows:

$$\begin{aligned} f - \varepsilon \leq & g \leq f + \varepsilon \\ \|f - g\| & < \varepsilon \end{aligned}$$

$\square_{6.2.7}$

Counterexample in the complex case:

$$E = [0,1] \times [0,1] \subseteq \mathbb{C}$$

Consider the set $\mathcal{A} = \mathcal{P}(z)$ of polynomials in z.

- The constant functions are in $\mathcal{A}$.
- $\mathcal{A}$ separates points:
  If $z_1 \neq z_2$ take $f(z) = z$ then $f(z_1) \neq f(z_2)$.

$$\overline{\mathcal{A}} = ?$$

By Morera's theorem we get:

$$\overline{\mathcal{A}} = \left\{ f \in C^0\left([0,1]^2\right) \middle| f|_{(0,1)^2} \text{ is holomorphic} \right\} \neq C^0\left([0,1]^2\right)$$

For example $f(x + \mathbf{i}y) = x - \mathbf{i}y$. We have $f \in C^0\left([0,1]^2\right)$, but $f \notin \overline{\mathcal{A}}$.

6.2.8 Theorem (Stone-Weierstraß, complex version)

Let $\mathcal{A} \subseteq C^0(E,\mathbb{C})$ be a subalgebra with the properties 1. and 2. from theorem 6.2.7 and additionally:

3. $f \in \mathcal{A} \Rightarrow \overline{f} \in \mathcal{A}$

Then $\mathcal{A}$ is dense in $C^0(E,\mathbb{C})$.

Proof

Consider the algebras:

$$\begin{aligned} \operatorname{Re}(\mathcal{A}) &= \left\{ f + \overline{f} \middle| f \in \mathcal{A} \right\} \subseteq \mathcal{A} \\ \operatorname{Im}(\mathcal{A}) &= \left\{ \frac{1}{\mathbf{i}}\left(f - \overline{f}\right) \middle| f \in \mathcal{A} \right\} \subseteq \mathcal{A} \end{aligned}$$

These are subalgebras of $C^0(E,\mathbb{R})$. By the real Stone-Weierstraß theorem we get:

$$\overline{\operatorname{Re}(\mathcal{A})} = \overline{\operatorname{Im}(\mathcal{A})} = C^0(E,\mathbb{R})$$

For given $f \in C^0(E,\mathbb{C})$ approximate $\operatorname{Re}(f)$ and $\operatorname{Im}(f)$. $\square_{6.2.8}$

6.3 Arzelà-Ascoli theorem

Let K be a compact metric space and E a Banach space.

$C^0(K,E)$ is the Banach space of continuous functions $f : K \to E$ with norm:

$$\|f\| := \sup_{x \in K} \|f(x)\|_E$$

Let $\mathcal{F} \subseteq C^0(K,E)$ be a subset. Is $\mathcal{F}$ compact?

6.3.1 Definition (relatively compact)

A subset A of a metric space is called *relatively compact*, if $\overline{A}$ is compact.

6.3.2 Definition (equicontinuous)

A family $\mathcal{F} \subseteq C^0(K,E)$ is called *equicontinuous* (gleichgradig stetig) if for all $x \in K$ and all $\varepsilon \in \mathbb{R}_{>0}$ there exists a $\delta \in \mathbb{R}_{>0}$ such that for all $y \in B_\delta(x)$ and for all $f \in \mathcal{F}$ holds:

$$\|f(x) - f(y)\| < \varepsilon$$

(Thus δ is independent of $f \in \mathcal{F}$.)

6.3.3 Theorem (Arzelà-Ascoli)

$\mathcal{F} \subseteq C^0(K,E)$ is relatively compact if and only if the following two conditions holds:

i) $\mathcal{F}$ is equicontinuous.

ii) For every $x \in K$ the set

$$\mathcal{F}(x) := \left\{ f(x) \,\middle|\, f \in \mathcal{F} \right\}$$

is relatively compact in E.

Proof

„$\Rightarrow$“: Assume that $\mathcal{F} \subseteq C^0(K,E)$ is relatively compact.

i) Assume that $\mathcal{F}$ is *not* equicontinuous. Then there exists an $\varepsilon \in \mathbb{R}_{>0}$ and sequences $x_n \in K$, $f_n \in \mathcal{F}$ and $y_n \in B_{\frac{1}{n}}(x_n)$ such that:

$$\|f_n(x_n) - f_n(y_n)\| \geq \varepsilon$$

After choosing subsequences (with the same notation), we can arrange:

$$\begin{array}{lll} x_n \to x & y_n \to x & \text{(use that } K \text{ is compact)} \\ f_n \to f & & \text{(use that } \mathcal{F} \text{ is relatively compact)} \end{array}$$

This means that there is a $N \in \mathbb{N}$ such that for all $n \in \mathbb{N}_{>N}$ holds for all $y \in K$:

$$\|f_n(y) - f(y)\| < \frac{\varepsilon}{3}$$

(Since convergence in $C^0(K,E)$ is the same as uniform convergence $f_n \rightrightarrows f$.)
Since f is continuous there exists a $\delta \in \mathbb{R}_{>0}$ such that for all $y \in B_\delta(x)$:

$$\|f(x) - f(y)\| < \frac{\varepsilon}{3}$$

With this we get:

$$\|f_n(x) - f_n(y)\| \leq \underbrace{\|f_n(x) - f(x)\|}_{<\frac{\varepsilon}{3}} + \underbrace{\|f(x) - f(y)\|}_{<\frac{\varepsilon}{3}} + \underbrace{\|f(y) - f_n(y)\|}_{<\frac{\varepsilon}{3}} < \varepsilon$$

This is a contradiction to $\|f_n(x_n) - f_n(y_n)\| \geq \varepsilon$. $\square_{\text{i)}}$

ii) Consider $y_n \in \mathcal{F}(x) \subseteq E$ (to show that y_n has a convergent subsequence in E).
Then there are functions $f_n \in \mathcal{F}$ with $f_n(x) = y_n$. Since $\mathcal{F}$ is relatively compact, a subsequence is a Cauchy sequence in $C^0(K,E)$, i.e. $\left\|f_{n_l} \to f_{n_{l'}}\right\| \xrightarrow{l,l'\to\infty} 0$.

$$\left\|f_{n_l} - f_{n_{l'}}\right\| = \sup_{z\in K}\left\|f_{n_l}(z) - f_{n_{l'}}(z)\right\|_E \geq \left\|f_{n_l}(x) - f_{n_{l'}}(x)\right\|_E = \left\|y_{n_l} - y_{n_{l'}}\right\|$$

Therefore we get+:

$$\left\|y_{n_l} - y_{n_{l'}}\right\| \xrightarrow{l,l'\to\infty} 0$$

Thus (y_{n_l}) is a Cauchy sequence in E. $\square_{\text{ii)}}$

„$\Leftarrow$“: Let (f_l) be a sequence in $\mathcal{F}$ and show that a subsequence (g_l) converges in $C^0(K,E)$:
Since K is compact, there is a countable dense subset $\{x_1,x_2,\ldots\} \subseteq K$. Since $\mathcal{F}(x_1)$ is relatively compact, there is a subsequence $f_l^{(1)} \in \mathcal{F}$ of (f_l) such that $f_l^{(1)}(x_1)$ converges in E. Since $\mathcal{F}(x_2)$ is relatively compact, there is a subsequence $f_l^{(2)}$ of $f_l^{(1)}$ such that $f_l^{(2)}(x_2)$ converges.
Inductively choose a subsequence $\left(f_l^{(n+1)}\right)$ of $\left(f_l^{(n)}\right)$ such that $f_l^{(n+1)}(x_{n+1})$ converges in E. Take the diagonal sequence $g_l := f_l^{(l)}$. This is for $l \geq n$ a subsequence of $f_l^{(n)}$, so for all $n \in \mathbb{N}$ converges $g_l(x_n) \xrightarrow{l\to\infty} y_n$.

Claim: g_n is a Cauchy sequence in $C^0(K,E)$, i.e. for all $\varepsilon \in \mathbb{R}_{>0}$ exists a $N \in \mathbb{N}$ such that for all $n,m \in \mathbb{N}_{>N}$ and all $x \in K$ holds:

$$|g_n(x) - g_m(x)| \leq \varepsilon$$

Proof: Since $\mathcal{F}$ is equicontinuous, for all $x \in E$ exists a $\delta \in \mathbb{R}_{>0}$ such that for all $z,z' \in B_{\delta(x)}(x)$ and all $f \in \mathcal{F}$ holds:

$$\left\|f(z) - f(z')\right\| < \frac{\varepsilon}{3}$$

We cover K by a finite number of such balls $B_1,\ldots,B_L$. In every Ball B_l there is at least one point of $\{x_1,x_2,\ldots\}$. We choose such a point $\xi_l \in B_l$. Since $(g_n(\xi_l))$ converges for every $l \in \{1,\ldots,L\}$ we can choose a $N \in \mathbb{N}$ such that for all $l \in \{1,\ldots,L\}$ and all $m,n \in \mathbb{N}_{>N}$ holds:

$$\|g_n(\xi_l) - g_m(\xi_l)\| < \frac{\varepsilon}{3}$$

For every $x \in K$ exists a $l \in \{1,\ldots,L\}$ with $x \in B_l$.

$$\|g_n(x) - g_m(x)\| \leq \underbrace{\|g_n(x) - g_n(\xi_l)\|}_{<\frac{\varepsilon}{3}} + \underbrace{\|g_n(\xi_l) - g_m(\xi_l)\|}_{<\frac{\varepsilon}{3}} + \underbrace{\|g_m(\xi_l) - g_m(x)\|}_{<\frac{\varepsilon}{3}}$$

$\square_{\text{Claim}}$

Therefore the subsequence (g_l) for (f_l) converges in $C^0(K,E)$, since $C^0(K,E)$ is complete, because E is a Banach space. $\square_{6.3.3}$

Application to integral operators

Let $K \subseteq \mathbb{R}^n$ be compact. Consider an integral operator $A : C^0(K,\mathbb{R}) \to C^0(K,\mathbb{R})$, i.e.:

$$(Af)(x) = \int_K A(x,y)\, f(y)\, \mathrm{d}^n y$$

$\mathcal{F} := A\left(C^0(K,\mathbb{R})\right)$ is equicontinuous provided that $A(.,y)$ is continuous.

6.4 The Riesz representation theorem

Let K again be a compact metric space. $E = C^0(K,\mathbb{R})$ with the sup-norm is a Banach space.
Question: What is E^*?
Consider $l \in E^*$, i.e.

$$l : E \to \mathbb{R}$$

and for all $f \in C^0(K)$ holds:

$$|l(f)| \leq C \|f\|$$

This means f is bounded or equivalently continuous.

6.4.1 Examples

Consider $K = [0,1] \subseteq \mathbb{R}$. For any $\varphi \in L^1([0,1])$, the functional

$$l(f) := \int_0^1 \varphi(x) f(x) \,\mathrm{d}x$$

is linear and bounded:

$$|l(f)| \leq \int_0^1 |\varphi(x)| \cdot |f(x)| \,\mathrm{d}x \leq \underbrace{\sup_{x\in[0,1]} |f|}_{=\|f\|} \cdot \underbrace{\int_0^1 |\varphi(x)| \,\mathrm{d}x}_{=\|\varphi\|_{L^1}}$$

It is convenient to identify $l \in E^*$ with the function $\varphi \in L^1$. We have represented l by an L^1-function φ.
This can also be written as a *signed measure* (signiertes Maß):

$$\mathrm{d}\mu := \varphi(x) \,\mathrm{d}x$$

But not every $l \in E^*$ can be represented in this form.

Example

$$l(f) := f\left(\frac{1}{2}\right)$$

is bounded:

$$|l(f)| = \left|f\left(\frac{1}{2}\right)\right| \leq \sup_{[0,1]} |f| = \|f\|$$

It can be represented by the Dirac measure:

$$l(f) = \int_0^1 f(x)\,\delta\left(x - \frac{1}{2}\right) \mathrm{d}x = \int_0^1 f(x) \,\mathrm{d}\mu$$

Here $\delta(x)$ is the δ-Distribution. $\mu = \delta_{\frac{1}{2}}$ is the Dirac measure.

$$\delta_{x_0}(\Omega) = \begin{cases} 1 & \text{if } x_0 \in \Omega \\ 0 & \text{otherwise} \end{cases}$$

6.4.2 Definition (bounded, positive, regular measure)

Let $X \neq \emptyset$ be a set. A σ-algebra $\mathcal{M}$ over X is a set of subsets of X such that holds:

i) $\emptyset \in \mathcal{M}$

ii) $A \in \mathcal{M} \Rightarrow \complement A := X \setminus A \in \mathcal{M}$

iii) For a countable family $(A_j)_{j\in\mathbb{N}}$ holds:

$$\bigcup_{j=1}^{\infty} A_j \in \mathcal{M}$$

The elements of $\mathcal{M}$ are called *measurable sets* (messbare Mengen).

Let K be a compact metric space. Denote by $\mathfrak{M}$ the *Borel algebra*, i.e. the smallest σ-algebra over K, which contains all open and therefore all closed subsets of K.

A *bounded (signed) measure* is a mapping

$$\mu : \mathfrak{M} \to \mathbb{R}$$

(not $\mu : \mathfrak{M} \to \mathbb{R}^+ \cup \{0\} \cup \{\infty\}$ as before in measure theory) with the following properties:

– The empty set measures zero:

$$\mu(\emptyset) = 0$$

– σ-additivity: For $M_j \in \mathfrak{M}$ with $M_i \cap M_j = \emptyset$ for all $i \neq j$ holds:

$$\mu\left(\dot{\bigcup_{j=1}^{\infty}} M_j\right) = \sum_{j=1}^{\infty} \mu(M_j)$$

μ is *positive*, if $\mu(M) \geq 0$ for all $M \in \mathfrak{M}$.
μ is *regular*, if for all $A \in \mathfrak{M}$ holds:

$$\mu(A) = \sup_{\substack{B\subseteq A \\ B \text{ compact}}} \mu(B) = \inf_{\substack{\Omega\supseteq A \\ \Omega \text{ open}}} \mu(\Omega)$$

Example

The Lebesgue measure $\mathrm{d}^n x$ restricted to the Borel algebra on $[0,1]^n$ is a bounded, positive and regular measure.

6.4.3 Theorem (Riesz representation theorem)

Consider $l \in C^0(K,\mathbb{R})^*$. Then there is a unique bounded regular Borel measure μ (i.e. a measure on the Borel algebra $\mathfrak{M}$) such that for all $f \in C^0(K,\mathbb{R})$ holds:

$$l(f) = \int_K f \mathrm{d}\mu$$

Here we only prove the case $K = [0,1]$. (We also need it for $K = [0,1]^2$.)

How can one construct positive regular Borel measures on $[0,1]$?

Lebesgue-Stieltjes integral

Let $\alpha : [0,1] \to \mathbb{R}$ be monotonically increasing (not necessarily continuous).

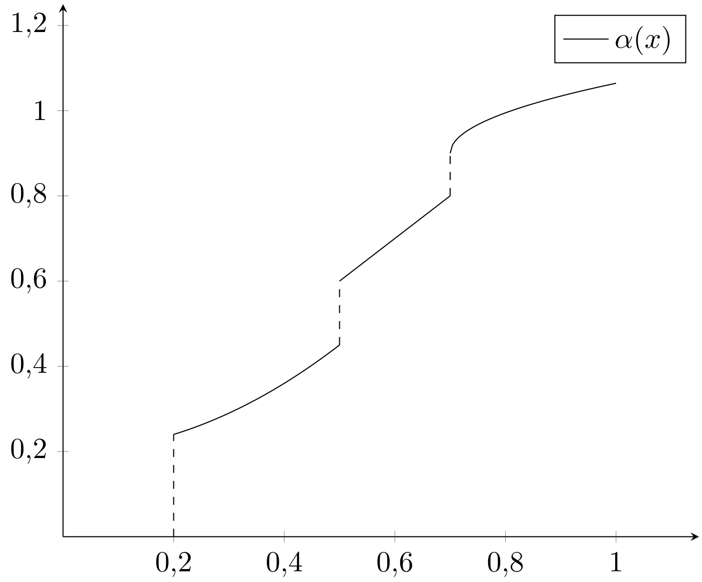


Figure 6.5: α is monotonically increasing, but not continuous

The two one-sided limits

$$\lim_{x \nearrow x_0} \alpha(x), \quad \lim_{x \searrow x_0} \alpha(x)$$

exist. In general:

$$\lim_{x \nearrow x_0} \alpha(x) \leq \alpha(x_0) \leq \lim_{x \searrow x_0} \alpha(x)$$

But equality does not need to hold. Define:

$$\mu((a,b)) := \lim_{x \nearrow b} \alpha(x) - \lim_{x \searrow a} \alpha(x)$$

By σ-additivity, this measure can be extended to a positive regular bounded Borel measure. (This can be proven exactly as for the Lebesgue integral.) The corresponding integral

$$\int_0^1 f \mathrm{d}\mu$$

is called Lebesgue-Stieltjes integral. If $\alpha(x) = x + c$, the Lebesgue-Stieltjes integral reduces to the Lebesgue integral

6.4.4 Example

Let $\alpha \in C^1([0,1])$ be monotonically increasing. Then holds:

$$\mu((a,b)) = \alpha(b) - \alpha(a) = \int_a^b \alpha'(x) \, \mathrm{d}x = \int_0^1 \chi_{(a,b)} \alpha'(x) \, \mathrm{d}x$$

The corresponding Lebesgue-Stieltjes integral is:

$$\int f \mathrm{d}\mu = \int_0^1 f(x) \cdot \alpha'(x) \,\mathrm{d}x$$

The following short notation is used in general:

$$\mathrm{d}\mu = \alpha'(x) \,\mathrm{d}x$$
$$\mathrm{d}\mu = \mathrm{d}\alpha$$

If $\alpha \in C^1([0,1])$ is not monotone, we can still set:

$$\int_0^1 f \mathrm{d}\mu := \int_0^1 f \cdot \alpha'(x) \,\mathrm{d}x$$

$\mathrm{d}\mu$ is a signed measure.

In order to extend the Lebesgue-Stieltjes construction to functions α, which are *not* monotone (such as to obtain signed measures), we need to assume, that α has bounded variation.

6.4.5 **Definition** (total variation)

Let $\alpha : [0,1] \to \mathbb{R}$ be a function (not necessarily continuous).
The *total variation* (Totalvariation) is defined by:

$$(\mathrm{TV}(\alpha))(x) := \sup_{\substack{N \in \mathbb{N} \\ 0 = x_0 < \ldots < x_N = x}} \sum_{i=1}^{N} |\alpha(x_1) - \alpha(x_{i-1})| \in \mathbb{R}_{\geq 0} \cup \{\infty\}$$

α is of *bounded variation* (beschränkte Totalvariation), $\alpha \in \mathcal{BV}([0,1])$, if $(\mathrm{TV}(f))(1) < \infty$.

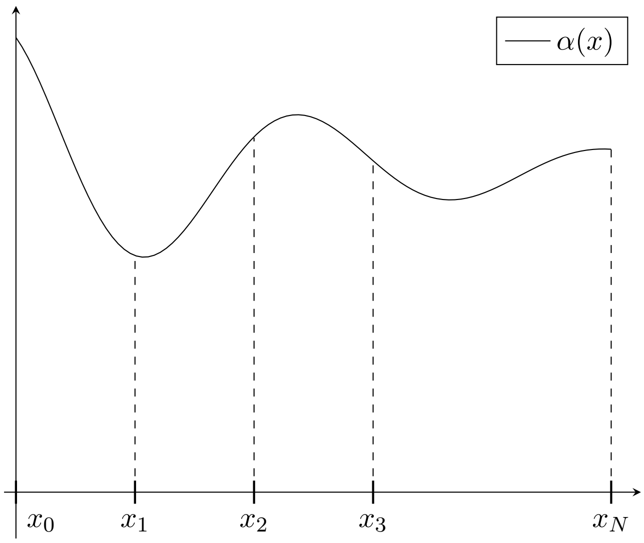


Figure 6.6: total variation of α

Note: If α is monotonically increasing, then holds:

$$(\mathrm{TV}(\alpha))(x) = \alpha(x) - \alpha(0) < \infty$$

Thus every monotonically function has bounded variation.

But there are even continuous functions, which have unbounded variation, e.g. for large enough $p \in \mathbb{R}_{>0}$:

$$\alpha(x) = x \sin\left(\frac{1}{x^p}\right)$$

For $\alpha \in C^1([0,1])$ holds:

$$\mathrm{TV}(\alpha)(x) = \int_0^x \left|\alpha'(\tau)\right| \mathrm{d}\tau$$

Lemma (Properties of the total variation)

$\mathrm{TV}(\alpha)(x)$ is monotonically increasing and:

$$\mathrm{TV}(\alpha)(0) = 0$$

$\mathrm{TV}(\alpha)(x) \pm \alpha(x)$ is also monotonically increasing.

Proof

Assume that $y \in \mathbb{R}_{>x}$.

$$\begin{aligned}\mathrm{TV}(\alpha)(y) = \sup_{\substack{N \in \mathbb{N} \\ 0 = x_0 < \ldots < x_N = y}} \sum_{i=1}^{N} |\alpha(x_i) - \alpha(x_{i-1})| &\geq \sup_{\substack{N \in \mathbb{N}_{\geq 2} \\ 0 = x_0 < \ldots < x_{N-1} = x < x_N = y}} \sum_{i=1}^{N} |\alpha(x_i) - \alpha(x_{i-1})| \geq \\ &\geq \sup_{\substack{N \in \mathbb{N}_{\geq 2} \\ 0 = x_0 < \ldots < x_{N-1} = x < x_N = y}} \sum_{i=1}^{N-1} |\alpha(x_i) - \alpha(x_{i-1})| = \mathrm{TV}(\alpha)(x)\end{aligned}$$

$$\mathrm{TV}(\alpha)(x) \pm \alpha(x) = \pm\alpha(0) + \sup_{\substack{N \in \mathbb{N} \\ 0 = x_0 < \ldots < x_N = x}} \sum_{i=1}^{N} \underbrace{|\alpha(x_i) - \alpha(x_{i-1})| \pm (\alpha(x_i) - \alpha(x_{i-1}))}_{\geq 0}$$

Just as before this implies that

$$\mathrm{TV}(\alpha)(x) \pm \alpha(x)$$

is monotonically increasing. $\square_{6.4.5}$

Suppose that $f \in \mathcal{BV}([0,1])$. Then the functions

$$\begin{aligned} f_+ &= \frac{1}{2}(\mathrm{TV}(f) + f) \\ f_- &= \frac{1}{2}(\mathrm{TV}(f) - f) \end{aligned}$$

are monotonically increasing and:

$$f = f_+ - f_-$$

Let $\mathrm{d}\mu_\pm = \mathrm{d}f_\pm$ be the bounded positive regular Borel measures of the corresponding Lebesgue-Stieltjes integrals. Then

$$\mu := \mu_+ - \mu_-$$

defines a bounded regular Borel measure with the property:

$$\begin{aligned}\mu((a,b)) &= \mu_+((a,b)) - \mu_-((a,b)) = \lim_{x\nearrow b} f_+(x) - \lim_{x\searrow a} f_+(x) - \lim_{x\nearrow b} f_-(x) + \lim_{x\searrow a} f_-(x) = \\ &= \lim_{x\nearrow b} f(x) - \lim_{x\searrow a} f(x)\end{aligned}$$

6.4.6 Example

Consider the Heaviside function:

$$f := \begin{cases} 0 & \text{if } x \leq \frac{1}{2} \\ 1 & \text{if } x > \frac{1}{2} \end{cases}$$

$\mathrm{d}\mu := \mathrm{d}f$ has the form $\mu = \delta_{\frac{1}{2}}$.

Proof of Theorem 6.4.3 in the case $K = [0,1]$

$\mathcal{PC}([0,1])$ are the piecewise continuous functions, i.e. for all $f \in \mathcal{PC}([0,1])$ exists a $N \in \mathbb{N}$ and points $0 = x_0 < \ldots < x_N = 1$ such that $f\big|_{(x_{i-1},x_i)}$ is continuous and has a continuous continuation to $[x_{i-1},x_i]$ for all $i \in \{1,\ldots,N\}$.
On $\mathcal{PC}$ we introduce the norm:

$$\|f\| = \sup_{x\in[0,1]} |f(x)|$$

This makes $\mathcal{PC}([0,1])$ a Banach space.

$$C^0([0,1]) \subseteq \mathcal{PC}([0,1])$$

is a subspace, which is closed, since it is complete.
Consider $l \in C^0([0,1])^*$, i.e.

$$l : C^0([0,1]) \to \mathbb{R}$$

with:

$$|l(f)| \leq C\,\|f\|_{C^0}$$

According to the Hahn-Banach theorem, there is an extension

$$\tilde{l} : \mathcal{PC}([0,1]) \to \mathbb{R}$$

with $\tilde{l}\big|_{C^0} = l$ and $|l(f)| \leq C\,\|f\|_{\mathcal{PC}([0,1])}$. Define $\alpha : [0,1] \to \mathbb{R}$ by:

$$\alpha(x) := \begin{cases} \tilde{l}\left(\chi_{[0,x)}\right) & \text{if } x < 1 \\ \tilde{l}\left(\chi_{[0,1]}\right) & \text{if } x = 1 \end{cases}$$

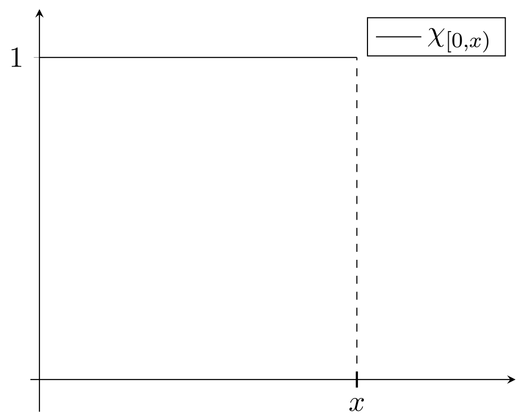


Figure 6.7: $\chi_{[0,x)}$

$l\left(\chi_{[0,x)}\right)$ is ill-defined, because $\chi_{[0,x)}$ is *not* continuous.

$\tilde{l}\left(\chi_{[0,x)}\right)$ is well-defined, because $\chi_{[0,x)}$ is piecewise-continuous.

– α has bounded variation: Consider:

$$0 = x_0 < \ldots < x_N = 1$$

We need to show:

$$\sum_{i=1}^{N} \left|\alpha\left(x_i\right) - \alpha\left(x_{i-1}\right)\right| < C$$

C has to be independent of N and the (x_i).
Define $s_i \in \{\pm 1\}$ by:

$$s_i := \begin{cases} +1 & \text{if } \alpha\left(x_i\right) - \alpha\left(x_{i-1}\right) \geq 0 \\ -1 & \text{if } \alpha\left(x_i\right) - \alpha\left(x_{i-1}\right) < 0 \end{cases}$$

Then holds:

$$\sum_{i=1}^{N} \left|\alpha\left(x_i\right) - \alpha\left(x_{i-1}\right)\right| = \sum_{i=1}^{N} s_i\left(\alpha\left(x_i\right) - \alpha\left(x_{i-1}\right)\right) = \tilde{l}\left(\sum_{i=1}^{N-1} s_i \chi_{[x_{i-1},x_i)} + s_N \chi_{[x_{N-1},1]}\right)$$

Since $\tilde{l}$ is bounded by construction, we know:

$$\begin{aligned}\sum_{i=1}^{N} \left|\alpha\left(x_i\right) - \alpha\left(x_{i-1}\right)\right| &\leq \left|\tilde{l}\left(\sum_{i=1}^{N-1} s_i \chi_{[x_{i-1},x_i)} + s_N \chi_{[x_{N-1},1]}\right)\right| \leq \\ &\leq C\left\|\sum_{i=1}^{N-1} s_i \chi_{[x_{i-1},x_i)} + s_N \chi_{[x_{N-1},1]}\right\| = C\end{aligned}$$

Therefore we have $\alpha \in \mathcal{BV}\left([0,1]\right)$.

– Consider $\mathrm{d}\mu := \mathrm{d}\alpha_+ - \mathrm{d}\alpha_-$ for the corresponding bounded regular Borel measure, where $\alpha = \alpha_+ - \alpha_-$ and $\alpha_\pm$ are monotonically increasing.

Claim: For all $f \in C^0([0,1])$ holds:

$$l(f) = \int_0^1 f \mathrm{d}\mu$$

Proof: Consider $f \in C^0([0,1])$. Set:

$$f_n(x) := \begin{cases} \sum_{i=1}^n f\left(\frac{i}{n}\right) \cdot \chi_{\left[\frac{i-1}{n}, \frac{i}{n}\right)} & \text{if } x < 1 \\ f(1) & \text{if } x = 1 \end{cases}$$

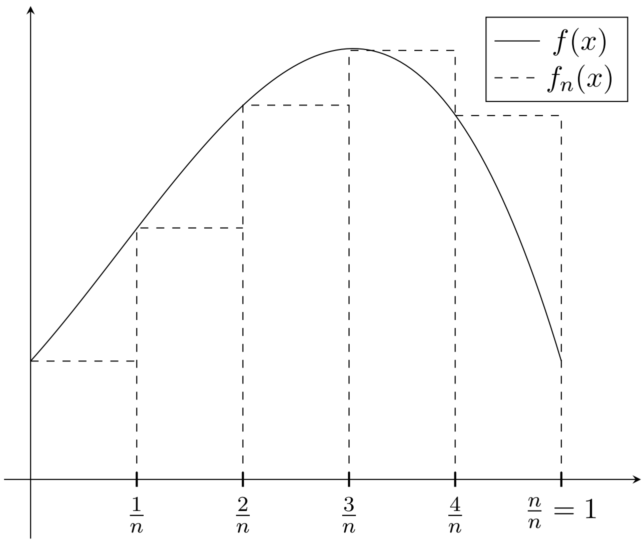


Figure 6.8: Approximation of f by $f\left(\frac{i}{n}\right)$ for $n = 5$

Since f_n is uniformly continuous, i.e. $f_n \rightrightarrows f$, we get:

$$l(f) = \tilde{l}(f) = \tilde{l}\left(\lim_{n\to\infty} f_n\right) \overset{\tilde{l} \text{ continuous}}{=} \lim_{n\to\infty} \tilde{l}(f_n) =$$
$$\overset{\text{by construction}}{=} \lim_{n\to\infty} \int_0^1 f_n \mathrm{d}\mu \overset{(*)}{=} \int_0^1 \lim_{n\to\infty} f_n \mathrm{d}\mu = \int_0^1 f \mathrm{d}\mu$$

For $(*)$ consider:

$$\left| \int_0^1 (f_n - f)\,\mathrm{d}\mu \right| \leq \underbrace{\sup |f - f_n|}_{\to 0} \cdot \underbrace{\mathrm{TV}(\alpha)(1)}_{<\infty} \xrightarrow{n\to\infty} 0$$

$\square_{\text{Claim}}$

$\square_{6.4.6}$

Remarks

- Our proof only works in the case $K = [a,b] \subseteq \mathbb{R}$. (see Reed, Simon: Appendix "The Riesz-Markov Theorem")

- In general dimension the idea would be:

$$\mu(\Omega) := \tilde{l}(\chi_\Omega)$$

But how to extend l? So choose $f_n \to \chi_\Omega$ and define:

$$\mu(\Omega) := \lim_{n\to\infty} l(f_n)$$

(see Rudin: *Real and complex analysis*)

- Total variation of a bounded Borel measure:

$$|\mu|(\Omega) := \sup_{\substack{N\in\mathbb{N} \\ \Omega_1,\dots,\Omega_N \\ \text{with } \Omega_1 \dot\cup \dots \dot\cup \Omega_N = \Omega}} \sum_{i=1}^{N} |\mu(\Omega_i)|$$

$|\mu|$ is a positive bounded Borel measure. (see Rudin)
Then we can write:

$$\left| \int_K (f - f_n)\,\mathrm{d}\mu \right| \leq \int_K |f - f_n| \cdot \mathrm{d}\,|\mu| \leq \sup_K |f - f_n| \cdot |\mu|(K)$$

7 The Spectral Theorem for symmetric bounded operators

Let $A \in L(H)$ be symmetric and H be a separable Hilbert space. Let $p(A)$ be a polynomial in A, for example the characteristic polynomial for $A \in L\left(\mathbb{C}^N\right)$ with $p(A) = 0$.
Extend this idea to functions $f(A)$ with $f \in C^0(\sigma(A))$. (Stone-Weierstraß)
Then for

$$\langle u, f(A) u \rangle =: l(f)$$

holds $l \in C^0(\sigma(A))^*$. Using the Riesz representation theorem we can write:

$$\langle u, f(A) u \rangle = \int_{\sigma(A)} f(\lambda) \, \mathrm{d}\mu_u(\lambda)$$

$$\mathrm{d}\mu_u(\lambda) = \langle u, \mathrm{d}E_\lambda u \rangle$$

$\mathrm{d}E_\lambda$ is the so-called *spectral measure*. Then holds the spectral theorem:

$$A = \int_{\sigma(A)} \lambda \mathrm{d}E_\lambda$$

7.1 The Spectrum of symmetric bounded operators

Let $A \in L(H)$ be symmetric, i.e. $\langle u, Av \rangle = \langle Au, v \rangle$ for all $u, v \in H$. The resolvent set is:

$$\begin{aligned} \varrho(A) &= \left\{ \lambda \in \mathbb{C} \middle| \, (\lambda - A) \text{ has a continuous inverse} \right\} \\ \sigma(A) &= \mathbb{C} \setminus \varrho(A) \end{aligned}$$

$\varrho(A) \subseteq \mathbb{C}$ is open and so the spectrum $\sigma(A) \subseteq \mathbb{C}$ is closed. The spectral radius is:

$$r(A) = \sup_{\lambda \in \sigma(A)} |\lambda| = |||A|||$$

Warning

Consider $\lambda \in \sigma(A)$, i.e. $\lambda - A$ has no continuous inverse. This does not mean $\ker(\lambda - A)$ is non-trivial. Thus λ does *not* need to be an eigenvalue!

7.1.1 Theorem

Let $A \in L(H)$ be self-adjoint. Then $\sigma(A) \subseteq \mathbb{R}$.

Proof

Consider $\lambda = \alpha + \mathbf{i}\beta$ with $\alpha,\beta \in \mathbb{R}$ and $\beta \neq 0$. We need to show that $\lambda - A$ has a continuous inverse. Introduce the following bilinear form:

$$B(x,y) = \left\langle x, \left(A - \overline{\lambda}\right) y \right\rangle = \langle (A - \lambda) x, y \rangle$$

This bilinear form satisfies the assumptions of the Lax-Milgram theorem:

i) The sesquilinearity is clear, since the scalar product is sesquilinear.

ii) B is bounded:

$$\left|\left\langle x, \left(A - \overline{\lambda}\right) y \right\rangle\right| \leq \|x\| \cdot \underbrace{\left\|A - \overline{\lambda}\right\|}_{\leq \|A\| + |\lambda|} \cdot \|y\| \leq C \|x\| \|y\|$$

iii) B is bounded from below, i.e. there exists an $\varepsilon \in \mathbb{R}_{>0}$ such that for all $x \in H$ holds:

$$|B(x,x)| \geq \varepsilon \|x\|^2$$

We know:

$$B(x,x) = \left\langle x, \left(A - \overline{\lambda}\right) x \right\rangle = \underbrace{\langle x, Ax \rangle}_{\text{real}} - \underbrace{\mathrm{Re}(\lambda \langle x, x \rangle)}_{\text{real}} - \underbrace{\mathbf{i}\mathrm{Im}(\lambda \langle x, x \rangle)}_{\text{imaginary}}$$

$$|B(x,x)| \geq |\mathrm{Im}(\lambda \langle x, x \rangle)| = |\beta| \cdot \|x\|^2$$

Set $\varepsilon := |\beta| \neq 0$.

The Lax-Milgram theorem yields that the linear functional $l(x) = \langle z, x \rangle$ can be represented as

$$l(x) = B(y,x)$$

with a unique $y = y(z) \in H$. Thus we get for all $x \in H$:

$$\begin{aligned} \langle z, x \rangle &= \langle (A - \lambda) y, x \rangle \\ \Rightarrow \quad z &= (A - \lambda) y \end{aligned}$$

Therefore, for all $z \in H$ exists a unique $y \in H$ su ch that $(A - \lambda) y = x$. Thus $A - \lambda$ is invertible. The inverse $(A - \lambda)^{-1}$ is continuous due to the open mapping theorem (see Corollary 2.4.8). $\square_{7.1.1}$

7.1.2 Theorem

It holds $\sigma(A) \subseteq [a,b]$ and $a,b \in \sigma(A)$ with:

$$a := \inf_{\|u\|=1} \langle u,Au\rangle$$
$$b := \sup_{\|u\|=1} \langle u,Au\rangle$$

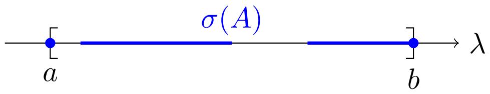


Figure 7.1: $\sigma(A) \subseteq [a,b]$ and $a,b \in \sigma(A)$

Proof

For $\lambda \in \mathbb{R}_{<a}$ holds:

$$\langle x, (A-\lambda)\, x\rangle = \langle x,Ax\rangle - \lambda \|x\|^2 \geq a\|x\|^2 - \lambda\|x\|^2 = \underbrace{(a-\lambda)}_{>0} \|x\|^2$$

Thus

$$\langle .,.\rangle_A := \langle ., (A-\lambda)\, .\rangle$$

is a scalar product on H. The corresponding norm

$$\|u\|_A := \sqrt{\langle u,u\rangle_A}$$

is equivalent to the norm $\|.\|$, because it holds:

$$(a-\lambda)\|u\|^2 \leq \|u\|_A = \langle u, (A-\lambda)\, u\rangle \leq (|||A||| - \lambda)\|u\|^2$$

For $u \in H$ and $l(w) := \langle u,w\rangle$ is $l \in H^*$. According to the Fréchet-Riesz theorem 3.1.3 (for the scalar product $\langle .,.\rangle_A$) there is a unique vector $v \in H$, such that for all $w \in H$ holds:

$$l(w) = \langle v,w\rangle_A$$

Thus we get for all $w \in H$:

$$\langle u,w\rangle = l(w) = \langle v,w\rangle_A = \langle v, (A-\lambda)\, w\rangle \overset{A-\lambda \text{ symmetric}}{=} \langle (A-\lambda)\, v,w\rangle$$

$$\Rightarrow \quad u = (A-\lambda)\, v$$

Thus there exists a

$$\varphi : H \to H$$
$$u \mapsto v$$

such that $u = (A - \lambda)\varphi(u)$, i.e. $A - \lambda \in L(H)$ is surjective. φ is linear and bounded according to the open mapping theorem 2.4.8. Thus we have

$$\varphi = (A - \lambda)^{-1} \in L(H)$$

and therefore $\lambda \in \varrho(A)$.
Applying the same argument to the operator $(-A)$, one sees that $(b,\infty) \subseteq \varrho(A)$.
Therefore holds $\sigma(A) \subseteq [a,b]$.

Only prove that $b \in \sigma(A)$. For $a \in \sigma(A)$ consider similarly the operator $-A$. Furthermore replace $A \to A - a$ to get $\sigma(A) \subseteq [0,b]$. We know:

$$|||A||| = r(A) = \sup_{\lambda \in \sigma(A)} |\lambda| = \sup_{\lambda \in \sigma(A)} \lambda = \sup \sigma(A)$$

As a consequence we get $|||A||| \leq b$. On the other hand we have:

$$b = \sup_{\|u\|=1} \langle u,Au \rangle \leq \sup_{\|u\|=1} \|Au\| \cdot \underbrace{\|u\|}_{=1} = |||A|||$$

Thus we have $b = |||A||| = r(A)$, especially b is a limit point of the spectrum of A. Since $\sigma(A)$ is closed, it follows that $b \in \sigma(A)$. $\square_{7.1.2}$

7.2 The continuous functional calculus

7.2.1 Theorem (continuous functions of operators)

Let $A \in L(H)$ be symmetric. Then there is a unique mapping $\Phi : C^0(\sigma(A),\mathbb{C}) \to L(H)$ (remember $\sigma(A) \subseteq [a,b]$) with the following properties:

i) Φ is an involutive algebra homomorphism, i.e.:
  - Φ is linear.
  - $\Phi(f \cdot g) = \Phi(f) \cdot \Phi(g)$
  - $\Phi(\overline{f}) = (\Phi(f))^*$ (involution)

ii) Φ is continuous:

$$\|\Phi(f)\|_{L(H)} \leq C \|f\|_\infty$$

iii) If $f(t) = t$, then $\Phi(f) = A$.

iv) If $Au = \lambda u$, i.e. $u \in H$ is an eigenvector of A, then $\Phi(f)u = f(\lambda)u$.

v) If $f \geq 0$, then $\Phi(f) \geq 0$, meaning that $\Phi(f)$ is a positive semi-definite operator, i.e. $\langle u,\Phi(f)u \rangle \geq 0$ for all $u \in H$.

vi) $\sigma(\Phi(f)) = f(\sigma(A))$ (spectral mapping theorem (spektraler Abbildungssatz))

vii) $\|\Phi(f)\|_{L(H)} = \|f\|_\infty$

Often we just write $\Phi(f) = f(A)$.

What if $f(t) = p(t) = a_n t^n + a_{n-1}t^{n-1} + \ldots + a_0$ is a polynomial?

$$\Phi(t) \overset{\text{iii)}}{=} A$$

From i) follows:

$$\Phi(1) = \Phi(1 \cdot 1) = \Phi(1) \cdot \Phi(1)$$

Therefore we get:

$$\Phi(1) = \mathbb{1}$$

Now follows:

$$\begin{aligned}
\Phi\left(t^2\right) &= \Phi(t \cdot t) = \Phi(t) \cdot \Phi(t) = A \cdot A = A^2 \\
\Phi\left(t^l\right) &= A^l \\
\Phi(p) &= p(A) = a_n A^n + a_{n-1}A^{n-1} + \ldots + a_0 \mathbb{1}
\end{aligned}$$

7.2.2 **Lemma** (spectral mapping theorem for polynomials)

For $p \in \mathfrak{P}(\mathbb{C})$, i.e. p is a complex polynomial, holds:

$$\sigma(p(A)) = p(\sigma(A))$$

Proof

- If $p = c \in \mathbb{C}$ is constant, then the lemma is trivial:

  $$p(\sigma(A)) = c = \sigma(c\mathbb{1}) = \sigma(p(A))$$

  So further on let p be not constant.
- $p(\sigma(A)) \subseteq \sigma(p(A))$: For $\lambda \in \sigma(A)$ and $z \in \mathbb{C}$ yields the fundamental theorem of algebra:

  $$p(z) - p(\lambda) = (z - \lambda)\, q(z)$$

  Here $q(z)$ is a new polynomial with $\deg(q) = \deg(p) - 1$. This also holds if we set $z = A$:

  $$p(A) - p(\lambda) = (A - \lambda)\, q(A)$$

  Assume $p(\lambda) \in \varrho(p(A))$, i.e. $p(A) - p(\lambda)$ has a bounded inverse. Then holds:

  $$\begin{aligned}
  & \mathbb{1} = (p(A) - p(\lambda)) \cdot (p(A) - p(\lambda))^{-1} = (A - \lambda) \cdot q(A) \cdot (p(A) - p(\lambda))^{-1} \\
  \Rightarrow \quad & (A - \lambda)^{-1} = \underbrace{q(A)}_{\in L(H)} \cdot \underbrace{(p(A) - p(\lambda))^{-1}}_{\in L(H)} \in L(H)
  \end{aligned}$$

  This gives $\lambda \in \varrho(A)$ in contradiction to $\lambda \in \sigma(A)$ and so $p(\lambda) \in \sigma(p(A))$.

– $\sigma(p(A)) \subseteq p(\sigma(A))$: Consider $\mu \in \sigma(p(A))$ and set $n := \deg(p)$. Using the fundamental theorem of algebra we get:

$$q(z) := p(z) - \mu = a(z-\lambda_1) \cdot \ldots \cdot (z-\lambda_n)$$
$$q(A) := p(A) - \mu = a(A-\lambda_1) \cdot \ldots \cdot (A-\lambda_n)$$

If all the operators $A - \lambda_i$ had a continuous inverse, then this would hold also for their product in contradiction to the assumption $\mu \in \sigma(p(A))$. Thus one of the λ_i is in the spectrum of A. Because one of the linear factors vanishes, follows:

$$0 = q(\lambda_i) = p(\lambda_i) - \mu$$
$$\Rightarrow \quad \mu = p(\lambda_i) \in p(\sigma(A))$$

$\square_{7.2.2}$

Let $p \in \mathfrak{P}(\mathbb{C})$ be a complex polynomial.

$$(p(A))^* = \overline{p}(A)$$

Thus $p(A)$ is not symmetric.

7.2.3 Definition (normal operator)

$A \in L(H)$ is called *normal*, if $[A,A^*] = 0$.

7.2.4 Theorem

For a normal $A \in L(H)$ holds $r(A) = |||A|||$.

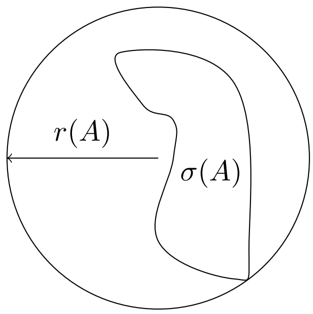


Figure 7.2: $r(A) = |||A|||$

Proof

We already proved for a general $A \in L(H)$:

$$r(A) = \sup_{\lambda \in \sigma(A)} |\lambda| = \lim_{n\to\infty} |||A^n|||^{\frac{1}{n}} \tag{7.1}$$

For symmetric operators, we know furthermore:

$$r(A) = |||A||| = \sup_{\|u\|=1} |\langle u, Au\rangle| \tag{7.2}$$

For *normal* operators, we conclude the following: A^*A is symmetric and thus:

$$\begin{aligned}|||A|||^2 &= \sup_{\|u\|=1} \|Au\|^2 = \sup_{\|u\|=1} \langle Au,Au\rangle = \sup_{\|u\|=1} \langle u,A^*Au\rangle \overset{(7.2)}{=} |||A^*A||| = \\ &\overset{(7.2)}{=} r\left(A^*A\right) \overset{(7.1)}{=} \lim_{n\to\infty} |||(A^*A)^n|||^{\frac{1}{n}}\end{aligned}$$

$$(A^*A)^n = \underbrace{A^*A \cdot A^*A \cdot \ldots \cdot A^*A}_{n\text{-times}} \overset{A \text{ normal}}{=} (A^*)^n \cdot A^n$$

With

$$|||A|||^2 = \sup_{\|u\|=1} \langle Au,Au\rangle = \sup_{\|u\|=1} \langle u,A^*Au\rangle \overset{A \text{ normal}}{=} \sup_{\|u\|=1} \langle u,AA^*u\rangle = \sup_{\|u\|=1} \langle A^*u,A^*u\rangle = |||A^*|||^2$$

we get:

$$|||(A^*A)^n||| \leq |||(A^*)^n||| \cdot |||A^n||| = |||A^n|||^2$$

It follows:

$$|||A|||^2 = \lim_{n\to\infty} |||(A^*A)^n|||^{\frac{1}{n}} \leq \lim_{n\to\infty} \left(|||A^n|||^2\right)^{\frac{1}{n}} \leq |||A|||^2$$

This gives:

$$|||A|||^2 = \lim_{n\to\infty} \left(|||A^n|||^{\frac{1}{n}}\right)^2 = \left(\lim_{n\to\infty} |||A^n|||^{\frac{1}{n}}\right)^2 = \left(r\left(A\right)\right)^2$$

$$\Rightarrow \quad r\left(A\right) = |||A|||$$

$\square_{7.2.4}$

7.2.5 Lemma

Let $A \in L\left(H\right)$ be symmetric and $p \in \mathfrak{P}\left(\mathbb{C}\right)$ a complex polynomial. Then holds:

$$|||p\left(A\right)||| = \sup_{\lambda\in\sigma(A)} \left|p\left(\lambda\right)\right|$$

Proof

$p\left(A\right)$ is normal and thus, according to Theorem 7.2.4 holds:

$$|||p\left(A\right)||| = \sup_{\mu\in\sigma(p(A))} |\mu| \overset{7.2.2}{=} \sup_{\lambda\in\sigma(A)} \left|p\left(\lambda\right)\right|$$

$\square_{7.2.5}$

Proof of theorem 7.2.1

– For complex polynomials, we set $\Phi(p) = p(A)$. Then holds:

$$|||\Phi(p)||| = |||p(A)||| = r(p(A)) = \sup_{\lambda\in\sigma(A)} |p(\lambda)| = \|p\|_{C^0(\sigma(A),\mathbb{C})}$$

Thus $\Phi : \mathfrak{P}(\mathbb{C}) \to L(H)$ is an isometry. ($\mathfrak{P}(\mathbb{C}) \subseteq C^0(\sigma(A),\mathbb{C})$)
Remark: If we had considered $C^0([a,b],\mathbb{C})$ with

$$a = \inf_{\|u\|=1} \langle u,Au\rangle$$
$$b = \sup_{\|u\|=1} \langle u,Au\rangle$$

then we would only have an inequality:

$$|||\Phi(p)||| \le \|p\|_{C^0([a,b])}$$

– Moreover holds:

$$\Phi(p\cdot q) = (p\cdot q)(A) = p(A)\cdot q(A) = \Phi(p)\cdot\Phi(q)$$
$$(\Phi(p))^* = \Phi(\overline{p})$$

– Using the Stone-Weierstraß approximation theorem, Φ uniquely extends to an isometry:

$$\Phi : C^0(\sigma(A),\mathbb{C}) \to L(H)$$

This yields i), ii), iii), vii).

– More specifically, consider $f \in C^0(\sigma(A),\mathbb{C})$. Then there exist $p_n \in \mathfrak{P}(\mathbb{C})$ such that $p_n \rightrightarrows f$ on $\sigma(A)$. ($K = \sigma(A)$ is a compact metric space.) This means:

$$\|p_n - f\|_{C^0(\sigma(A),\mathbb{C})} = \sup_{z\in\sigma(A)} |p_n(z) - f(z)| \xrightarrow{n\to\infty} 0$$

$$|||\Phi(p_n) - \Phi(p_m)||| \overset{\text{isometry}}{=} \|p_n - p_m\| \xrightarrow{n,m\to\infty} 0$$

Thus the operators $\Phi(p_n)$ form a Cauchy sequence in $L(H)$ and since $L(H)$ is a Banach space, this sequence converges to:

$$\Phi(f) := \lim_{n\to\infty} \Phi(p_n)$$

iv) For $Au = \lambda u$ holds:

$$\Phi(f)u = \lim_{n\to\infty} \Phi(p_n)u = \lim_{n\to\infty} p_n(A)u = \lim_{n\to\infty} p_n(\lambda)u = f(\lambda)u$$

vi) Now we prove the spectral mapping theorem:
„$\subseteq$“: Assume $\mu \in \sigma(f(A))$, but $\mu \notin f(\sigma(A))$. Then holds $f - \mu \neq 0$ on $\sigma(A)$ and we can invert:

$$\frac{1}{f-\mu} \in C^0(\sigma(A),\mathbb{C})$$

Now follows:

$$\mathbb{1} = \Phi(1) = \Phi\left(\frac{1}{f-\mu}(f-\mu)\right) = \underbrace{\Phi\left(\frac{1}{f-\mu}\right)}_{\in L(H)} \cdot \underbrace{\Phi(f-\mu)}_{=f(A)-\mu\mathbb{1}}$$

So $f(A) - \mu\mathbb{1}$ has a bounded inverse in contradiction to the assumption $\mu \in \sigma(f(A))$.
„$\supseteq$“: Consider $\lambda \in \sigma(A)$. Choose polynomials $p_n \in \mathfrak{P}(\mathbb{C})$ with $p_n \rightrightarrows f$. Then converges in $L(H)$:

$$p_n(A) - p_n(\lambda)\,\mathbb{1} \xrightarrow{n\to\infty} f(A) - f(\lambda)\,\mathbb{1}$$

Assume that $f(\lambda) \notin \sigma(f(A))$. Then $f(A) - f(\lambda)\,\mathbb{1}$ has a bounded inverse.
According to Theorem 2.5.3, the invertible operators are open in $L(H)$. Therefore there exists a $\delta \in \mathbb{R}_{>0}$ such that B has a bounded inverse for all $B \in B_\delta(f(A) - f(\lambda)\,\mathbb{1})$. In particular, the operators $p_n(A) - p_n(\lambda)\,\mathbb{1}$ have a bounded inverse for sufficiently large n. This is a contradiction to the spectral mapping theorem for polynomials 7.2.2.

v) Claim: $f \geq 0 \quad \Rightarrow \quad \Phi(f) \geq 0$
Let $f \in C^0(\sigma(A), \mathbb{R})$ be real-valued and $f \geq 0$. Then $g := \sqrt{f} \in C^0(\sigma(A), \mathbb{R})$ and $f = g^2$.

$$\langle u, \Phi(f)\,u\rangle = \left\langle u, \Phi\left(g^2\right)u\right\rangle = \langle u, \Phi(g)\,\Phi(g)\,u\rangle = \langle \Phi(\overline{g})\,u, \Phi(g)\,u\rangle = \langle \Phi(g)\,u, \Phi(g)\,u\rangle \geq 0$$

$\square_{7.2.5}$

$\chi_\Omega(A)$ would be the projector onto the invariant subspace corresponding to the spectrum in Ω. Formally we can compute:

$$(\chi_\Omega(A))^* = \overline{\chi_\Omega}(A) = \chi_\Omega(A)$$
$$\chi_\Omega(A)\,\chi_\Omega(A) = \chi_\Omega^2(A) = \chi_\Omega(A)$$

This motivates, why we would like to form $f(A)$ for a bounded Borel function f on $\sigma(A)$.

7.3 Spectral Measures

Let $A \in L(H)$ be symmetric. Choose a $u \in H$ (fixed).

$$\Phi_u : C^0(\sigma(A), \mathbb{R}) \to \mathbb{R} \subseteq \mathbb{C}$$
$$f \mapsto \langle u, \Phi(f)\,u\rangle$$

$$|\Phi_u(f)| = |\langle u, \Phi(f)\,u\rangle| \leq |||\Phi(f)||| \cdot \|u\|^2 = \|f\|_{C^0(\sigma(A),\mathbb{R})} \cdot \|u\|^2$$

Thus ϕ_u is a bounded linear functional on $C^0(\sigma(A), \mathbb{R})$. According to the Riesz representation theorem there exists a unique regular bounded Borel measure μ_u such that:

$$\langle u, f(A)\,u\rangle = \int_{\sigma(A)} f(\lambda)\,\mathrm{d}\mu_u(\lambda)$$

The measure μ_u is even positive, because if $f \geq 0$, set $g = \sqrt{f}$ to get:

$$\int_{\sigma(A)} f(\lambda)\,\mathrm{d}\mu_u(\lambda) = \langle u, f(A)\,u\rangle = \langle g(A)\,u, g(A)\,u\rangle \geq 0 \qquad \forall_{f\in C^0(\sigma(A),\mathbb{R}),\ f\geq 0}$$

Hence by approximation follows $\mu_u(\Omega) \geq 0$ for all Borel sets $\Omega \subseteq \sigma(A)$. So μ_u is a positive measure.
The resulting integral can be defined for a more general class of functions.
A *Borel function* f is a function, which is measurable for the Borel algebra, i.e. $f^{-1}(\Omega)$ is a Borel function for all open $\Omega \subseteq \mathbb{C}$.
We use the following notation: $\mathfrak{M}$ is the set of all Borel sets in $\sigma(A)$.
$\mathcal{B}(\sigma(A),\mathbb{R}) = L^\infty(\mathrm{d}\mu_u)$ are the bounded Borel functions on $\sigma(A)$. We always assume:

$$\sup_{\sigma(A)} |f| < \infty$$

We define:

$$\begin{aligned}\phi_u : \mathcal{B}(\sigma(A),\mathbb{R}) &\to \mathbb{R}\\ \phi_u(f) &:= \int_{\sigma(A)} f(\lambda)\,\mathrm{d}\mu_u(\lambda)\end{aligned}$$

7.3.1 Lemma

$$|\phi_u(f)| \leq \|f\|_{L^\infty} \cdot \|u\|^2$$

Proof

For $f \in \mathcal{B}(\sigma(A),\mathbb{R})$ choose $\varphi_n \in C^0(\sigma(A),\mathbb{R})$ such that $\varphi_n \to f$ converges point-wise and $\|\varphi_n\|_\infty \leq \|f\|_\infty$. (Approximate f by step-functions and then approximate the step functions by continuous functions.)
Due to $|\varphi_n| \leq C$ and

$$\int_{\sigma(A)} C\mathrm{d}\mu_u = C\mu_u(\sigma(A)) = C\langle u,\Phi(1)u\rangle = C\langle u,\mathbb{1}u\rangle = C\|u\|^2 < \infty$$

we can use the dominated convergence theorem:

$$\begin{aligned}\left|\int_{\sigma(A)} f\mathrm{d}\mu_u\right| \overset{\text{dominated}}{\underset{\text{convergence}}{=}} \lim_{n\to\infty}\left|\int_{\sigma(A)}^{\infty} \varphi_n\mathrm{d}\mu_n\right| &= \lim_{n\to\infty} |\langle u,\Phi(\varphi_n)u\rangle| \leq\\ \leq \lim_{n\to\infty}\|u\|^2 \cdot \|\Phi(\varphi_n)\| = \lim_{n\to\infty}\|u\|^2 \cdot \|\varphi_n\| &\leq \|f\| \cdot \|u\|^2\end{aligned}$$

$\square_{7.3.1}$

Define using the Fréchet-Riesz theorem the unique Operator $\Phi(f)$ by:

$$\Phi_u(f) := \langle u,\Phi(f)u\rangle$$

By polarization we get:

$$B_f(u,v) = \Phi_{\frac{u+v}{2}}(f) - \Phi_{\frac{u-v}{2}}(f) - \mathbf{i}\Phi_{\frac{u+\mathbf{i}v}{2}}(f) + \mathbf{i}\Phi_{\frac{u-\mathbf{i}v}{2}}(f)$$

Alternatively define for $f \in C^0(\sigma(A), \mathbb{C})$:

$$\Phi_{u,v}(f) := \langle u, \Phi(f)\, v\rangle = \int_{\sigma(A)} f(\lambda)\, \mathrm{d}\mu_{u,v}(\lambda)$$

$$B_f(u,v) := \int_{\sigma(A)} f(\lambda)\, \mathrm{d}\mu_{u,v}(\lambda)$$

$\mathrm{d}\mu_{u,v}$ is a only a *complex-valued,* bounded, regular Borel measure.

7.3.2 Lemma

$B_f(u,v)$ is a *sesquilinear form*, i.e. linear in the second and anti-linear in the first argument, and it holds:

$$|B_f(u,v)| \leq \|f\| \cdot \|u\| \cdot \|v\|$$

Proof

This follows from the polarization formula and Lemma 7.3.1. $\square_{7.3.2}$

7.3.3 Theorem

Let B be a bounded sesquilinear form, i.e.:

$$|B(u,v)| \leq C \cdot \|u\| \cdot \|v\| \qquad \underset{u,v \in H}{\forall}$$

Then there is a unique operator $D \in L(H)$ with $|||D||| \leq C$ such that:

$$B(u,v) = \langle u, Dv\rangle$$

Proof

For $v \in H$ the map

$$\psi := \overline{B(.,v)}$$

is a bounded linear form. According to the Fréchet-Riesz theorem 3.1.3 there exists a $w \in H$ such that for all $u \in H$ holds:

$$\psi(u) = \langle w, u\rangle$$

Then follows:

$$B(u,v) = \overline{\langle w, u\rangle} = \langle u, w\rangle$$

Thus D is uniquely determined by $Dv = w$. So $D : H \to H$ is linear and bounded by the open mapping principle 2.4.7, i.e. $D \in L(H)$ and for all $v \in H$ holds:

$$B(u,v) = \langle u, Dv\rangle$$

Choose $u = Dv$ to get:

$$\begin{aligned} B(Dv,v) &= \langle Dv,Dv\rangle = \|Dv\|^2 \\ &\leq C\cdot\|Dv\|\cdot\|v\| \end{aligned}$$

Therefore we have for all $v \in H$:

$$\begin{aligned} \|Dv\| &\leq C\cdot\|v\| \\ |||D||| &\leq C \end{aligned}$$

$\square_{7.3.3}$

We conclude: For $f \in \mathcal{B}(\sigma(A),\mathbb{C})$ we construct $B_f(u,v)$. Then there exists a $\Phi(f) \in L(H)$ such that for all $u,v \in H$ holds:

$$\langle u,\Phi(f)\,v\rangle = B_f(u,v)$$

So $\Phi : \mathcal{B}(\sigma(A),\mathbb{C}) \to L(H)$ gives a functional calculus on $\mathcal{B}(\sigma(A),\mathbb{C})$, i.e. we can calculate $f(A)$ for an arbitrary Borel function.

7.3.4 Theorem (Spectral theorem in functional calculus form)

Let $A \in L(H)$ be symmetric. Then there is a unique mapping $\Phi : \mathcal{B}(\sigma(A)) \to L(H)$ with the following properties:

i) Φ is an involutive algebra homomorphism, i.e.:

$$\begin{aligned} \Phi(f)\cdot\Phi(g) &= \Phi(f\cdot g) \\ \Phi(f)^* &= \Phi\left(\overline{f}\right) \end{aligned}$$

If $f \in C^0(\sigma(A),\mathbb{C})$, then $\Phi(f)$ agrees with the corresponding operator of the continuous functional calculus.

ii) $|||\Phi(f)||| \leq \|f\|_\infty$

iii) If $f_n \to f$ converges point-wise and it holds $\|f_n\|_\infty < C$, then $\Phi(f_n) \to \Phi(f)$ converges strongly, i.e. for all $u \in H$ converges in H:

$$\Phi(f_n)\,u \to \Phi(f)\,u$$

iv) From $Au = \lambda u$ follows:

$$\Phi(f)\,u = f(\lambda)\,u$$

v) If $f \geq 0$ holds, then $\Phi(f) \geq 0$ is positive semidefinite.

vi) If $B \in L(H)$ commutes with A, i.e. $[A,B] = AB - BA = 0$, then $[B,\Phi(f)] = 0$. We write also $f(A) = \Phi(f)$.

Note: There is no spectral mapping theorem.

Proof

i) Prove the homomorphism property by approximation:
First step: Assume $f \in C^0(\sigma(A),\mathbb{C})$ and $g \in \mathcal{B}(\sigma(A),\mathbb{C})$. Then there exists a series $g_n \in C^0$ such that $g_n \to g$ converges point-wise and $\|g_n\|_\infty < C$. Then follows the point-wise convergence:

$$f g_n \to f g$$

We use the notation:

$$\begin{aligned} &\phi_{u,v}(h) := \langle u,\Phi(h)\,v\rangle \\ \Rightarrow \qquad &\phi_{u,u}(h) = \phi_u(h) \end{aligned}$$

Since μ_u is a regular bounded Borel measure, we can apply the dominated convergence theorem:

$$\begin{aligned} \phi_{u,u}(f\cdot g) &\overset{\text{Definition}}{=} \int_{\sigma(A)} f\cdot g \mathrm{d}\mu_u \underset{\text{convergence}}{\overset{\text{dominated}}{=}} \lim_{n\to\infty} \int_{\sigma(A)} f\cdot g_n \mathrm{d}\mu_u = \lim_{n\to\infty} \phi_{u,u}(f,g_n) = \\ &= \lim_{n\to\infty} \langle u,\Phi(f\cdot g_n)\,u\rangle = \lim_{n\to\infty} \langle u,f(A)\cdot g_n(A)\,u\rangle = \\ &= \lim_{n\to\infty} \langle (f(A))^* u,g_n(A)\,u\rangle = \lim_{n\to\infty} \phi_{(f(A))^* u,u}(g_n) \end{aligned}$$

We know for all $u \in H$ using dominated convergence (see above):

$$\phi_{u,u}(g_n) \to \phi_{u,u}(g)$$

By polarization follows for all $u,v \in H$:

$$\phi_{v,u}(g_n) \to \phi_{v,u}(g)$$

This gives:

$$\phi_{u,u}(f\cdot g) = \lim_{n\to\infty} \phi_{(f(A))^* u,u}(g_n) = \phi_{(f(A))^* u,u}(g) = \langle (f(A))^* u,\Phi(g)\,u\rangle$$

$$\Rightarrow \quad \langle u,\Phi(f\cdot g)\,u\rangle = \langle u,f(A)\cdot g(A)\,u\rangle$$

Polarization yields:

$$\Phi(fg) = \Phi(f)\cdot\Phi(g)$$

Second Step: Consider $f,g \in \mathcal{B}$. We choose $f_n \in C^0$ with $f_n \to f$ and $\|f_n\| < C$. Then $f_n \cdot g \to f\cdot g$ converges point-wise.

$$\begin{aligned} \langle u,\Phi(f\cdot g)\,u\rangle &\underset{\text{convergence}}{\overset{\text{dominated}}{=}} \lim_{n\to\infty} \langle u,\Phi(f_n\cdot g)\,u\rangle \overset{\text{First step}}{=} \lim_{n\to\infty} \langle u,\Phi(f_n)\cdot\Phi(g)\,u\rangle = \\ &= \lim_{n\to\infty} \phi_{u,g(A)u}(f_n) = \phi_{u,g(A)u}(f) = \langle u,f(A)\,g(A)\,u\rangle \end{aligned}$$

$$\Rightarrow \qquad \langle u, (\Phi(fg) - \Phi(f)\,\Phi(g))\,u\rangle = 0 \qquad \underset{u\in H}{\forall}$$

By polarization follows:

$$\Phi(fg) = \Phi(f)\,\Phi(g)$$

The involution property follows similarly. $\square_{\text{i)}}$

iii) Claim: From point-wise convergence $f_n \to f$ and $\|f_n\| < C$ follows strong convergence $f_n(A) \to f(A)$.

a) From the dominated convergence theorem it is clear that holds:

$$\begin{aligned}\phi_u(f_n) &\to \phi_u(f)\\ \langle u, f_n(A)\, u\rangle &\to \langle u, f(A)\, u\rangle\end{aligned}$$

Polarization gives for all $u,v \in H$:

$$\langle u, f_n(A)\, v\rangle \to \langle u, f(A)\, v\rangle$$

In other words for all $v \in H$ holds:

$$f_n(A)\, v \rightharpoonup f(A)\, v$$

b) It holds:

$$\begin{aligned}\|f_n(A)\, v\|^2 &= \langle f_n(A)\, v, f_n(A)\, v\rangle = \langle v, (f_n(A))^* f_n(A)\, v\rangle = \\ &= \left\langle v, \overline{f_n}(A)\, f_n(A)\, v\right\rangle = \left\langle v, |f_n(A)|^2\, v\right\rangle \underset{\text{convergence}}{\overset{\text{dominated}}{\to}} \left\langle v, |f|^2(A)\, v\right\rangle = \\ &= \left\langle v, \overline{f}(A)\, f(A)\, v\right\rangle = \langle f(A)\, v, f(A)\, v\rangle = \|f(A)\, v\|^2\end{aligned}$$

c) Now apply the following general Lemma:

Lemma: $u_n \rightharpoonup u$ and $\|u_n\| \to \|u\|$ imply $u_n \to u$.

Proof:

$$\begin{aligned}\|u - u_n\| &= \langle u - u_n, u - u_n\rangle = \\ &= \|u\|^2 - 2\mathrm{Re}\underbrace{\langle u, u_n\rangle}_{\substack{\to \langle u,u\rangle \\ \text{because } u \rightharpoonup u_n}} + \underbrace{\|u_n\|^2}_{\substack{\to \|u\|^2 \\ \text{because } \|u_n\| \to \|u\|}} \to \|u\|^2 - 2\|u\|^2 + \|u\|^2 = 0\end{aligned}$$

$\square_{\text{Lemma}}$

d) This gives:

$$f_n(A)\, v \to f(A)\, v$$

$\square_{\text{i)}}$

ii) Claim: $\|f(A)\| \leq \|f\|_\infty$ for $f \in \mathcal{B}$.
Choose $f_n \in C^0$ which converge point-wise to f and $\|f_n\|_\infty < \|f\|$.

$$\|f(A)\, u\| \overset{\text{iii)}}{=} \lim_{n\to\infty} \|f_n(A)\, u\| \leq \lim_{n\to\infty} \underbrace{|||f_n(A)|||}_{=\|f_n\|_\infty} \cdot \|u\| = \lim_{n\to\infty} \|f_n\|_\infty \cdot \|u\| = \|f\|_\infty \cdot \|u\|$$

$$\Rightarrow \qquad |||f(A)||| \leq \|f\|_\infty$$

$\square_{\text{i)}}$

iv) - vi) follow immediately by approximation.

$\square_{7.3.4}$

7.3.5 Remark

So far we considered Borel measures on $\sigma(A) \subseteq \mathbb{R}$. These measures can be extended to Borel measures on $\mathbb{R}$ by defining for a Borel set $\Omega \in \mathfrak{M}(\mathbb{R})$:

$$\mu(\Omega) := \mu(\Omega \cap \sigma(A))$$

$\Omega \cap \sigma(A)$ is a Borel set of $\sigma(A)$, since $\sigma(A)$ is closed.

Now let $M \subseteq \mathfrak{M}(\mathbb{R})$ be a Borel set. $f(A)$ is well defined for any $f \in \mathcal{B}(\mathbb{R})$. With the characteristic function χ_M of M define:

$$E_M := \chi_M(A)$$

Then we get:

$$\begin{aligned} E_M^* &= \overline{\chi_M}(A) = \chi_M(A) = E_M \\ E_M^2 &= \chi_M(A) \cdot \chi_M(A) = (\chi_M \cdot \chi_M)(A) = \chi_M(A) = E_M \end{aligned}$$

Thus E_M is symmetric and idempotent, in other words E_M is a projection operator.

The mapping $M \mapsto E_M$ is the spectral measure.

7.3.6 Definition (projection operator, spectral measure)

$P \in L(H)$ is a *projection operator* if $P^2 = P = P^*$.
An operator-valued *spectral measure* E is a mapping

$$\begin{aligned} E : \mathfrak{M}(\mathbb{R}^n) &\to L(H) \\ M &\mapsto E_M := E(M) \end{aligned}$$

with the following properties:

i) E_M is a projection operator for all $M \in \mathfrak{M}$.

ii) $E_\emptyset = 0$, $E_{\mathbb{R}^n} = \mathbb{1}$

iii) For $M = \dot{\bigcup}_{n=1}^{\infty} M_n$ the operator E_M is the strong limes of the partial sums $\sum\limits_{n=1}^{k} E_{M_n}$:

$$E_M = \operatorname*{s-lim}_{k\to\infty} \sum_{n=1}^{k} E_{M_n}$$

This means that for all $u \in H$ holds:

$$E_M u = \sum_{n=1}^{\infty} (E_{M_n} u)$$

The series does not necessarily converge in the operator norm!

iv) $E_M \cdot E_N = E_{M\cap N}$

v) For all $u \in H$, the mapping $M \mapsto \langle u, E_M u\rangle \in \mathbb{R}$ is a (real) bounded regular Borel measure.

$\operatorname{supp}(E)$ is the complement of the largest open set Ω with $E_\Omega = 0$, which exists due to the σ-additivity.
E is called a *compact* spectral measure if $\operatorname{supp}(E)$ is compact.

7.3.7 Theorem

Let $A \in L(H)$ be symmetric. Then the mapping

$$E : M \mapsto \chi_M(A)$$

is a spectral measure on $\mathbb{R}$ with $\operatorname{supp}(E) \subseteq \sigma(A)$.

Proof

We have to show the properties from the definition 7.3.6.

i) is clear.

$$\chi_\emptyset(A) = 0(A) = 0$$
$$\chi_\mathbb{R}(A) = \Phi(1) = \mathbb{1}$$

So ii) is shown.

iv) follows from:

$$\chi_M(A) \cdot \chi_N(A) = (\chi_M \cdot \chi_N)(A) = \chi_{M \cap N}(A)$$

For v) consider:

$$\langle u, E_M u \rangle = \langle u, \chi_M(A) u \rangle = \phi_u(\chi_M) = \int \chi_M \mathrm{d}\mu_u = \mu_u(M)$$

It remains to show iii) and $\operatorname{supp}(E) \subseteq \sigma(A)$.

Appendix

Acknowledgements

My special thanks goes to Professor Finster, who gave this lecture and allowed me to publish this script of the lecture.

I would also like to thank all those, who found errors by careful reading and told me of them.

Andreas Völklein

GNU Free Documentation License

0. Preamble

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. Applicability and definitions

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "**Document**", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "**you**". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A “**Modified Version**” of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A “**Secondary Section**” is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document’s overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The “**Invariant Sections**” are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The “**Cover Texts**” are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A “**Transparent**” copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not “Transparent” is called “**Opaque**”.

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The “**Title Page**” means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, “Title Page” means the text near the most prominent appearance of the work’s title, preceding the beginning of the body of the text.

The “**publisher**” means any person or entity that distributes copies of the Document to the public.

A section “**Entitled XYZ**” means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as “**Acknowledgements**”, “**Dedications**”, “**Endorsements**”, or “**History**”.) To “**Preserve the Title**” of such a section when you modify the Document means that it remains a section “Entitled XYZ” according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. Verbatim copying

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. Copying in quantity

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. Modifications

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.

B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.

C. State on the Title page the name of the publisher of the Modified Version, as the publisher.

D. Preserve all the copyright notices of the Document.

E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.

F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.

G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.

H. Include an unaltered copy of this License.

I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.

J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.

K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.

L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.

M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.

N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.

O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties—for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. Combining documents

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements".

6. Collections of documents

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. Aggregation with independent works

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. Translation

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. Termination

You may not copy, modify, sublicense, or distribute the Document except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense, or distribute it is void, and will automatically terminate your rights under this License.

However, if you cease all violation of this License, then your license from a particular copyright holder is reinstated (a) provisionally, unless and until the copyright holder explicitly and finally terminates your license, and (b) permanently, if the copyright holder fails to notify you of the violation by some reasonable means prior to 60 days after the cessation.

Moreover, your license from a particular copyright holder is reinstated permanently if the copyright holder notifies you of the violation by some reasonable means, this is the first time you have received notice of violation of this License (for any work) from that copyright holder, and you cure the violation prior to 30 days after your receipt of the notice.

Termination of your rights under this section does not terminate the licenses of parties who have received copies or rights from you under this License. If your rights have been terminated and not permanently reinstated, receipt of a copy of some or all of the same material does not give you any rights to use it.

10. Future revisions of this license

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See `https://www.gnu.org/copyleft/`.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation. If the Document specifies that a proxy can decide which future versions of this License can be used, that proxy's public statement of acceptance of a version permanently authorizes you to choose that version for the Document.

11. Relicensing

"Massive Multiauthor Collaboration Site" (or "MMC Site") means any World Wide Web server that publishes copyrightable works and also provides prominent facilities for anybody to edit those works. A public wiki that anybody can edit is an example of such a server. A "Massive Multiauthor Collaboration" (or "MMC") contained in the site means any set of copyrightable works thus published on the MMC site.

"CC-BY-SA" means the Creative Commons Attribution-Share Alike 3.0 license published by Creative Commons Corporation, a not-for-profit corporation with a principal place of business in San Francisco, California, as well as future copyleft versions of that license published by that same organization.

"Incorporate" means to publish or republish a Document, in whole or in part, as part of another Document.

An MMC is "eligible for relicensing" if it is licensed under this License, and if all works that were first published under this License somewhere other than this MMC, and subsequently incorporated in whole or in part into the MMC, (1) had no cover texts or invariant sections, and (2) were thus incorporated prior to November 1, 2008.

The operator of an MMC Site may republish an MMC contained in the site under CC-BY-SA on the same site at any time before August 1, 2009, provided the MMC is eligible for relicensing.

ADDENDUM: How to use this License for your documents

To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page:

> Copyright © YEAR YOUR NAME.
> Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation;
>
> with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.
>
> A copy of the license is included in the section entitled "GNU Free Documentation License".

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts, replace the "with ... Texts." line with this:

> with the Invariant Sections being LIST THEIR TITLES, with the Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST.

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.