Funktionalanalysis

Jan Fuhrmann

Wintersemester 2023/24

Inhaltsverzeichnis

Literatur:

Zu den Themen unseres Kurses

Prinzipiell können Sie jedes Lehrbuch zur (linearen) Funktionalanalysis begleitend zur Vorlesung konsultieren. Die folgende Liste enthält eine Auswahl von Büchern, die ich selbst mehr oder weniger regelmäßig nutze.

- W. Alt, Lineare Funktionalanalysis, 6. Auflage, Springer (2012)
  (online verfügbar via Uni-Bibliothek, auch englisch: "Linear Functional Analysis")
- H. Brezis, Functional Analysis, Sobolev Spaces and Partial Differential Equations, Springer (2011)
- J. B. Conway, A Course in Functional Analysis, 2nd edition, Springer (2007)
- M. Reed, B. Simon, Methods of Modern Mathematical Physics I: Functional Analysis, Academic Press (1972)
- F. Riesz, B. Sz.-Nagy, Vorlesungen über Funktionalanalysis, 4. Auflage, Harri Deutsch (1982)
  (auch englisch "Functional Analysis", französisch "Leçons d'analyse fonctionnelle")
- D. Werner, Funktionalanalysis, 8. Auflage, Springer (2018)
  (online verfügbar via Uni-Bibliothek)
- E. Zeidler, Applied Functional Analysis – Applications to Mathematical Physics, Springer (1995)
- E. Zeidler, Applied Functional Analysis – Main Principles and Their Applications, Springer (1995)

Zum Nachschlagen von Vorkenntnissen

Zum Nachschlagen, Wiederholen oder Vertiefen der benötigten Vorkenntnisse können Sie auf eine riesige Auswahl von Lehrbüchern der Analysis und (linearen) Algebra zurück greifen. Die hier aufgeführten begleiten mich (ggf. in anderen Auflagen) zuverlässig seit meinem eigenen Studium.

- M. Artin, Algebra, Springer (1993)
  (Vektorräume und lineare Abbildungen)
- K. Königsberger, Analysis 2, 5. Auflage (2004)
  (Metrische und topologische Räume, Lebesgue-Integral)
- W. Rudin, Analysis, 5. Auflage, De Gruyter Oldenbourg (2022)
  (Metrische und topologische Räume, Lebesgue-Integral)

Notation:

Viele unserer Aussagen gelten für reelle ebenso wie für komplexe Vektorräume, und wir bezeichnen mit $\mathbb{K}$ stets einen der Körper $\mathbb{R}$ oder $\mathbb{C}$.

Für reelle Zahlen a, b bezeichnen $a \wedge b$ das Minimum und $a \vee b$ das Maximum von a und b. Diese Schreibweise erweist sich insbesondere für reellwertige Funktionen als nützlich, indem etwa $f \wedge g$ als punktweises Minimum von f und g gelesen wird.

Für eine Abbildung $f : V \to W$ zwischen Vektorräumen V und W bezeichnen wir mit $\ker f = \{v \in V : f(v) = 0_W\}$ den Kern und mit $\operatorname{im} f = \{f(v) : v \in V\}$ das Bild von f. Letzteres ergibt natürlich auch für Abbildungen zwischen allgemeinen Mengen Sinn.

Mit $(v_n)_n$ werden wir stets Folgen von Elementen v_n einer gegebenen Menge V bezeichnen. Es handelt sich also einfach um eine Kurzschreibweise für $(v_n)_{n\in\mathbb{N}} = (v_n)_{n=1}^{\infty} = (v_1, v_2, \dots)$.

Für eine beliebige Teilmenge A einer gegebenen Grundmenge Ω bezeichnen wir mit $\mathbf{1}_A : \Omega \to \mathbb{R}$ die durch

$$\mathbf{1}_A(x) := \begin{cases} 1 & \text{für } x \in A, \\ 0 & \text{für } x \in \Omega \backslash A \end{cases}$$

gegebene Indikatorfunktion[1].

Sind V ein $\mathbb{K}$-Vektorraum, $v \in V$, $A \subset V$ und $\alpha \in K$, so schreiben wir

$$v + A := \{v + w : w \in A\}, \quad \alpha A := \{\alpha w : w \in A\}$$

für die Translation von A um v bzw. die Multiplikation von A mit α.
Ferner bezeichnen wir mit 0_V den Nullvektor im Vektorraum V, lassen das Subskript V aber auch gelegentlich weg, wenn aus dem Kontext klar ist, welches Nullelement gemeint ist.

Ist V ein $\mathbb{K}$-Vektorraum und $A \subset V$ eine nicht leere Teilmenge, so bezeichnen wir mit

$$< A >= \left\{ \sum_{j=1}^{n} \alpha_j v_j : n \in \mathbb{N},\ \alpha_1, \dots, \alpha_n \in \mathbb{K},\ v_1, \dots, v_n \in A \right\}$$

die lineare Hülle von A. Man beachte, dass nur endliche Linearkombinationen von Elementen von A zugelassen sind.

[1] In der Analysis finden Sie dafür oft auch die Bezeichnung charakteristische Funktion unter dem Namen χ_A.

KAPITEL 1

Grundlagen: Topologie normierter Vektorräume

In dieser Vorlesung befassen wir uns mit linearer Funktionalanalysis, also vor allem mit linearen Räumen und Abbildungen zwischen diesen. Aus der Analysis 2 wissen wir, dass lineare Abbildungen $f : \mathbb{R}^n \to \mathbb{R}^m$ bzw. $f : \mathbb{C}^n \to \mathbb{C}^m$ notwendigerweise bezüglich der euklidischen Topologie stetig und sogar unter jeder Norm (beliebig oft) differenzierbar sind. Im Kern werden wir uns mit der Verallgemeinerung dieser Beobachtung auf den unendlichdimensionalen Fall beschäftigen, werden aber schnell sehen, dass bereits die Stetigkeit nicht mehr notwendigerweise garantiert ist. Auch andere uns bekannte Aussagen lassen sich nicht direkt verallgemeinern. Folgende aus dem endlichdimensionalen Fall bekannte Aussagen werden wir genauer untersuchen müssen:

- Auf einem endlichdimensionalen $\mathbb{K}$-Vektorraum V erzeugen alle Normen die gleiche Topologie.

- In einem endlichdimensionalen normierten Vektorraum sind abgeschlossene und beschränkte Mengen kompakt[1]. Insbesondere sind abgeschlossene Einheitskugeln kompakt.

- Lineare Unterräume eines endlichdimensionalen normierten Vektorraums sind abgeschlossen (bzgl. der von der Norm erzeugten Topologie).

- Eine lineare Abbildung $f : \mathbb{K}^n \to \mathbb{K}^n$ ist genau dann injektiv, wenn sie surjektiv ist. Genauer gilt für lineare Abbildungen $f : V \to V$ die Dimensionsformel

$$\dim_{\mathbb{K}} V = \dim_{\mathbb{K}}(\ker f) + \dim_{\mathbb{K}}(\operatorname{im} f).$$

[1] Für die euklidischen n-dimensionalen Räume $\mathbb{K}^n$ ist das der Satz von Heine-Borel. Nach dem vorherigen Punkt gilt das dann auch unter allen Normen auf $\mathbb{K}^n$.

In gewissem Sinn mag das zwar noch richtig sein, wenn wir auf beiden Seiten ∞ als Wert zulassen, aber die Aussagekraft für einen der Summanden auf der rechten Seite ist recht gering, wenn der andere Summand und die linke Seite unendlich sind.

1.1 Definition und grundlegende Eigenschaften

Wir beginnen mit einer Erinnerung an das Konzept normierter linearer Räume. Dabei handelt es sich um eine spezielle Klasse metrischer Räume, die Sie in der Analysis 2 kennen gelernt haben. Diese wiederum hatten wir als spezielle topologische Räume identifiziert. Viele der hier angegebenen Definitionen und Aussagen sollten Ihnen also bekannt vorkommen. Falls das nicht der Fall ist, können Sie gern in den einschlägigen Lehrbüchern zur Analysis 2 oder elementarer Topologie nachschlagen.

Definition 1.1. *Ein normierter linearer Raum (oder normierter Vektorraum) ist ein $\mathbb{K}$-Vektorraum V zusammen mit einer Norm $\|\cdot\| : V \to [0, \infty)$ derart, dass für alle $v, w \in V$ und $\alpha \in \mathbb{K}$*

(i) $v = 0_V \iff \|v\| = 0$ *(Definitheit)*

(ii) $\|\alpha v\| = |\alpha|\|v\|$ *(Homogenität)*

(iii) $\|v + w\| \leq \|v\| + \|w\|$ *(Dreiecksungleichung)*

gelten.

Da Normen nur auf Vektorräumen Sinn ergeben (wir brauchen die Addition in V inklusive der Null und die Multiplikation mit Skalaren), sprechen wir manchmal kurz von normierten Räumen und wissen dann sofort, dass es sich um normierte *lineare* Räume handeln muss.

Bemerkung. *In der Analysis 2 haben wir gelernt, dass jeder normierte lineare Raum $(V, \|\cdot\|)$ vermöge*

$$d(v, w) = \|v - w\|, \qquad v, w \in V,$$

auch ein metrischer Raum ist. Umgekehrt hatten wir aber gesehen, dass auf jeder nicht leeren Menge etwa die diskrete Metrik definiert werden konnte, auch wenn gar keine lineare Struktur vorlag.

Bemerkung. *Ist in Punkt (i) nur die Implikation $v = 0 \implies \|v\| = 0$ erfüllt, kann $\|v\|$ also auch für nicht triviale Vektoren verschwinden, so heißt $\|\cdot\|$ eine Halbnorm (oder Seminorm). Die Menge $N = \{v \in V : \|v\| = 0\}$ ist dann wegen der Homogenität und der Dreiecksungleichung ein Vektorraum, und wir erhalten durch Faktorisierung einen Vektorraum V/N, auf dem $\|\cdot\|$ zu einer Norm wird[2]. Sie kennen diese Konstruktion aus der Maßtheorie von der Konstruktion der Räume $L_p(\mu)$.*

[2] Genauer ist für jede Restklasse $v + N$ durch $\|v + N\| = \inf_{w \in N} \|v + w\|_V$ eine Norm definiert.

Die üblichen Verdächtigen $\mathbb{K}^n$ zusammen mit den p-Normen

$$\|v\|_p = \begin{cases} \left(\sum_{k=1}^n |v_k|^p\right)^{\frac{1}{p}} & \text{für } p \in [1,\infty), \\ \max\limits_{k=1,\dots,n} |v_k| & \text{für } p = \infty \end{cases}$$

kennen wir schon. Das verallgemeinern wir schnell auf die natürlichen unendlichdimensionalen Versionen.

Beispiel 1. *Mit $\mathbb{K}^{\mathbb{N}}$ bezeichnen wir den Vektorraum der $\mathbb{K}$-wertigen Zahlenfolgen $v = (v_k)_{k\geq 1}$. Wir sehen sofort, dass $\mathbb{K}^n$ mit dem linearen Unterraum der spätestens nach n Gliedern abbrechenden Folgen identifizierbar ist. Jedes $x = (x_1, \dots, x_n)^T \in \mathbb{K}^n$ liefert nämlich eine eindeutige Folge $v \in \mathbb{K}^{\mathbb{N}}$:*

$$v_k = \begin{cases} x_k & \textit{für } k = 1, \dots, n, \\ 0 & \textit{für } k > n. \end{cases}$$

Dass Summen und skalare Vielfache solcher abbrechenden Folgen wieder spätestens bei n abbrechen, ist offensichtlich, also handelt es sich tatsächlich um einen linearen Unterraum.
Auf $\mathbb{K}^{\mathbb{N}}$ definieren wir analog zu den bekannten p-Normen

$$\|v\|_p := \begin{cases} \left(\sum_{k=1}^\infty |v_k|^p\right)^{\frac{1}{p}} & \textit{für } p \in [1,\infty), \\ \sup\limits_{k\in\mathbb{N}} |v_k| & \textit{für } p = \infty \end{cases}$$

und stellen sofort fest, dass wir im allgemeinen nicht davon ausgehen können, dass $\|v\|_p \in [0,\infty)$ gilt, wie es sich für eine Norm gehört. Das führt uns zur Definition der Räume der zur pten Potenz (absolut) summierbaren (bzw. im Fall $p = \infty$: der beschränkten) Folgen:

$$\ell_p^{\mathbb{K}} := \left\{ v \in \mathbb{K}^{\mathbb{N}} : \|v\|_p < \infty \right\}.$$

Versehen mit $\|\cdot\|_p$ wird $\ell_p^{\mathbb{K}}$ zu einem normierten Vektorraum. Wir rechnen das für den Fall $p = \infty$ nach, für $p < \infty$ können Sie das als Übung unter Verwendung der Rechenregeln für absolut konvergente Reihen und einiger Eigenschaften der Potenzfunktionen selbst überprüfen.
Die Endlichkeit der Norm haben wir durch die Definition von $\ell_\infty^{\mathbb{K}}$ erzwungen, die müssen wir also nicht mehr untersuchen.
Zunächst ist die triviale Folge $0 \in \mathbb{K}^{\mathbb{N}}$ natürlich beschränkt, also ist $0 \in \ell_\infty$ (wir lassen das $\mathbb{K}$ oft weg, wenn entweder klar oder unerheblich ist, welcher Körper gemeint ist), und es gilt $\|0\|_\infty = \sup_k |0| = 0$. Ist umgekehrt $v \neq 0$, so existiert ein k_0 mit $v_{k_0} \neq 0$, und dann ist $\|v\|_\infty \geq |v_{k_0}| > 0$. Das zeigt die Definitheit.
Für $v \in \ell_\infty$ und $k \in \mathbb{N}$ gilt $|v_k| \leq \|v\|_\infty$. Ist nun $\alpha \in \mathbb{K}$, so folgt $|(\alpha v)_k| = |\alpha||v_k| \leq |\alpha|\|v\|_\infty$, also $\|\alpha v\| \leq |\alpha|\|v\|_\infty$ und insbesondere $\alpha v \in \ell^\infty$. Umgekehrt führen wir die Annahme $\|\alpha v\|_\infty < |\alpha|\|v\|_\infty$, d.h., $\|\alpha v\|_\infty = |\alpha|\|v\|_\infty - \delta$ für ein $\delta > 0$ zum Widerspruch.

Wäre $\alpha = 0$*, so stünde hier* $0 = 0 - \delta$*, was völlig unmöglich ist. Ist aber* $|\alpha| > 0$*, so fänden wir nach der Definition des Supremums ein* k_0 *derart, dass*

$$|v_{k_0}| \geq \|v\|_\infty - \frac{\delta}{2|\alpha|}, \quad \textit{also} \quad \|v\|_\infty \leq |v_{k_0}| + \frac{\delta}{2|\alpha|}$$

wäre. Damit berechnen wir nach unserer Annahme:

$$|\alpha||v_{k_0}| \leq |\alpha|\|v\|_\infty - \delta \leq |\alpha| \left(|v_{k_0|+\frac{\delta}{2|\alpha|}} \right) - \delta = |\alpha||v_{k_0}| - \frac{\delta}{2},$$

was auch ein Widerspruch ist. Damit haben wir auch die Homogenität gezeigt.
Zur Dreiecksungleichung stellen wir fest, dass für jedes $k \in \mathbb{N}$

$$|(v+w)_k| = |v_k + w_k| \leq |v_k| + |w_k| \leq \|v\|_\infty + \|w\|_\infty.$$

Nehmen wir das Supremum über alle k*, so erhalten wir sofort die Dreiecksungleichung, die uns auch nochmal sichert, dass die Summe zweier beschränkter Folgen wieder beschränkt ist.*
Wir haben also gezeigt, dass ℓ_∞ *tatsächlich ein Vektorraum ist und* $\|\cdot\|_\infty$ *eine Norm darauf darstellt.*

Bemerkung. *Die Konstruktion der Räume* ℓ_p *folgt einem allgemeinen Konzept. Ist auf einem Vektorraum* V *eine Abbildung* $\|\cdot\| : V \to [0, \infty]$ *definiert, die bis auf die Endlichkeit alle Eigenschaften einer Norm hat, so nennen wir* $\|\cdot\|$ *eine Quasinorm und stellen fest, dass wegen der Homogenität und der Dreiecksungleichung* $W = \{v \in V : \|v\| < \infty\}$ *ein Untervektorraum von* V *ist, auf dem* $\|\cdot\|$ *eine Norm ist.*

Man beachte den Unterschied zum endlichdimensionalen Fall. Bei weitem nicht alle Folgen $v \in \mathbb{K}^{\mathbb{N}}$ gehören zu einem der Räume ℓ_p, und die Räume enthalten nicht die gleichen Folgen als Elemente[3]. Schließlich sind unendliche Summen bzw. Suprema unendlicher Mengen nicht zwingend endlich. Das ist auch schon die Grundlage der in den einführenden Bemerkungen erwähnten scheiternden Verallgemeinerungen von Aussagen aus der endlichdimensionalen linearen Algebra bzw. Analysis 2.

Bemerkung. *In Erinnerung an die Maßtheorie (aus der höheren Analysis oder Wahrscheinlichkeitstheorie) stellen wir fest, dass es sich bei den* ℓ_p*-Räumen um Spezialfälle der* $L_p(\mu)$*-Räume handelt, den Räumen (von Äquivalenzklassen) jener messbaren Funktionen* $f : \Omega \to \mathbb{K}$*, für die*

$$\|f\|_p := \begin{cases} \left(\int_\Omega |f|^p \,\mathrm{d}\mu\right)^{\frac{1}{p}} & \textit{für } p \in [1, \infty), \\ \operatorname*{ess\,sup}_{x \in \Omega} |f(x)| & \textit{für } p = \infty \end{cases}$$

endlich ist. Die Räume ℓ_p *hatten wir als Spezialfälle für* $\Omega = \mathbb{N}$*, versehen mit der Potenzmenge als* σ*-Algebra und dem Zählmaß als Maß erkannt. Die Folgenräume haben die*

[3]Sie rechnen leicht nach, dass $\ell_p \subsetneq \ell_q$ für $p < q$ gilt. Denken Sie etwa an eine durch $v_k = k^{-r}$ gegebene Folge, wobei $r > 0$ so gewählt ist, dass $rp < 1 < rq$ gilt.

angenehme Eigenschaft, dass man auf die Identifizierung fast überall übereinstimmender Funktionen verzichten kann, da es bezüglich des Zählmaßes keine nicht trivialen Nullmengen gibt. Jede der Äquivalenzklassen besteht also aus genau einem Repräsentanten. In dieser Vorlesung werden wir uns nicht explizit mit Maß- und Integrationstheorie beschäftigen, aber in vielen Situationen auf das Lebesgue-Integral zurück greifen. Neben den Folgenräumen werden diese L_p-Räume für das Lebesgue-Maß auf einer offenen Menge $\Omega \subset \mathbb{R}^n$ sowie deren Abkömmlinge eine wichtige Rolle spielen. Die nötigen Aussagen werden wir im Anhang **??** *sammeln; dort können Sie jederzeit nachschlagen.*

Auch hier haben wir uns wieder auf lineare Unterräume des ursprünglichen Raums messbarer Funktionen (genauer: Äquivalenzklassen solcher) zu beschränken, auf denen die Normen endlich sind.

Notation. Wenn wir vom normierten Vektorraum ℓ_p oder $L_p(\mu)$ bzw. $L_p(\Omega)$ sprechen, meinen wir – sofern nicht ausdrücklich anders gesagt – den mit der natürlichen Norm $\|\cdot\|_p$ versehenen Raum. Die Norm gehört also, anders als im endlichdimensionalen Fall, praktisch zum Raum selbst.

Vom üblichen Betrag in $\mathbb{R}$ kennen wir folgende Aussage, deren Beweis wir aus der Analysis 1 praktisch wortgleich übernehmen können.

Lemma 1.2 (Dreiecksungleichung nach unten)**.** *Ist $(V, \|\cdot\|)$ ein normierter Vektorraum, so gilt $|\|v\| - \|w\|| \leq \|v - w\|$ für alle $v, w \in V$.*

Beweis. Übung (siehe Analysis 1). □

Folgende Definition dürfte eine Wiederholung sein, aber der Vollständigkeit halber müssen wir sie hier noch einmal angeben.

Definition 1.3. *In einem normierten Vektorraum $(V, \|\cdot\|)$ definieren wir zu $v \in V$ und $r > 0$ die offene Kugel vom Radius r um v durch*

$$B_r(v) = \{w \in V : \|v - w\| < r\},$$

und die abgeschlossene Kugel vom Radius r um v durch

$$\bar{B}_r(v) = \{w \in V : \|v - w\| \leq r\}.$$

Die offene bzw. abgeschlossene Einheitskugel in V sind durch

$$B_V = B_1(0_V) \qquad \text{bzw.} \qquad \bar{B}_V = \bar{B}_1(0_V)$$

gegeben.

Wie in jedem metrischen Raum[4] erlauben uns die offenen Kugeln die Definition offener Mengen.

[4]In einem metrischen Raum wird eine offene Kugel durch $B_r(v) = \{w \in V : d(v, w) < r\}$ definiert.

Definition 1.4. *Eine Teilmenge $U \subset V$ eines normierten Vektorraums $(V, \|\cdot\|)$ heißt offen, falls zu jedem $v \in U$ ein $r > 0$ derart existiert, dass $B_r(v) \subset U$ gilt.*
Die Familie $\mathcal{T}$ der offenen Mengen ist die von $\|\cdot\|$ erzeugte Topologie über V.

Manchmal schreiben wir $\mathcal{T}_{\|\cdot\|}$ für die von $\|\cdot\|$ erzeugte Topologie. Das wird spätestens dann relevant, wenn wir verschiedene Normen oder auch gar nicht von einer Norm erzeugte Topologien auf einem gegebenen Vektorraum betrachten.
Dass es sich bei den so definierten Systemen offener Mengen tatsächlich um Topologien handelt, besagt das folgende Lemma, dessen Aussagen gerade die definierenden Eigenschaften eines Systems offener Mengen (also einer Topologie) sind.

Lemma 1.5. *Für die von $\|\cdot\|$ erzeugte Topologie $\mathcal{T}$ über einem normierten Vektorraum V gelten:*

(i) $\emptyset \in \mathcal{T}$, $V \in \mathcal{T}$ *(die leere Menge und der gesamte Raum sind stets offen)*

(ii) *Ist $\mathcal{I}$ eine beliebige (nicht leere) Indexmenge und sind $U_i \in \mathcal{T}$, $i \in \mathcal{I}$, so ist $\bigcup_{i\in\mathcal{I}} U_i \in \mathcal{T}$ (beliebige Vereinigungen offener Mengen sind offen)*

(iii) *Sind $U_1, U_2 \in \mathcal{T}$, so ist $U_1 \cap U_2 \in \mathcal{T}$ (endliche Durchschnitte offener Mengen sind offen)*

Beweis. Übungsblatt 0 □

Selbstverständlich sind (nach Dreiecksungleichung) alle offenen Kugeln offen (also in $\mathcal{T}$), wie Sie sich schnell überzeugen. Wir werden gleich sehen, dass auch der Begriff der *abgeschlossenen* Kugel durchaus sinnvoll gewählt ist.

Definition 1.6. *Die Norm $\|\cdot\|$ auf dem Vektorraum V heißt stärker als die Norm $|[\cdot]|$ auf V, wenn die von $|[\cdot]|$ erzeugte Topologie in der von $\|\cdot\|$ erzeugten enthalten ist.*
Zwei Normen $\|\cdot\|$, $|[\cdot]|$ auf dem gleichen Vektorraum heißen äquivalent, wenn sie die gleiche Topologie erzeugen.

Zwei Normen sind also äquivalent, wenn jede der beiden stärker als die andere ist.

Lemma 1.7. *Zwei Normen $\|\cdot\|$, $|[\cdot]|$ auf einem Vektorraum sind genau dann äquivalent, wenn Konstanten $C \geq c > 0$ derart existieren, dass für alle $v \in V$*

$$c\|v\| \leq |[v]| \leq C\|v\| \tag{1.1}$$

gilt.

Beweis. Zum Beweis der Äquivalenz der Normen genügt es zu zeigen, dass in jeder offenen Kugel bezüglich $\|\cdot\|$ eine offene Kugel bezüglich $|[\cdot]|$ liegt und umgekehrt. Dann finden wir nämlich zu einer Menge U und einem $v \in U$ stets genau dann eine in U liegende ε-Kugel bezüglich $\|\cdot\|$ um v, wenn wir eine solche bezüglich $|[\cdot]|$ finden.

Zu $\varepsilon > 0$ und $v \in V$ gilt

$$B_{\varepsilon,|[\cdot]|}(v) = \{w : |[v - w]| < \varepsilon\} \subset \{w : c\|v - w\| < \varepsilon\} = B_{\frac{\varepsilon}{c},\|\cdot\|}(v)$$

und umgekehrt $B_{\varepsilon,\|\cdot\|}(v) \subset B_{C\varepsilon,|[\cdot]|}(v)$, wie verlangt.

Zur Notwendigkeit der Bedingung nehmen wir an, wir hätten zu jedem $C_n = n$ ein $v_n \in V$ derart, dass $|[v_n]| > n\|v_n\|$ ist, dass also die rechte Ungleichung nicht gilt. Dank der Homogenität beider Normen gilt dann auch für $w_n := \frac{v_n}{\|v_n\|}$ (man beachte, dass die v_n notwendigerweise von 0 verschieden sind):

$$|[w_n]| > n \left\| \frac{v_n}{\|v_n\|} \right\| = n.$$

Die Einheitskugel $U = B_{1,|[\cdot]|}(0)$ ist offen in $(V, |[\cdot]|)$, also nach Annahme der Äquivalenz auch in $(V, \|\cdot\|)$. Zu $0 \in U$ müssten wir also ein $\varepsilon > 0$ derart finden, dass $B_{\varepsilon,\|\cdot\|}(0) \subset U$ gilt. In $B_{\varepsilon,\|\cdot\|}(0)$ liegen nun aber alle $\frac{\varepsilon}{2}w_n$, also gilt für alle n:

$$\frac{\varepsilon}{2}w_n \in B_{\varepsilon,\|\cdot\|}(0) \subset U = B_{1,|[\cdot]|}(0), \quad \text{also} \quad \frac{\varepsilon}{2}|[w_n]| \leq 1.$$

Das steht aber für $n > \frac{2}{\varepsilon}$ im Widerspruch zu $|[w_n]| > n$.
Völlig symmetrisch führen wir auch die Annahme, die linke Ungleichung gälte nicht, zu einem Widerspruch. $\square$

In diesem Beweis haben wir einige sehr mächtige Tricks benutzt, die Beweise zu normierten Vektorräumen oft erleichtern. Zum einen haben wir uns im zweiten Teil auf Kugeln um den Ursprung beschränkt. Das funktioniert, weil V dank der Vektorraumstruktur überall lokal gleich aussieht. Die Kugel $B_\varepsilon(0)$ unterscheidet sich also nicht von $B_\varepsilon(v)$. Tatsächlich liegt ja w genau dann in $B_\varepsilon(v)$, wenn $w - v$ in $B_\varepsilon(0)$ liegt. Zweitens haben wir benutzt, dass Kugeln eines gegebenen Radius wegen der Homogenität einfach durch Aufblähen oder Schrumpfen aus solchen mit Radius 1 hervorgehen: $v \in B_1(0) \iff \varepsilon v \in B_\varepsilon(0)$, d.h.:

$$B_\varepsilon(w) = \{w + \varepsilon v : v \in B_1(0)\}.$$

Das Studium der Einheitskugel eines normierten Vektorraums verrät uns also schon alles über, was wir über Kugeln in diesem Raum wissen müssen, insbesondere charakterisiert die Einheitskugel die Norm und die von ihr erzeugte Topologie vollständig. Da Kugeln auch Umgebungen eindeutig charakterisieren, erhalten wir insbesondere folgende Aussage.

Korollar 1.8. *Die Normen $\|\cdot\|$ und $|[\cdot]|$ auf dem Vektorraum V sind genau dann äquivalent, wenn für jede Folge $(v_n)_n$ in V gilt:*

$$\lim_{n\to\infty} \|v_n\| = 0 \quad \iff \quad \lim_{n\to\infty} |[v_n]| = 0.$$

Beweis. Übung.
Man erinnere sich daran, dass Konvergenz bzgl. einer Topologie über V bedeutet:

$$v_n \to v \quad :\Longleftrightarrow \quad \text{Für jede Umgebung } U \text{ von } v \text{ ex. } n_0 \in \mathbb{N} \text{ mit } v_n \in U \text{ f.a. } n \geq n_0.$$

Eine Umgebung von v ist im topologischen Sinn einfach eine Menge U, die eine offene Menge W mit $v \in W \subset U$ enthält. In normierten Räumen können wir uns um kugelförmige Umgebungen beschränken, also:

$$v_n \to v \iff \text{für alle } \varepsilon > 0 \text{ ex. } n_0 \in \mathbb{N} \text{ mit } v_n \in B_\varepsilon(v) \text{ für alle } n \geq n_0.$$

Äquivalente Normen (bzw. Metriken) ergeben also für genau die gleichen Folgen Konvergenz (mit gleichem Grenzwert)[5]. □

Da wir normierte Vektorräume als spezielle topologische Räume identifiziert haben, können wir uns über Stetigkeit unterhalten. Dabei benutzen wir ohne Beweis, dass in normierten Vektorräumen wie in allen metrischen Räumen ε-δ-Stetigkeit (wie auch Folgenstetigkeit, aber dafür brauchen wir noch Grenzwerte) äquivalent zur topologischen Definition (Urbilder offener Mengen unter stetigen Funktionen sind offen) ist.

Korollar 1.9. *Die Norm auf einem normierten Vektorraum ist eine stetige Abbildung $\|\cdot\| : V \to \mathbb{R}$, wenn wir $\mathbb{R}$ mit seinem Standardbetrag versehen.*

Beweis. Das ist gerade die Aussage der Dreiecksungleichung nach unten. □

Ebenso schnell rechnen wir folgende Aussage nach.

Lemma 1.10. *Ist $(V, \|\cdot\|)$ ein normierter $\mathbb{K}$-Vektorraum, so sind die Vektorraumoperationen*

$$(v, w) \mapsto v + w \quad (V \times V \to V), \qquad (\alpha, v) \mapsto \alpha v \quad (\mathbb{K} \times V \to V)$$

stetig bezüglich $\|\cdot\|$.

Beweis. Zu gegebenem $\varepsilon > 0$ gilt für $v, w, \tilde{v}, \tilde{w} \in V$ mit $\|v - \tilde{v}\|, \|w - \tilde{w}\| < \delta = \frac{\varepsilon}{2}$:

$$|\|\tilde{w} + \tilde{v}\| - \|w + v\|| \leq \|(\tilde{w} + \tilde{v}) - (w + v)\| \leq \|\tilde{w} - w\| + \|\tilde{v} - v\| < 2\delta = \varepsilon.$$

Die Addition ist also sogar gleichmäßig stetig. Analog gilt zu gegebenen $\alpha \in \mathbb{K}$, $v \in V$ und $\varepsilon > 0$ für $\tilde{\alpha} \in \mathbb{K}$, $\tilde{v} \in V$ mit

$$\|\tilde{v} - v\|, |\tilde{\alpha} - \alpha| < \delta = \frac{\varepsilon}{2(1 + |\alpha| + \|v\|)} \wedge \frac{\sqrt{\varepsilon}}{2} :$$

[5] Kleine Warnung: In allgemeinen topologischen Räumen (ohne Metrik) kann es vorkommen, dass verschiedene Topologien die gleichen konvergenten Teilfolgen liefern.

$$\begin{aligned}
\|\tilde{\alpha}\tilde{v} - \alpha v\| &= \|(\tilde{\alpha}\tilde{v} - \tilde{\alpha}v) + (\tilde{\alpha}v - \alpha v)\| \\
&\leq \|\tilde{\alpha}(\tilde{v} - v)\| + \|(\tilde{\alpha} - \alpha)v\| \\
&= |\tilde{\alpha} - \alpha + \alpha|\|\tilde{v} - v\| + |\tilde{\alpha} - \alpha|\|v\| \\
&< (\delta + |\alpha|)\delta + \delta\|v\| \; < \; \delta^2 + \delta(1 + |\alpha| + \|v\|) \\
&\leq \frac{\varepsilon}{4} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

□

Die Aussage dieses Lemmas bedeutet schlicht, dass V unter der von der Norm erzeugten Topologie ein topologischer Vektorraum ist.
Wir erinnern auch noch an die Definition abgeschlossener Mengen. In einem topologischen Raum $(X, \mathcal{T})$ heißt eine Teilmenge $A \subset X$ abgeschlossen, wenn ihr Komplement $X \backslash A$ offen ist. In metrischen Räumen und erst recht in normierten Vektorräumen haben wir eine oft deutlich leichter zu überprüfende Bedingung. Dazu müssen wir noch die Konzepte konvergenter Folgen wiederholen, die wir bereits angedeutet haben.

Definition 1.11. *Eine Folge $(v_n)_n$ in einem metrischen Raum (V, d) heißt Cauchyfolge, falls für alle $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ derart existiert, dass $d(v_n, v_m) < \varepsilon$ für alle $n, m \geq n_0$ gilt. Die Folge $(v_n)_n$ heißt konvergent, falls ein $v \in V$ derart existiert, dass $\lim_{n\to\infty} d(v, v_n) = 0$ gilt. In diesem Fall heißt v der*[a] *Grenzwert der Folge, und wir schreiben $\lim_{n\to\infty} v_n = v$ oder $v_n \to v$ für $n \to \infty$.*

[a]Wir erinnern uns, dass eine Folge in einem metrischen Raum höchstens einen Grenzwert haben kann.

An dieser Stelle ist eine kleine Warnung angebracht. Haben wir auf einer gegebenen Menge verschiedene Metriken, so muss die Konvergenz bezüglich einer Metrik keineswegs die bezüglich der anderen implizieren. Wir hatten bereits festgestellt: Sollten zwei Metriken (bzw. im Fall normierter linearer Räume zwei Normen) für genau die gleichen Folgen Konvergenz (gegen die gleichen Grenzwerte) liefern, so bedeutet das, dass sie die gleiche Topologie erzeugen, also äquivalent sind. Diese Aussage formulieren wir nochmal in anderer Form als ein Lemma, das ebenfalls nur der Wiederholung dient.

Lemma 1.12. *Eine Teilmenge $A \subset V$ eines metrischen Raums (V, d) ist genau dann abgeschlossen, wenn für jede konvergente Folge $(v_n)_n$ in A der Grenzwert $v = \lim_{n\to\infty} v_n$ ebenfalls in A liegt.*

Das liefert auch eine geeignete Definition des Abschlusses von Teilmengen metrischer Räume.

Definition 1.13. *Ist $M \subset V$ Teilmenge des metrischen Raums (V, d), so bezeichnen wir den Abschluss $\bar{M}$ von M (in (V, d)) als die Menge aller Grenzwerte von Folgen in M:*

$$\bar{M} = \{v \in V : \lim_{n\to\infty} v_n = v \;\; \text{für eine Folge } (v_n)_n \text{ mit } v_n \in M \text{ für alle } n\}.$$

Selbstverständlich ist $\bar{M}$ damit abgeschlossen[6], und es gilt $M \subset \bar{M}$ (man betrachte konstante Folgen). Genauer ist $\bar{M}$ die nach Mengeninklusion kleinste abgeschlossene Teilmenge von V, die M enthält[7]:

$$\bar{M} = \bigcap_{M \subset A \subset V, A \text{ abgeschl.}} A.$$

Eine weitere nützliche Charakterisierung des Abschlusses nähert sich dem Problem über kleine Abstände. Der Abschluss einer Menge ist die Menge aller Berührungsspunkte von M. Das sind all jene Punkte, für die jede (ε-)Umgebung mindestens einen Punkt aus M enthält.

Lemma 1.14. *Der Abschluss $\bar{M}$ einer Teilmenge $M \subset V$ eines metrischen Raums (V, d) ist durch*

$$\bar{M} = \{v \in V : \text{ Für alle } \varepsilon > 0 \text{ existiert ein } w \in M \text{ mit } d(v, w) < \varepsilon.\}$$

gegeben.

Beweis. Übungsblatt 0 □

Auch hier haben wir wieder gute Nachrichten mit Blick auf normierte Vektorräume.

Lemma 1.15. (i) *Sind U ein Untervektorraum eines Vektorraums V und $\|\cdot\|$ eine Norm auf V, so ist auch $\bar{U}$ ein Untervektorraum von V.*
(ii) *In einem normierten Vektorraum $(V, \|\cdot\|)$ gilt für alle $v \in V$ und $\varepsilon > 0$:*

$$\bar{B}_\varepsilon(v) = \overline{B_\varepsilon(v)}.$$

Beweis. Übung (man überprüfe für (i), dass $\bar{U}$ unter den Vektorraumoperationen abgeschlossen ist). □

Bemerkung. *Eine beliebige Menge M mit mindestens zwei Elementen, versehen mit der diskreten Metrik, zeigt sofort, dass Aussage (ii) in allgemeinen metrischen Räumen nicht gelten muss.*

1.2 Vollständigkeit und Kompaktheit

Dank der Dreiecksungleichung ist natürlich jede konvergente Folge in einem metrischen Raum eine Cauchyfolge. Wir hatten auch schon eine Bezeichnung für jene metrischen Räume eingeführt, in denen die Umkehrung gilt.

[6] Machen Sie sich das klar. Benutzen Sie, dass Sie für jede Folge $(v_n)_n$ in $\bar{M}$, die gegen ein $v \in V$ konvergiert, zu jedem n eine Folge in M finden, die ihrerseits gegen v_n konvergiert. Aus je einem Element jeder dieser Folgen generieren Sie eine Folge in M, die gegen v konvergiert.

[7] Unsere Definition stimmt also mit der Definition des Abschlusses für allgemeine topologische Räume überein.

Definition 1.16. *Ein metrischer Raum (V, d) heißt vollständig, falls jede Cauchyfolge in V (gegen einen in V liegenden Grenzwert) konvergiert.*
Ein vollständiger normierter Vektorraum heißt Banachraum.

Beispiel 2. *Sie überzeugen sich schnell, dass $\mathbb{R}^n$ und $\mathbb{C}^n$ unter allen p-Normen $\|\cdot\|_p$, $p \in [1, \infty]$, zu Banachräumen werden. Kraft Isomorphie ist dann jeder endlichdimensionale normierte $\mathbb{K}$-Vektorraum ein Banachraum.*

Mit der durch $d(x, y) = |x - y|$ gegebenen Metrik ist der Raum $\mathbb{Q}$ der rationalen Zahlen hingegen nicht vollständig. Denken Sie etwa an die durch $x_n = (1 + \frac{1}{n})^n$ gegebene Cauchyfolge. Das erklärt auch, warum wir nur $\mathbb{R}$- und $\mathbb{C}$-Vektorräume betrachten. Über nicht vollständigen Körpern würde es uns schwer fallen, vollständige Vektorräume zu finden.

Beispiel 3. *Erfreulicherweise sind die Räume $\ell_p^{\mathbb{K}}$ unter ihren natürlichen Normen $\|\cdot\|_p$ ($p \in [1, \infty]$) wie ihre endlichdimensionalen Geschwister auch Banachräume. Wir zeigen das hier für den Fall $p = \infty$, der Fall $p \in [1, \infty)$ ist eine (konzeptionell nicht schwierige, aber für $p > 1$ recht rechenintensive) Übung.*
Angenommen, wir hätten eine $\|\cdot\|_\infty$-Cauchyfolge $(v^{(k)})_k$ beschränkter Folgen. Zu gegebenem $\varepsilon > 0$ finden wir also ein k_0 derart, dass für alle $k, l \geq k_0$ und alle $n \in \mathbb{N}$

$$|v_n^{(k)} - v_n^{(l)}| \leq \|v^{(k)} - v^{(l)}\| < \varepsilon$$

gilt. Zu festem n ist damit $(v_n^{(k)})_k$ eine Cauchyfolge in $\mathbb{K}$, die einen Grenzwert v_n hat. Die Folge $v = (v_n)_n$ dieser Grenzwerte ist dann beschränkt und erfüllt $\|v^{(k)} - v\|_\infty \to 0$ für $k \to \infty$. Ist nämlich k_0 zu gegebenem $\varepsilon > 0$ wie oben gewählt, so gilt für $k \geq k_0$ und alle $n \in \mathbb{N}$:

$$|v_n^{(k)} - v_n| = \lim_{l \to \infty} \underbrace{|v_n^{(k)} - v_n^{(l)}|}_{<\varepsilon \text{ f. } l \geq k_0} \leq \varepsilon.$$

Das gilt für alle n:

$$\|v^{(k)} - v\|_\infty = \sup_n |v_n^{(k)} - v_n| \leq \varepsilon \qquad \text{für } k \geq k_0.$$

Allgemeiner sind die Räume $L_p(\mathbb{R}^n)$ (oder noch allgemeiner $L_p(\mu)$) der zur $p$ten Potenz integrierbaren ($p \in [1, \infty)$) bzw. wesentlich beschränkten ($p = \infty$) (Äquivalenzklassen von) Funktionen ebenfalls Banachräume. Das haben wir in der Maßtheorie gezeigt (siehe Theorem **??**), also wird es hier nur erwähnt. Wieder sind die Folgenräume nur Spezialfälle, die sich für das Zählmaß über der Grundmenge $\mathbb{N}$ als "Integrationsgebiet" ergeben.

Beispiel 4. *Mit $c_{00}^{\mathbb{K}}$ bezeichnen wir den Raum der abbrechenden Folgen in $\mathbb{K}$:*

$$c_{00}^{\mathbb{K}} := \left\{ (a_n)_n \in \mathbb{K}^{\mathbb{N}} : \text{ es existiert ein } n_0 \in \mathbb{N} \text{ mit } a_n = 0 \text{ für alle } n > n_0 \right\}.$$

Da alle abbrechenden Folgen beschränkt sind und Summen sowie Vielfache abbrechender Folgen abbrechen, handelt es sich um einen linearen Unterraum von $\ell_\infty^{\mathbb{K}}$. Versehen mit *der Supremumsnorm $\|\cdot\|_\infty$ ist c_{00} aber keineswegs abgeschlossen. Wir finden nämlich eine Cauchyfolge $(v^{(k)})_k$, die wir folgendermaßen definieren:*

$$v_n^{(k)} := \begin{cases} \frac{1}{n} & \text{für } n \leq k, \\ 0 & \text{sonst,} \end{cases}$$

d.h., das $k$te Folgeglied ist jene Zahlenfolge, deren erste k Einträge $1, \frac{1}{2}, \ldots, \frac{1}{k}$ sind, gefolgt von Nullen. Dass es sich bei $(v^{(k)})$ um eine Cauchyfolge handelt sehen wir schnell. Zu $\varepsilon > 0$ wählen wir $k_0 \in \mathbb{N}$ derart, dass $\frac{1}{k_0} < \varepsilon$ ist. Dann stimmen $(v^{(k)})$ und $(v^{(l)})$ für $k, l \geq k_0$ bis zum $(k \wedge l)$ten Folgeglied überein. Ist ohne Beschränkung der Allgemeinheit $k > l \geq k_0$, so ist also

$$\|v^{(k)} - v^{(l)}\|_\infty = \left| \frac{1}{l+1} - 0 \right| \leq \frac{1}{k_0 + 1} < \varepsilon.$$

Nun überzeugen wir uns aber, dass

$$v = (v_n)_n \in \mathbb{K}^{\mathbb{N}}, \qquad v_n = \frac{1}{n} \text{ für alle } n \in \mathbb{N}$$

die einzig mögliche Grenzfolge ist. Nehmen wir an, es gäbe eine andere Grenzfolge $w = (w_n)_n$, so wäre $w_N \neq v_N = \frac{1}{N}$ für ein $N \in \mathbb{N}$. Zu $\varepsilon = \frac{1}{2}|w_N - \frac{1}{N}|$ gälte dann für alle $k \geq N$:

$$v_N^{(k)} = \frac{1}{N}, \quad \text{also: } \|v^{(k)} - w\|_\infty \geq \left| \frac{1}{N} - w_N \right| = 2\varepsilon > \varepsilon,$$

also kann w unmöglich Grenzwert der Folge $(v^{(k)})_k$ sein. Die Grenzfolge v bricht aber keineswegs ab, gehört also nicht zu c_{00}.

Auch hier gibt es wieder Analoga für Funktionen auf allgemeineren Grundmengen:

Beispiel 5. *Der Raum $V := C([a,b];\mathbb{R})$ der stetigen, reellwertigen Funktionen auf dem kompakten Intervall $[a,b]$ ist ein Vektorraum, und für alle $f \in V$ ist (sogar das Regel-)Integral*

$$\|f\|_1 := \int_a^b |f(t)| \, \mathrm{d}t$$

wohldefiniert. Sie rechnen schnell nach, dass es sich bei $\|\cdot\|_1$ um eine Norm auf V handelt. Für die Definitheit ist wesentlich, dass wir nur stetige Funktionen als Integranden betrachten – das relevante Argument werden wir unten noch einmal sehen. Der normierte Vektorraum $(V, \|\cdot\|_1)$ ist nicht vollständig, also kein Banachraum. Das klassische Beispiel für eine Cauchyfolge ohne (stetige) Grenzfunktion liefert der Fall $[a,b] = [0,2]$,

$$f_n(t) = \begin{cases} t^n & \text{für } 0 \leq t < 1, \\ 1 & \text{für } 1 \leq t \leq 2, \end{cases}$$

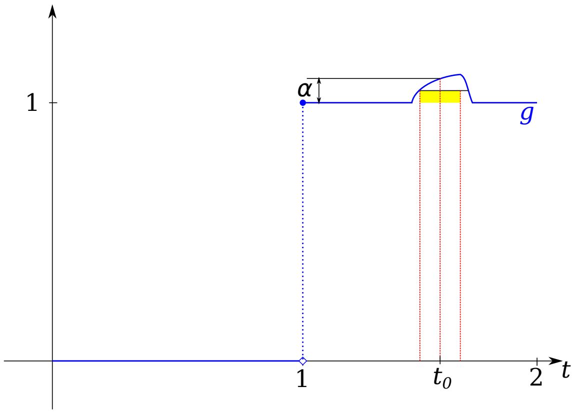
für die

$$\|f_n - f_k\|_1 = \int_0^1 |t^n - t^k| \, \mathrm{d}t \leq \frac{1}{(n \wedge k) + 1}$$

gilt. Als Grenzfunktion drängt sich $f = \mathbf{1}_{[1,2]}$ *auf, schließlich handelt es sich dabei um den punktweisen Grenzwert der* f_n*, und es gilt:*

$$\|f_n - f\|_1 = \int_0^1 t^n \,\mathrm{d}t = \frac{1}{n+1} \to 0 \quad \textit{für } n \to \infty,$$

jedoch ist f *keineswegs stetig.*
Tatsächlich kann die Folge $(f_n)_n$ *keinen stetigen Grenzwert haben. Wäre nämlich* g *eine solche stetige Grenzfunktion, die auf* $[1,2]$ *von* 1 *verschieden ist, so gäbe es ein* $t_0 \in [1,2]$ *mit* $g(t_0) := 1 + \alpha$ *für ein* $\alpha \neq 0$*. Dann existiert wegen der Stetigkeit ein* $\delta > 0$ *derart, dass* $|g(t) - 1| > \frac{|\alpha|}{2}$ *auf* $[1,2] \cap (t_0 - \delta, t_0 + \delta)$ *gilt.*



Wir können dabei δ *so klein wählen, dass mindestens eines der Teilintervalle* $(t_0 - \delta, t_0]$, $[t_0, t_0 + \delta)$ *ganz in* $[1,2]$ *liegt. Dann wäre aber z.B. im Fall* $(t_0 - \delta, t_0] \subset [1,2]$

$$\|f_n - g\|_1 = \int_0^2 |f_n(t) - g(t)| \,\mathrm{d}t \geq \int_{t_0-\delta}^{t_0} |1 - g(t)| \,\mathrm{d}t \geq \int_{t_0-\delta}^{t_0} \frac{|\alpha|}{2} \,\mathrm{d}t = \frac{|\alpha|}{2}\delta.$$

Das kriegen wir beim besten Willen nicht kleiner als etwa $\varepsilon = \frac{|\alpha|}{4}\delta$*, also gilt bezüglich* $\|\cdot\|_1$ *nicht* $f_n \to g$*. Der Fall,* g *sei auf* $[0,1)$ *nicht identisch Null wird analog behandelt*[8]*.*
Wir werden also keinen stetigen Grenzwert finden.
In einer Übungsaufgabe werden Sie sich überzeugen, dass V *unter der durch*

$$\|f\|_\infty := \sup_{0 \leq t \leq 2} |f(t)|$$

sehr wohl ein Banachraum ist[9]*. Schließlich ist die Konvergenz bezüglich dieser Norm (wohlgemerkt bei kompaktem* $[a,b]$*) die gleichmäßige Konvergenz, und gleichmäßige Limiten stetiger Funktionen sind stetig.*

[8] Wir müssen dann δ so klein wählen, dass $t_0 + \delta =: s < 1$ gilt, damit $f_n(t) \leq s^n$ für $t < t_0 + \delta$ gilt.

[9] Diese Norm kennen wir von den Räumen $L_\infty(\mu)$ der (Äquivalenzklassen von) wesentlich beschränkten messbaren Funktionen. Wir stellen also fest, dass V ein abgeschlossener Unterraum des Raums $L_\infty(\lambda^1|_{[a,b]})$, versehen mit der Norm $\|f\|_\infty := \operatorname*{ess\,sup}_t |f(t)|$, ist.

Notation. Den Raum $C(M;\mathbb{K})$ der stetigen Funktionen von einem kompakten metrischen Raum[10] (M,d), versehen mit der Maximumsnorm

$$\|v\|_\infty = \sup_{x\in M} |v(x)|$$

bezeichnen wir als $C^0(M;\mathbb{K})$. Das passt im reellen Fall in eine ganze Familie von Räumen $C^k(M;\mathbb{R})$ der auf kompaktem $M \subset \mathbb{R}^n$ $k$mal stetig differenzierbaren Funktionen, versehen mit der Norm

$$\|v\|_{C^k} = \sum_{|a|\leq k} \|D^a v\|_\infty,$$

bei denen die Summation über alle Multiindizes $a = (a_1,\ldots,a_n)$ der Länge $|a| = a_1 + \ldots + a_n \leq k$ läuft. Wir summieren also einfach die Maximumsnormen aller Ableitungen von v der Ordnungen $0,1,\ldots,k$ auf. Die so definierten Räume $C^k(M;\mathbb{R})$ sind auch Banachräume.

Wir merken uns: Der gleiche Vektorraum V kann unter manchen Normen ein Banachraum, unter anderen aber unvollständig sein.
Folgenden Zusammenhang zwischen abgeschlossenen Mengen und Vollständigkeit kennen wir auch schon: abgeschlossene Teilräume vollständiger Räume sind vollständig.

Lemma 1.17. *Eine Teilmenge $M \subset V$ eines vollständigen metrischen Raums (V,d) ist genau dann abgeschlossen, wenn $(M, d|_{M\times M})$ ebenfalls vollständig ist.*

Beweis. Übungsblatt 0 □

Das folgende Beispiel beantwortet bereits eine unserer Fragen vom Anfang: Sind lineare Unterräume eines normierten Vektorraums stets abgeschlossen?

Beispiel 6. *Wir folgern, dass c_{00} ein nicht abgeschlossener Teilraum von $(\ell_\infty, \|\cdot\|_\infty)$ ist. Im Rahmen einer Übungsaufgabe überzeugen Sie sich, dass der Abschluss von c_{00} in ℓ_∞ der Raum c_0 der Nullfolgen ist.*

In Erinnerung an die rationalen Zahlen hatten wir den Begriff dichter Teilmengen metrischer Räume eingeführt. Auch diesen wiederholen wir hier.

Definition 1.18. *Eine Teilmenge $M \subset V$ eines (nicht notwendigerweise vollständigen) metrischen Raums (V,d) heißt dicht in V, wenn $\bar{M} = V$ gilt.*

Beispiel 7. *Im Rahmen unseres Standardbeispiels folgt, dass c_{00} in ℓ_∞ nicht dicht liegen kann, da der Abschluss c_0 von c_{00} ein echter Teilraum von ℓ_∞ ist. Das hätten wir natürlich auch einfacher haben können. Versuchen Sie etwa mal, eine $\|\cdot\|_\infty$-Cauchyfolge in c_{00} zu finden, deren Grenzwert $\mathbb{1}_{l_\infty}$ (die nur aus Einsen bestehende Folge) ist.*

Dichte Teilmengen können wir dank Lemma 1.14 alternativ auch folgendermaßen charaktisieren.

[10] Die Definition sollten Sie schon kennen, wir werden sie aber gleich nochmal wiederholen.

Korollar 1.19. *Eine Teilmenge $M \subset V$ eines metrischen Raums (V, d) ist genau dann dicht, wenn zu jedem $v \in V$ und jedem $\varepsilon > 0$ ein $w \in M$ mit $d(v, w) < \varepsilon$ existiert.*

Mit anderen Worten: Ist $M \in V$ dicht, so gilt für jedes $\varepsilon > 0$:

$$V \subset \bigcup_{w \in M} B_\varepsilon(w).$$

Beispiel 8. *Der Approximationssatz von Stone-Weierstraß*[11] *besagt, dass der (Vektor-)Raum der Polynome auf $[a, b]$ dicht im Raum $C^0([a, b]; \mathbb{R})$ liegt, dass wir also stetige Funktionen auf kompakten Intervallen beliebig genau gleichmäßig durch Polynome approximieren können: Zu stetigem $f : [a, b] \to \mathbb{R}$ und $\varepsilon > 0$ existiert ein Polynom p derart, dass*

$$\|f - p\|_\infty = \sup_{a \le t \le b} |f(t) - p(t)| < \varepsilon$$

gilt.

Wir werden in Kürze auf dichte Teilmengen zurück kommen, wenn wir separable Räume betrachten.
Nun stellt sich noch die Frage, ob wir aus einem nicht vollständigen metrischen Raum (V, d) stets durch Hinzunahme weiterer Elemente einen vollständigen machen können, in dem V dann dicht liegt. Bei den rationalen Zahlen hat das ja ganz gut funktioniert; das Ergebnis waren die reellen Zahlen. Sofern der gegebene Raum Teilraum eines vollständigen Raums ist, lösen wir das Problem einfach gemäß Lemma 1.17 durch Abschluss. Wie sieht es nun aber aus, wenn wir (noch) gar keinen Raum haben, dessen Teilraum V ist?

Satz 1.20. *Zu jedem metrischen Raum (V, d) existiert ein vollständiger Raum $(\tilde{V}, \tilde{d})$ derart, dass V eine (bzgl. $\tilde{d}$) dichte Teilmenge von $\bar{V}$ ist und*

$$\text{für alle } v, w \in V \qquad \tilde{d}(v, w) = d(v, w)$$

gilt.

Beweis. Auf der Menge $\mathcal{C}_d$ aller Cauchyfolgen in (V, d) definieren wir die Äquivalenzrelation $\sim$ durch

$$(v_n)_n \sim (w_n)_n \;:\Longleftrightarrow\; \forall\, \varepsilon > 0\; \exists\, n_0 : d(v_n, w_n) < \varepsilon \text{ für } n \ge n_0.$$

Zwei Cauchyfolgen sind also äquivalent, wenn ihre Glieder für hinreichend große n beliebig nah beieinander liegen. Überzeugen Sie sich, dass es sich bei $\sim$ tatsächlich um eine Äquivalenzrelation handelt, und dass wir für äquivalente Folgen $(v_n)_n \sim (w_n)_n$ zu gegebenem $\varepsilon > 0$ sogar ein n_1 derart finden, dass $d(v_m, w_l) < \varepsilon$ für beliebige $k, n \ge n_1$ gilt.

[11] siehe etwa Satz I.2.11 bei Werner (2018)

Nun bezeichnen wir mit $\tilde{V} = \mathcal{C}_d/_\sim$ die Menge der Äquivalenzklassen von Cauchyfolgen in V und konstruieren $\tilde{d}$ wie folgt. Für zwei Cauchyfolgen $(v_n)_n$ und $(w_n)_n$ definieren wir

$$\hat{d}((v_n)_n, (w_n)_n) := \lim_{n\to\infty} d(v_n, w_n).$$

Das ist wohldefiniert, weil $\alpha_n := d(v_n, w_n)$ eine Cauchyfolge in $\mathbb{R}$ definiert, die selbstverständlich einen Grenzwert hat. Weiter stellen wir fest, dass dieser Pseudoabstand[12] nach Konstruktion genau dann verschwindet, wenn beide Folgen zur gleichen Äquivalenzklasse gehören, und dass dann wegen der Dreiecksungleichung für andere Repräsentanten $(x_n)_n \sim (v_n)_n$ und $(y_n)_n \sim (w_n)_n$ stets

$$\hat{d}((x_n)_n, (y_n)_n) = \hat{d}((v_n)_n, (w_n)_n)$$

gilt. Es ergibt also Sinn,

$$\tilde{d}\big([(v_n)_n], [(w_n)_n]\big) := \hat{d}((v_n)_n, (w_n)_n)$$

für die durch $(v_n)_n$ bzw. $(w_n)_n$ repräsentierten Äquivalenzklassen zu definieren.

Hat eine Cauchyfolge $(v_n)_n$ nun einen Grenzwert $v \in V$, so ist v auch Grenzwert einer jeden Cauchyfolge $(w_n)_n \sim (v_n)_n$:

$$d(w_n, v) \le d(w_n, v_n) + d(v_n, v) \to 0 \quad \text{für } n \to \infty,$$

wenn $(w_n)_n \sim (v_n)_n$ und $v_n \to v$. Ferner gibt es zu jedem $v \in V$ die durch $v_n = v$ gegebene konstante Cauchyfolge, die trivialerweise gegen v konvergiert. Indem wir v also mit der Äquivalenzklasse dieser konstanten Folge identifizieren, erhalten wir $V \subset \tilde{V}$ und

$$\tilde{d}(v, w) = \hat{d}((v_n = v)_n, (w_n = w)_n) = \lim_{n\to\infty} d(v, w) = d(v, w).$$

Wir haben noch zu zeigen, dass V dicht in $\tilde{V}$ liegt. Dazu sei $(v_n)_n$ Repräsentant einer beliebigen Äquivalenzklasse. Zu $\varepsilon > 0$ finden wir n_0 derart, dass $d(v_n, v_k) < \varepsilon$ für $n, k \ge n_0$ gilt. Dann betrachten wir die konstante Folge $w_n = w := v_{n_0}$ für alle n. Diese repräsentiert (per Identifikation von w mit $[(w_n)_n]$) ein Element in $V \subset \tilde{V}$, für das

$$\tilde{d}(w, [(v_n)_n]) = \hat{d}((w_n)_n, (v_n)_n) = \lim_{n\to\infty} \underbrace{d(v_{n_0}, v_n)}_{<\varepsilon \text{ f. } n \ge n_0} \le \varepsilon$$

gilt. □

Wir interessieren uns nun für kompakte Teilmengen metrischer Räume. Dazu definieren zwei Arten von Kompaktheit, die zunächst wenig miteinander zu tun zu haben scheinen.

Definition 1.21. (1) *Ein metrischer Raum (V, d) heißt* <u>*überdeckungskompakt,*</u> *wenn jede offenen Überdeckung $\{U_\alpha\}_{\alpha\in\mathcal{I}}$ von V eine endliche Teilüberdeckung $\{U_{\alpha_1}, \dots, U_{\alpha_n}\}$ enthält.*

(2) *Ein metrischer Raum (V, d) heißt* <u>*folgenkompakt,*</u> *wenn jede Folge in V eine in V konvergente Teilfolge enthält.*

[12] Überzeugen Sie sich, dass $\hat{d}$ nicht negativ und symmetrisch ist und die Dreiecksungleichung erfüllt.

Nach Teil (1) muss also jede, nicht notwendigerweise abzählbare, Familie offener Mengen, die ganz V ausfüllen, eine endliche Teilfamilie enthalten, die schon genügt, um V auszufüllen. Teil (2) besagt, dass V folgenkompakt ist, wenn jede Folge $(v_n)_n$ in V eine Teilfolge $(v_{n_k})_k$ enthält, die gegen ein Element von V konvergiert.
Die Definitionen ergeben auch in allgemeinen topologischen Räumen Sinn, allerdings haben metrische Räume die erfreuliche Eigenschaft, dass beide übereinstimmen.

Theorem 1.22. *Für einen metrischen Raum (V,d) sind folgende Aussagen äquivalent.*

(i) *(V,d) ist überdeckungskompakt.*

(ii) *(V,d) ist folgenkompakt.*

(iii) *(V,d) ist vollständig und zu jedem $\varepsilon > 0$ existieren endlich viele $v_1, \ldots, v_n$ derart, dass*

$$V \subset \bigcup_{j=1}^{n} B_\varepsilon(v_j)$$

gilt.

Beweis. *(i)* $\implies$ *(ii)*. Hat die Folge $(v_n)_n$ keine konvergente Teilfolge (d.h., kein $v \in V$ ist Grenzwert einer Teilfolge), so finden wir zu jedem $v \in V$ ein $\varepsilon_v > 0$ derart, dass $U_v := B_{\varepsilon_v}(v)$ höchstens endlich viele der v_n enthält (genauer enthält U_v nur v_n für endlich viele n). Die U_v bilden eine offene Überdeckung von V, denn jedes $v \in V$ liegt mindestens in seinem persönlichen U_v. Wäre V also kompakt, so genügten endlich viele der U_v, um V zu überdecken. In jedem dieser endlich vielen U_v liegen aber nur endlich viele der v_n im Widerspruch dazu, dass alle (unendlich vielen) v_n in V liegen.

(ii) $\implies$ *(iii)*. Angenommen, die Überdeckungsbedingung in *(iii)* wäre verletzt, wir hätten also ein $\varepsilon > 0$, für das wir V nicht mit endlich vielen ε-Kugeln überdecken können. Dann wählen wir $v_1 \in V$ beliebig und finden wegen $B_\varepsilon(v_1) \subsetneqq V$ ein v_2 mit $d(v_1, v_2) \geq \varepsilon$. Induktiv finden wir so stets ein

$$v_{n+1} \in V \setminus \bigcup_{k=1}^{n} B_\varepsilon(v_k),$$

das von allen bisher gefundenen v_k mindesten den Abstand ε hat. Wir haben also eine Folge $(v_n)_n$ konstruiert, für die $d(v_k, v_n) \geq \varepsilon$ für alle n gilt. Das gilt natürlich auch für jede Teilfolge, also kann keine Teilfolge von $(v_n)_n$ eine Cauchyfolge sein, geschweige denn konvergieren.
Die Vollständigkeit folgt einfach daraus, dass nach *(ii)* insbesondere jede Cauchyfolge eine konvergente Teilfolge besitzt, dann aber notwendigerweise (gegen den Grenzwert dieser Teilfolge) selbst konvergiert.

(iii) $\implies$ *(i)*. Angenommen $\{U_\alpha\}_{\alpha \in \mathcal{I}}$ sei eine Überdeckung von V, die keine endliche Teilüberdeckung enthält. Zu $k \in \mathbb{N}$ setzen wir $\varepsilon_k = 2^{-k}$ und finden nach *(iii)* jeweils

$v_1^k, \dots, v_{n(k)}^k$ derart, dass

$$V = \bigcup_{j=1}^{n(k)} B_{2^{-k}}(v_j^k)$$

gilt. Im ersten Schritt kann mindestens eines der $B_{2^{-1}}(v_j^1)$ nicht von endlich vielen der U_α überdeckt werden, ohne Beschränkung der Allgemeinheit möge dies das mit $j = 1$ sein.

Im zweiten Schritt können wir diese Kugel $B_1 = B_{2^{-1}}(v_1^1)$ als

$$B_1 = \bigcup_{j=1}^{n(2)} \left(B_1 \cap B_{2^{-2}}(v_j^2)\right)$$

darstellen. Da aber unendlich viele U_α benötigt werden um B_1 zu überdecken, brauchen wir auch für mindestens einen der $n(2)$ Durchschnitte unendlich viele U_α. Wir nehmen wieder ohne Beschränkung der Allgemeinheit an, es sei der mit $j = 1$ und setzen $B_2 := B_{2^{-2}}(v_1^2)$: $B_1 \cap B_2$ kann nicht von endlich vielen U_α überdeckt werden

Im dritten Schritt stellen wir dann analog fest, dass $B_1 \cap B_2 \cap B_3$ nicht von endlich vielen U_α überdeckt werden kann, wobei $B_3 = B_{2^{-3}(v_1^3)}$ ist. Wir erhalten so eine Folge $(w_k)_k$, gegeben durch $w_k := v_1^k$, derart, dass

$$\bigcap_{k=1}^{m} B_{2^{-k}}(w_k)$$

für kein m von endlich vielen der U_α überdeckt werden kann. Da nun also für alle k die Folgeglieder w_k und w_{k+1} insbesondere in $B_k \cup B_{k+1}$ mit $B_k \cap B_{k+1} \neq \emptyset$ liegen, gilt

$$d(w_k, w_{k+1}) \leq 2^{-k} 2 = 2^{-k+1},$$

also gilt für $l > m$:

$$d(w_m, w_l) \leq \sum_{k=m}^{l-1} d(w_k, w_{k+1}) \leq \sum_{k=m}^{l-1} 2^{-k+1} \leq \sum_{k=m-1}^{\infty} 2^{-k} = \frac{1}{1-\frac{1}{2}} - \frac{1-\left(\frac{1}{2}\right)^m}{1-\frac{1}{2}} = 2^{-m+1}.$$

Also ist $(w_n)_n$ eine Cauchyfolge, die wegen der Vollständigkeit von V gegen ein $w \in V$ konvergiert.

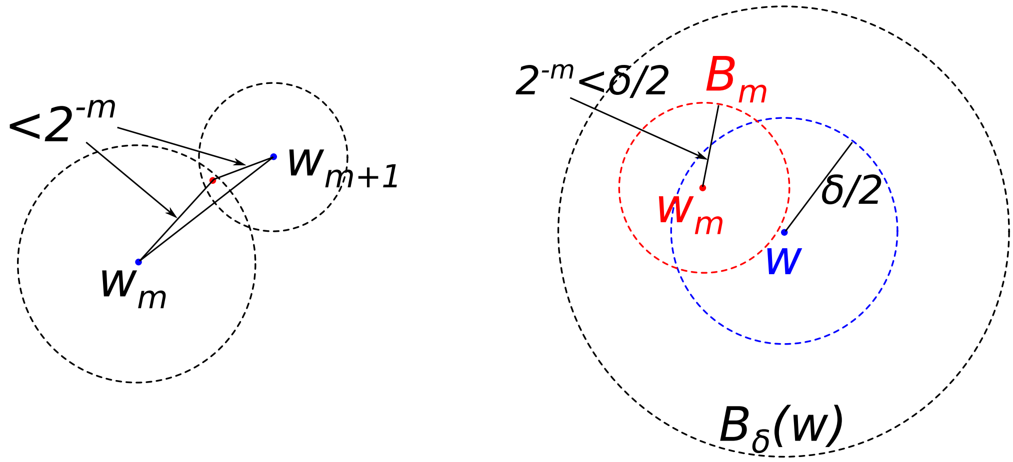


Illustration des Abstands $d(w_m, w_{m+1})$ via Dreiecksungleichung (links) und der Überdeckung aller B_m, $m \geq n$, durch $B_\delta(w)$ (rechts).

Dieses liegt in mindestens einem der U_α, sagen wir $w \in U_{\alpha_0}$, und da U_{α_0} offen ist, finden wir ein $\delta > 0$ derart, dass $B_\delta(w) \subset U_{\alpha_0}$ gilt. Wählen wir nun n so, dass
1. $d(w_m, w) < \frac{\delta}{2}$ für alle $m \geq n$ (möglich wegen der Konvergenz) und
2. $2^{-n} < \frac{\delta}{2}$
gelten, so liegen alle B_m, $m \geq n$, in $B_\delta(w) \subset U_{\alpha_0}$, werden also vom einzelnen U_{α_0} überdeckt:

$$u \in B_m \implies d(u, w) \leq \underbrace{d(u, w_m)}_{<2^{-m} \leq 2^{-n}} + d(w_m, w) < \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$

Insbesondere gilt also

$$B_1 \cap \cdots \cap B_n \subset B_n \subset U_{\alpha_0}$$

im Widerspruch zu Konstruktion der B_j. □

Die Äquivalenz von (i) und (ii) ermuntert uns zu folgender Definition.

Definition 1.23. *Eine Teilmenge $M \subset V$ eines metrischen Raums (V, d) heißt kompakt, wenn sie (als Teilraum $(M, d|_{M \times M})$) die Aussagen von Theorem 1.22 erfüllt.*

Wir werden also im folgenden bei Teilmengen metrischer Räume und insbesondere normierter Vektorräume nur noch von kompakten Mengen sprechen, beachten aber, dass in allgemeineren topologischen Räumen Überdeckungs- und Folgenkompaktheit sehr wohl verschiedenen Bedingungen sein können[13]. In den meisten Texten, vor allem im Englischen, wird Überdeckungskompaktheit als Definition der Kompaktheit genutzt. Der Begriff der Überdeckungskompaktheit wird dann oft gar nicht separat eingeführt.

[13] Aussage (iii) ergibt in allgemeinen topologischen Räumen ohne Metrik gar keinen Sinn.

Bemerkung. *Unabhängig davon, dass wir uns sehr für kompakte Teilmengen von Banachräumen interessieren werden, liefern kompakte metrische Räume die Grundlage für eine große Klasse von Banachräumen. Ist nämlich (M, d) ein kompakter metrischer Raum, so wissen wir, dass stetige reellwertige Funktionen $f : M \to \mathbb{R}$ Minimum und Maximum annehmen, und dass der Raum $C(M; \mathbb{K})$ der stetigen Funktionen von M nach $\mathbb{K}$, versehen mit der Maximumsnorm*

$$\|f\|_\infty = \max_{x \in M} |f(x)|$$

ein Banachraum ist. Das ist eine Verallgemeinerung des zweiten Teils von Beispiel 5.

Bedingung (iii) in Theorem 1.22 erinnert uns an den Satz von Heine-Borel, der uns sagte, dass die kompakten Teilmengen des $\mathbb{R}^n$ genau jene sind, die abgeschlossen und beschränkt sind. Nun ergibt das Konzept der Beschränktheit in allgemeinen metrischen Räumen nur wenig topologischen Sinn, da wir zur Metrik d auf V durch

$$\tilde{d}(v, w) = 1 \wedge d(v, w)$$

eine Metrik finden, welche die gleiche Topologie (also die gleichen konvergenten Folgen und die gleichen kompakten Mengen) erzeugt[14] und bezüglich der alle Mengen beschränkt sind.

Definition 1.24. *Eine Teilmenge M eines metrischen Raums (V, d) heißt total beschränkt (auch: präkompakt), falls zu jedem $\varepsilon > 0$ endlich viele $v_1, \ldots, v_n \in M$ derart existieren, dass*

$$M \subset \bigcup_{j=1}^{n} B_\varepsilon(v_j)$$

gilt.
M heißt relativ kompakt, wenn $\bar{M}$ kompakt ist.

Bemerkung. *Oft wird auch der Begriff der relativen Folgenkompaktheit benutzt. Eine Menge $M \subset V$ heißt relativ folgenkompakt, wenn jede Folge $(v_n)_n$ in M eine in V konvergente Teilfolge besitzt. Überzeugen Sie sich, dass relativ folgenkompakt und relativ kompakt für Teilmengen metrischer Räume äquivalente Eigenschaften sind.*

Das Theorem besagte also in Analogie zum Satz von Heine-Borel, dass Teilmengen eines vollständigen metrischen Raums (z.B. eines Banachraums) genau dann kompakt sind, wenn sie abgeschlossen und total beschränkt sind. Insbesondere gilt, da aus totaler Beschränktheit stets Beschränktheit folgt, eine Richtung des Satzes von Heine-Borel weiterhin: Kompakte Teilmengen metrischer Räume sind abgeschlossen und beschränkt. Wir fassen das zusammen.

Korollar 1.25. *Eine Teilmenge $M \subset V$ eines vollständigen metrischen Raums (V, d) ist genau dann kompakt, wenn sie abgeschlossen und total beschränkt ist.*
Insbesondere sind kompakte Teilmengen metrischer Räume abgeschlossen und beschränkt.

[14] Überzeugen Sie sich davon!

Für allgemeine topologische Räume ist selbst der erste Teil der zweiten Aussage nicht immer wahr. Versehen wir etwa die Menge $X = \{0, 1\}$ mit der Klumpentopologie $\{\emptyset, X\}$, so ist die Teilmenge $\{0\}$ offensichtlich kompakt (endliche Teilmengen topologischer Räume sind immer kompakt), jedoch nicht abgeschlossen (das sind nur $\emptyset$ und X).
Die zweite Aussage des Korollars ist besonders für normierte Vektorräume interessant, bei denen die geforderte Homogenität der Norm die künstliche Beschränkung unbeschränkter Mengen verhindert. Überlegen Sie sich, dass in *vollständigen* metrischen Räumen relative Kompaktheit und Präkompaktheit äquivalent sind.
Dass wir uns im allgemeinen auch bei normierten Vektorräumen nicht auf Beschränktheit und Abgeschlossenheit berufen dürfen um Kompaktheit zu zeigen, zeigt folgendes Beispiel.

Beispiel 9. *In ℓ_p betrachten wir die kanonischen Einheitsvektoren $b^{(n)}$, gegeben durch*

$$b_k^{(n)} = \delta_{k,n} = \begin{cases} 1 & \text{für } k = n, \\ 0 & \text{sonst.} \end{cases}$$

Diese haben die Normen $\|b^{(n)}\|_p = 1$, liegen also in der abgeschlossenen (und beschränkten) Einheitskugel $D = \bar{B}_1(0)$. Allerdings ist für $n \neq m$

$$\|b^{(n)} - b^{(m)}\|_p = 2^{\frac{1}{p}},$$

also kann keine Teilfolge eine Cauchyfolge sein, geschweige denn gegen ein $v \in D$ konvergieren. Also ist die abgeschlossene Einheitskugel in ℓ_p nicht (folgen-)kompakt.

Tatsächlich ist die abgeschlossene Einheitskugel in einem normierten $\mathbb{K}$-Vektorraum genau dann kompakt, wenn der Raum endlichdimensional ist.
Im nächsten Kapitel beschäftigen wir uns noch mit besonderen Banachräumen, in denen wir nicht nur Längen, sondern auch Winkel messen können. Die Grundlage dafür ist eine Verallgemeinerung des euklidischen Skalarprodukts.

KAPITEL 2

Hilberträume und die Sätze von Riesz

2.1 Skalarprodukte und Hilberträume

Definition 2.1. *Eine Abbildung* $\langle\cdot|\cdot\rangle : V \times V \to \mathbb{K}$ *auf einem* $\mathbb{K}$*-Vektorraum* V *heißt*

1. *Semiskalarprodukt, falls für alle* $u, v, w \in V$, $\alpha \in \mathbb{K}$

$$\begin{aligned} &\langle u|\alpha v + w\rangle = \alpha\langle u|v\rangle + \langle u|w\rangle && \textit{(Linearität im zweiten Argument)} \\ &\langle v|w\rangle = \langle w|v\rangle^* && \textit{(konjugiert symmetrisch)} \\ &\langle v|v\rangle \geq 0 && \textit{(positiv semidefinit)} \end{aligned}$$

   gilt[a].

2. *Gilt zusätzlich* $\langle v|v\rangle = 0 \implies v = 0$, *so ist* $\langle\cdot|\cdot\rangle$ *ein Skalarprodukt.*

3. *Ein mit einem Skalarprodukt versehener Vektorraum* V *heißt Prähilbertraum oder unitärer Raum.*

[a] Dabei steht der Asterisk für die komplexe Konjugation. Im Fall $\mathbb{K} = \mathbb{R}$ kann dieser natürlich entfallen.

Bemerkung. *Der Begriff des unitären Raums wird oft für komplexe Prähilberträume reserviert.*
Im Fall $\mathbb{K} = \mathbb{R}$ *können keine komplexen Werte auftreten. Die ersten beiden Bedinungen sagen dann einfach, dass es sich bei* $\langle\cdot|\cdot\rangle$ *um eine symmetrische Bilinearform handelt. Im komplexen Fall spricht man von einer konjugiert symmetrischen Sesquilinearform.*
Sie rechnen sofort nach, dass daraus sofort die konjugierte Linearität im ersten Argument folgt:

$$\langle\alpha u + v|w\rangle = \alpha^*\langle u|w\rangle + \langle v|w\rangle \quad \textit{für } u, v, w \in V,\ \alpha \in \mathbb{K}.$$

Ebenso leicht überzeugen Sie sich, dass $\langle 0_V|v\rangle = 0 = \langle v|0_V\rangle$ *für alle* $v \in V$ *gelten muss.*

Ein Skalarprodukt erhalten wir dann, wenn diese Form auch noch positiv definit ist.

Als kleine Warnung sei hier noch angemerkt, dass in manchen Texten die Linearität im ersten Argument gefordert wird. Dann ist $\langle\cdot|\cdot\rangle$ natürlich im zweiten Argument konjugiert linear. Diese Konvention tritt besonders gern bei der in der Quantenmechanik beliebten Bracket-Notation auf.

Lemma 2.2 (Cauchy-Schwarz-Ungleichung)**.** *In einem Prähilbertraum $(V, \langle\cdot|\cdot\rangle)$ gilt für alle $v, w \in V$*

$$|\langle v|w\rangle|^2 \leq \langle v|v\rangle \, \langle w|w\rangle.$$

Beweis. Dank der Symmetriebedingung hat $\langle w|v\rangle$ den Betrag $|\langle v|w\rangle|$, also finden wir ein $\alpha \in \mathbb{K}$ mit $|\alpha| = 1$ derart, dass $\langle w|v\rangle = \alpha|\langle v|w\rangle|$ ist. Dann rechnen wir für alle $\lambda \in \mathbb{R}$

$$\begin{aligned} 0 &\leq \langle v - \lambda\alpha w | v - \lambda\alpha w\rangle \\ &= \langle v|v\rangle - \lambda\alpha^* \underbrace{\langle w|v\rangle}_{=\alpha|\langle v|w\rangle|} - \lambda\alpha \underbrace{\langle v|w\rangle}_{=(\alpha|\langle v|w\rangle|)^*} + \lambda^2 \underbrace{\alpha^*\alpha}_{=1}\langle w|w\rangle \\ &= \langle v|v\rangle - 2\lambda|\langle v|w\rangle| + \lambda^2\langle w|w\rangle. \end{aligned}$$

Das so definierte quadratische Polynom in λ hat also höchstens eine reelle Wurzel, muss also eine nicht positive Diskriminante

$$D = (2|\langle v|w\rangle|)^2 - 4\langle v|v\rangle\langle w|w\rangle \leq 0$$

haben. Das ergibt aber genau die gesuchte Ungleichung. □

Bewaffnet mit dieser Ungleichung können wir aus (Semi-)Skalarprodukten sofort (Halb-)Normen machen.

Satz 2.3. *Auf jedem Prähilbertraum $(V, \langle\cdot|\cdot\rangle)$ ist durch*

$$\|v\| := \sqrt{\langle v|v\rangle}, \qquad v \in V$$

eine Norm definiert.

Beweis. Die Wohldefiniertheit und positive Definitheit der Norm sind sofort klar.
Für die Homogenität rechnen wir in einer Zeile

$$\|\alpha v\| = \sqrt{\langle\alpha v|\alpha v\rangle} = \sqrt{\alpha^*\alpha\langle v|v\rangle} = \sqrt{|\alpha|^2\|v\|^2}.$$

Und schließlich ergibt die Cauchy-Schwarz-Ungleichung die Dreiecksungleichung:

$$\begin{aligned} \|v + w\|^2 &= \langle v|v\rangle + \underbrace{\langle v|w\rangle + \langle w|v\rangle}_{=2\mathrm{Re}\langle v|w\rangle} + \langle w|w\rangle \\ &\leq \|v\|^2 + 2|\langle v|w\rangle| + \|w\|^2 \\ \text{(C.-S.)} \quad &\leq \|v\|^2 + 2\sqrt{\langle v|v\rangle\langle w|w\rangle} + \|w\|^2 \\ &= (\|v\| + \|w\|)^2. \end{aligned}$$

Ziehen wir auf beiden Seiten die Wurzel und beachten, dass alle Terme an den Enden der Ungleichungskette nicht negativ sind, steht die Dreiecksungleichung schon da. □

Beispiel 10. *Der $\mathbb{K}^n$ mit dem euklidischen Skalarprodukt*

$$\langle v|w\rangle = \sum_{k=1}^{n} v_k^* w_k$$

ist uns schon hinlänglich bekannt. Im Fall $\mathbb{K} = \mathbb{R}$ können wir die wirkungslose komplexe Konjugation einfach ignorieren.
Wie so oft können wir dieses Beispiel auf Folgenräume erweitern. Auf $\ell_2^{\mathbb{K}}$ ist durch

$$\langle v|w\rangle = \sum_{k=1}^{\infty} v_k^* w_k$$

ein Skalarprodukt gegeben. Die Sesquilinearität und konjugierte Symmetrie sind anhand der Rechenregeln für (absolut konvergente Reihen) klar.
Wegen $v_k^* v_k = |v_k|^2$ folgt auch die positive Semidefinitheit.
Ist $\langle v|v\rangle = 0$, so müssen alle Summanden $|v_k|^2 = 0$ sein, wir haben es also mit der trivialen Folge 0_{ℓ_2} zu tun.
*Wir müssen noch zeigen, dass $\langle v|w\rangle$ tatsächlich durch eine absolut konvergente Reihe gegeben und damit wohldefiniert ist. Das ist aber gerade die Hölderungleichung (siehe Lemma **??**) für das Zählmaß auf $\mathcal{P}(\mathbb{N})$.*
Die von den hier angegebenen Skalarprodukten erzeugten Normen sind genau die bereits bekannten $\|\cdot\|_2$-Normen.

Beispiel 11. *Für allgemeinere Maßräume erhalten wir völlig analog zu ℓ_2 auf $\mathcal{L}_2(\mu;\mathbb{K})$ das Semiskalarprodukt*

$$\langle f|g\rangle = \int_\Omega f^* g \,\mathrm{d}\mu,$$

das auf dem Restklassenraum $L_2(\mu;\mathbb{K})$ zum Skalarprodukt wird, durch das auch hier die bereits bekannte Norm $\|\cdot\|_2$ induziert wird.

All diese Räume hatten wir als Banachräume identifiziert. Das motiviert folgende Definition.

Definition 2.4. *Ein unitärer Raum $(V, \langle\cdot|\cdot\rangle)$ heißt Hilbertraum, wenn er unter der durch das Skalarprodukt induzierten Norm vollständig ist.*

Im Rahmen einer Übungsaufgabe rechnen Sie noch folgende Gleichungen nach:

Lemma 2.5 (Polarisierungsidentitäten)**.** *In einem Prähilbertraum $(V, \langle\cdot|\cdot\rangle)$ gilt mit der durch das Skalarprodukt induzierte Norm $\|\cdot\|$ für alle $v, w \in V$*

$$\begin{aligned}
&(i)\ \langle v|w\rangle = \frac{1}{4}\left(\|v+w\|^2 - \|v-w\|^2\right) && (\mathbb{K} = \mathbb{R})\\
&(ii)\ \langle v|w\rangle = \frac{1}{4}\left(\|v+w\|^2 - \|v-w\|^2 - \imath\|v+\imath w\|^2 + \imath\|v-\imath w\|^2\right) && (\mathbb{K} = \mathbb{C})
\end{aligned}$$

Beweis. Auf der rechten Seite beginnend nachrechnen (Übung). □

Die Polarisierungsidentitäten stellen sicher, dass eine gegebene Norm nicht durch verschiedene Skalarprodukte induziert werden kann. Umgekehrt stellt sich aber die Frage, ob jeder Norm ein Skalarprodukt zugrunde liegt, das durch die jeweilige Polarisierungsidentität gegeben ist. Die Antwort ist negativ, aber wir können ein Kriterium angeben.

Satz 2.6 (Jordan-von Neumann)**.** *Die Norm* $\|\cdot\|$ *auf dem* $\mathbb{K}$*-Vektorraum wird genau dann durch ein Skalarprodukt induziert, wenn für alle* $v, w \in V$ *die Parallelogrammgleichung*

$$\|v+w\|^2 + \|v-w\|^2 = 2(\|v\|^2 + \|w\|^2) \tag{2.1}$$

gilt. In diesem Fall ist das Skalarprodukt durch die Polarisierungsidentitäten eindeutig bestimmt.

Beweis. Von der Gültigkeit der Parallelogrammgleichung in einem unitären Raum überzeugen Sie sich durch Nachrechnen. Dabei hilft die Skizze eines ganz elementar im euklidischen $\mathbb{R}^2$ von zwei Vektoren aufgespannten Parallelogramms.
Die Umkehrung zeigen wir für den Fall $\mathbb{K} = \mathbb{R}$, müssen also zeigen, dass die durch

$$\langle v|w\rangle \;=\; \frac{1}{4}\left(\|v+w\|^2 - \|v-w\|^2\right)$$

definierte Abbildung $\langle\cdot|\cdot\rangle : V \times V \to \mathbb{R}$ ein Skalarprodukt ist. Der rechnerisch etwas aufwendige Teil ist die Linearität.
(I) Für $x_1, x_2, y \in V$ berechnen wir

$$\begin{aligned}
\|x_1+x_2+y\|^2 \;&=\; \frac{1}{2}\left(\|(x_1+y)+x_2)\|^2 + \|(x_2+y)+x_1\|^2\right) \\
&\quad \text{(2.1) mit } v = x_2, w = x_1+y \text{ bzw. } v = x_2+y, w = x_1 \\
&=\; \frac{1}{2}\left(2\|x_1+y\|^2 + 2\|x_2\|^2 - \|x_1-x_2+y\|^2\right) \\
&\quad + \frac{1}{2}\left(2\|x_2+y\|^2 + 2\|x_1\|^2 - \|-x_1+x_2+y\|^2\right) \\
&=\; \|x_1+y\|^2 + \|x_2+y\|^2 + \|x_1\|^2 + \|x_2\|^2 \\
&\quad - \frac{1}{2}\left(\|x_1-x_2+y\|^2 + \|-x_1+x_2+y\|^2\right)
\end{aligned} \tag{2.2}$$

und durch Ersetzen von y durch $-y$ analog

$$\begin{aligned}
\|x_1+x_2-y\|^2 \;&=\; \|x_1-y\|^2 + \|x_2-y\|^2 + \|x_1\|^2 + \|x_2\|^2 \\
&\quad - \frac{1}{2}\left(\|x_1-x_2-y\|^2 + \|-x_1+x_2-y\|^2\right)
\end{aligned} \tag{2.3}$$

Das nutzen wir zum Nachrechnen der Additivität (im ersten Argument):

$$\begin{aligned}
\langle x_1 + x_2 \mid y\rangle &= \frac{1}{4}\left(\|x_1+x_2+y\|^2 - \|x_1+x_2-y\|^2\right)\\
\text{(via (2.2))} \quad &= \frac{1}{4}\Bigg[\|x_1+y\|^2 + \|x_2+y\|^2 + \|x_1\|^2 + \|x_2\|^2\\
&\qquad - \frac{1}{2}\left(\|x_1 - x_2 + y\|^2 + \|-x_1+x_2+y\|^2\right)\Bigg]\\
\text{(via (2.3))} \quad &\quad - \frac{1}{4}\Bigg[\|x_1-y\|^2 + \|x_2-y\|^2 + \|x_1\|^2 + \|x_2\|^2\\
&\qquad - \frac{1}{2}\left(\|x_1 - x_2 - y\|^2 + \|-x_1+x_2-y\|^2\right)\Bigg]\\
&= \frac{1}{4}\left(\|x_1+y\|^2 - \|x_1-y\|^2\right) + \frac{1}{4}\left(\|x_2+y\|^2 - \|x_2-y\|^2\right)\\
&= \langle x_1|y\rangle + \langle x_2|y\rangle.
\end{aligned}$$

(II) Für den Rest der Linearität (im reellen Fall entfällt der Teil mit der komplexen Konjugation) zeigen wir zunächst für $\alpha = 0$:

$$\langle 0v|w\rangle = \frac{1}{4}\left(\|w\|^2 - \|w\|^2\right) = 0\langle v|w\rangle$$

und für $\alpha = 2$ ($\alpha = 1$ ist trivial):

$$\langle 2v|w\rangle = \langle v+v|w\rangle = \langle v|w\rangle + \langle v|w\rangle = 2\langle v|w\rangle,$$

und dann induktiv $\langle \alpha v|w\rangle = \alpha\langle v|w\rangle$ für $\alpha \in \mathbb{N}_0$. Für nicht negative rationale Zahlen $\alpha = \frac{k}{m}$, $k \in \mathbb{N}_0, m \in \mathbb{N}$ erhalten wir

$$m\langle \alpha v|w\rangle = \langle m\alpha v|w\rangle = \langle kv|w\rangle = k\langle v|w\rangle,$$

was nach Teilen durch m die Aussage für $\alpha \in \mathbb{Q}_+$ ergibt. Zu den negativen rationalen Zahlen α kommen wir dann via

$$0 = \langle 0_V|w\rangle = \langle \alpha v + (-\alpha v)|w\rangle = \langle \alpha v|w\rangle + \underbrace{\langle -\alpha v|w\rangle}_{-\alpha\langle v|w\rangle}.$$

Wegen der Stetigkeit der Normen (und damit auch von $\langle\cdot|\cdot\rangle$ nach der Definition) folgt daraus $\langle \alpha v|w\rangle = \alpha\langle v|w\rangle$ für beliebige $\alpha \in \mathbb{R}$ und $v, w \in V$.

(III) Für die positive Definitheit rechnen wir nach

$$\langle v|v\rangle = \frac{1}{4}\left(\|2v\|^2 - \|0\|^2\right) = \|v\|^2 \geq 0$$

mit Gleichheit genau dann, wenn $v = 0$.

Die Symmetrie ist dank der Homogenität der Norm auch leicht nachzurechnen

$$\langle w|v\rangle = \frac{1}{4}\big(\|w+v\|^2 - \|\underbrace{w-v}_{=-(v-w)}\|^2\big) = \frac{1}{4}\left(\|v+w\|^2 - \|v-w\|\right) = \langle v|w\rangle.$$

Zusammen haben wir also gezeigt, dass es sich um ein Skalarprodukt handelt.
Im komplexen Fall werden die Rechnungen deutlich länger, aber im Prinzip ganz ähnlich.

□

2.2 Orthogonale Projektionen, erster Satz von Riesz

Wir hatten erwähnt, dass das Skalarprodukt uns bei der Bestimmung von Winkeln helfen soll. Der erste Schritt ist folgende Definition.

Definition 2.7. *Zwei Vektoren $v, w \in V$ in einem Prähilbertraum $(V, \langle\cdot|\cdot\rangle)$ heißen orthogonal, geschrieben $v \perp w$, falls $\langle v|w\rangle = 0$ gilt.*
Die Familie $\{v_i\}_{i\in\mathcal{I}} \subset V$ heißt Orthogonalsystem, falls die v_i paarweise orthogonal sind. Gilt zusätzlich noch $\langle v_i|v_i\rangle = 1$ für alle $i \in \mathcal{I}$, so sprechen wir von einem Orthonormalsystem.

Beispiel 12. *In ℓ_2, versehen mit dem Skalarprodukt aus Beispiel 10, ist die Familie $\{b^{(n)}\}_{n\in\mathbb{N}}$, gegeben durch $b_k^{(n)} = \delta_{k,n}$, ein Orthonormalsystem. Im endlichdimensionalen Fall sind das gerade die Standardeinheitsvektoren.*

Beispiel 13. *In $L_2([-\pi,\pi];\mathbb{K})$ sind sowohl $S = \{v_k = \sin(k\cdot)\}_{k\in\mathbb{N}}$ als auch $C = \{w_k = \cos(k\cdot)\}_{k\in\mathbb{N}_0}$ Orthogonalsysteme, wie wir schnell durch Integration nachrechnen[1]. Sie rechnen sogar nach, dass $S \cup C$ ein Orthogonalsystem ist. Ferner berechnen wir*

$$\langle v_k|v_k\rangle = \int_{-\pi}^{\pi} \sin^2(kt)\,\mathrm{d}t = \pi,$$

$$\langle w_k|w_k\rangle = \int_{-\pi}^{\pi} \cos^2(kt)\,\mathrm{d}t = \begin{cases} 2\pi & \text{für } k = 0, \\ \pi & \text{für } k \geq 1. \end{cases}$$

Normieren wir die v_k und w_k entsprechend und $\sqrt{\pi}$ oder $\sqrt{2\pi}$, so erhalten wir sogar ein Orthonormalsystem.

Bemerkung. *Wir finden sofort den aus der euklidischen Ebene bekannten Satz des Pythagoras wieder. Sind $v, w \in V$ mit $v \perp w$, so gilt*

$$\|v-w\|^2 = \langle v-w|v-w\rangle = \langle v|v\rangle - \langle v|w\rangle - \langle w|v\rangle + \langle w|w\rangle = \|v\|^2 + \|w\|^2.$$

Die gleiche Rechnung funktioniert natürlich auch für $v + w$.

[1] Für die diese stetigen Funktionen genügt auf dem Kompaktum $[-\pi,\pi]$ das Riemann- oder Regelintegral.

Lemma 2.8. *Sind $(V, \langle\cdot|\cdot\rangle)$ ein Hilbertraum, $v \in V$ beliebig und $M \subset V$ dicht in V, so gilt*

$$w \perp v \text{ für alle } w \in M \quad \Longrightarrow \quad v = 0.$$

Beweis. Übung (man nutze die Stetigkeit des Skalarprodukts). □

Definition 2.9. *Für zwei $\mathbb{K}$-Hilberträume $(V_i, \langle\cdot|\cdot\rangle_i)$, $i = 1, 2$, definieren wir die orthogonale (direkte) Summe $(V_1 \oplus V_2, \langle\cdot|\cdot\rangle)$ von V_1 und V_2 durch*

$$V_1 \oplus V_2 = \{(v_1, v_2) : v_1 \in V_1, v_2 \in V_2\}, \quad \langle(v_1, v_2)|(w_1, w_2)\rangle := \langle v_1|w_1\rangle_1 + \langle v_2|w_2\rangle_2.$$

Dass es sich dabei um einen Hilbertraum handelt, zeigt das folgende Lemma.

Lemma 2.10. *Sind $(V_i, \langle\cdot|\cdot\rangle_i)$, $i = 1, 2$, Hilberträume, so ist die ihre orthogonale Summe ebenfalls ein Hilbertraum.*

Beweis. Dass es sich bei $\langle\cdot|\cdot\rangle$ um ein Skalarprodukt handelt, rechnen Sie schnell nach. Zur Vollständigkeit stellen wir fest, dass die erzeugte Norm durch

$$\|(v_1, v_2)\| = \sqrt{\langle(v_1, v_2)\rangle} = \sqrt{\langle v_1|v_1\rangle_1 + \langle v_2|v_2\rangle_2} = (\|v_1\|_1^2 + \|v_2\|_2^2)^{\frac{1}{2}}$$

gegeben ist. Ist also $\left((v_1^{(n)}, v_2^{(n)})\right)_n$ eine Cauchyfolge in $V_1 \oplus V_2$, so sind notwendigerweise $(v_i^{(n)})_n$ Cauchyfolgen in V_i, $i = 1, 2$, haben also Grenzwerte $v_i = \lim_{n\to\infty} v_i^{(n)}$. Dann gilt aber zu gegebenem $\varepsilon > 0$ für n so groß, dass $\|v_i - v_i^{(n)}\|_i < \frac{\varepsilon}{2}$

$$\left\|(v_1, v_2) - \left(v_1^{(n)}, v_2^{(n)}\right)\right\| = \sqrt{\left\|v_1 - v_1^{(n)}\right\|_1^2 + \left\|v_2 - v_2^{(n)}\right\|_2^2} \le \sqrt{\frac{\varepsilon^2}{4} + \frac{\varepsilon^2}{4}} = \frac{\varepsilon}{\sqrt{2}} < \varepsilon,$$

also ist (v_1, v_2) Grenzwert der Folge $\left(v_1^{(n)}, v_2^{(n)}\right)_n$. □

Beispiel 14. *Die klassischen Fälle sind $\mathbb{R}^{n+m} = \mathbb{R}^n \oplus \mathbb{R}^m$. Sogar das "richtige" Skalarprodukt erhalten wir.*

Beispiel 15. *Als weniger triviales Beispiel gilt für $a < b < c$:*

$$L_2([a, b]; \mathbb{K}) \oplus L_2([b, c]; \mathbb{K}) = L_2([a, c]; \mathbb{K}).$$

Tatsächlich ist mit $v = v_1 \mathbf{1}_{[a,b)} + v_2 \mathbf{1}_{[b,c]}$, also

$$v(t) = \begin{cases} v_1(t) & \text{für } t \in [a, b), \\ v_2(t) & \text{für } t \in [b, c], \end{cases}$$

(wir beachten, dass die Änderung auf der Nullmenge $\{b\}$ die Funktion f im L_2-Sinn nicht ändert), das Skalarprodukt durch

$$\begin{aligned}\langle v|w\rangle &= \langle v_1|w_1\rangle_1 + \langle v_2|w_2\rangle_2\\ &= \int_{[a,b]} v_1(t)^* w_1(t)\,\mathrm{d}t + \int_{[b,c]} v_2(t)^* w_2(t)\,\mathrm{d}t\\ &= \int_{[a,c]} v_1(t)^* w_1(t)\mathbf{1}_{[a,b)}(t) + v_2(t)^* w_2(t)\mathbf{1}_{[b,c]}(t)\,\mathrm{d}t\\ &= \int_{[a,c]} v(t)^* w(t)\mathbf{1}_{[a,b)}(t) + v(t)^* w(t)\mathbf{1}_{[b,c]}(t)\,\mathrm{d}t\\ &= \int_{[a,c]} v(t)^* w(t)\,\mathrm{d}t.\end{aligned}$$

Offensichtlich wir die Summanden einer solchen orthogonalen Summe als Unterräume von $V_1 \oplus V_2$ auffassen. Da es sich bei ihnen kraft der Inklusionen

$$\iota_1 : V_1 \ni v_1 \mapsto (v_1, 0_{V_2}) \in V_1 \oplus V_2, \qquad \iota_2 : V_2 \ni v_2 \mapsto (0_{V_1}, v_2) \in V_1 \oplus V_2$$

um vollständige Räume nicht nur bezüglich ihres jeweils eigenen Skalarprodukts, sondern auch bezüglich $\langle\cdot|\cdot\rangle$ handelt, sind sie sogar abgeschlossen. Umgekehrt ist jeder abgeschlossene Unterraum U eines beliebigen Hilbertraums V kraft Lemma 1.17 ebenfalls vollständig unter der durch die Einschränkung des Skalarprodukts definierten Norm. Wir bezeichnen abgeschlossene Untervektorräume von Hilberträumen daher auch als Unterhilberträume oder kurz als Unterräume des gegebenen Hilbertraums.

Für die Summanden einer orthogonalen Summe gilt mittels der oben angebenen Inklusionen sogar noch mehr:

$$\langle (v_1, 0_{V_2}) | (0_{V_1}, v_2)\rangle = \langle v_1, 0\rangle_1 + \langle 0, v_2\rangle = 0 \qquad \text{für alle } v_1 \in V_1, v_2 \in V_2.$$

Definition 2.11. *Zwei Unterräume U_1, U_2 eines Hilbertraums $(V, \langle\cdot|\cdot\rangle)$ heißen orthogonal, geschrieben $U_1 \perp U_2$, falls*

$$\langle v_1|v_2\rangle = 0 \qquad \textit{für alle } v_1 \in U_1, v_2 \in V_2$$

gilt.

Die obige Identifizierung der Summanden V_1, V_2 mit den Unterräumen

$$\tilde{V}_1 := \iota_1(V_1) = \{(v_1, 0_{V_2}) : v_1 \in V_1\} \quad \text{bzw.} \quad \tilde{V}_2 := \iota_2(V_2) = \{(0_{V_1}, v_2) : v_2 \in V_2\}$$

von $V_1 \oplus V_2$ war bisher rein formal über die Inklusionen ι_1, ι_2 gegeben. Noch ist uns nicht ganz klar, ob diese Identifizierung gerechtfertigt ist.

Definition 2.12. *Eine Abbildung $J : V \to Y$ zwischen den Hilberträumen V und Y heißt isometrischer Isomorphismus, wenn sie linear und bijektiv ist und*

$$\langle J(v)|J(w)\rangle_Y = \langle v|w\rangle_V \qquad \textit{für alle } v, w \in V$$

erfüllt. Existiert ein isometrischer Isomorphismus von V nach Y, so heißen V und Y isometrisch isomorph.

Zwei isometrisch isomorphe Hilberträume sind als Hilberträume nicht von einander zu unterscheiden. Selbstverständlich gilt durch die Bijektivität auch für die Umkehrabbildung $J^{-1} : Y \to V$ eines isometrischen Isomorphismus

$$\langle J^{-1}(x)|J^{-1}(y)\rangle_V = \langle x|y\rangle_Y,$$

also ist auch J^{-1} ein isometrischer Isomorphismus ist.
Nach unserer bisherigen Diskussion ist sofort klar, dass $\iota_i : V_i \to \tilde{V}_i$, $i = 1, 2$, isometrische Isomorphismen sind.
Wir merken uns also, dass die Summanden V_1, V_2 einer orthogonalen Summe $V_1 \oplus V_2$ (isometrisch isomorph zu und damit nicht unterscheidbar von den) Unterhilberträumen $\tilde{V}_1, \tilde{V}_2$ sind, die ihrerseits orthogonal sind und die Summe vollständig aufspannen, da nach Konstruktion jedes $v \in V_1 \oplus V_2$ als $v = (v_1, v_2)$ mit gewissen $v_1 \in V_1$, $v_2 \in V_2$ gegeben ist. Wir werden also ab sofort auf die Unterscheidung von V_i und $\tilde{V}_i$ verzichten und einfach V_i als Unterraum von $V_1 \oplus V_2$ auffassen.
Induktiv können wir nun auch endliche orthogonale Summen mehrerer Hilberträume definieren. Mit etwas Arbeit geht sogar noch mehr.

Definition 2.13. *Für $n \in \mathbb{N}$ seien die Hilberträume $(V_n, \langle\cdot|\cdot\rangle_n)$ über dem gleichen Körper $\mathbb{K}$ gegeben. Wir setzen*

$$V := \left\{ (v_n)_{n\in\mathbb{N}} : v_n \in V_n \text{ für alle } n \text{ und } \sum_{n=1}^{\infty} \langle v_n|v_n\rangle < \infty \right\}.$$

Auf V definieren wir das Skalarprodukt

$$\langle (v_n)_n|(w_n)_n\rangle := \sum_{n=1}^{\infty} \langle v_n|w_n\rangle_n$$

und nennen $V = \bigoplus_{n=1}^{\infty} V_n$ die orthogonale Summe der V_n.

Beispiel 16. *Im Fall $V_n = \mathbb{K}^1 (= \mathbb{K})$, jeweils versehen mit dem (Skalar-)produkt $\langle v|w\rangle_n = v^* w$, ist das genau die Definition von $\ell_2^{\mathbb{K}}$.*

Bemerkung. *Tatsächlich ist das so definierte V nach Lemma 2.2 ein Vektorraum. Dass $\langle\cdot|\cdot\rangle$ ein Skalarprodukt ist, rechnen wir ganz analog zum ℓ_2 nach. Und auch die Vollständigkeit folgt exakt wie beim ℓ_2 (dort aus der Vollständigkeit von $\mathbb{K}$ unter dem Betrag, hier aus der Vollständigkeit der V_n unter den von $\langle\cdot|\cdot\rangle_n$ definierten Normen). Die Definition ergibt also Sinn, denn $(V, \langle\cdot|\cdot\rangle)$ ist in der Tat ein Hilbertraum.*

Wieder können wir jedes der V_n per isometrischer Isomorphie mit einem Unterhilbertraum der abzählbaren orthogonalen Summe identifizieren.
Wir kommen nun zur umgekehrten Frage: Finden wir zu jedem Unterhilbertraum U eines Hilbertraums V stets einen Unterraum W derart, dass $U \perp W$ und $V = U \oplus W$ ist? Die Antwort wird der erste Satz von Riesz liefern, für den wir aber noch ein paar Vorbereitungen brauchen.

Zur Erinnerung wiederholen wir noch einmal das Konzept einer konvexen Teilmenge eines Vektorraums.

Definition 2.14. *Eine Teilmenge $A \subset V$ eines $\mathbb{K}$-Vektorraums heißt konvex, wenn für alle $v, w \in A$ und $\lambda \in [0,1]$ die Konvexkombination $\lambda v + (1-\lambda)w$ auch in A liegt.*

Lassen wir λ alle Werte in $[0,1]$ durchlaufen, erhalten wir die griffige Formulierung, dass A genau dann konvex ist, wenn für je zwei Punkte in A auch deren Verbindungsstrecke in A liegt.

Beispiel 17. (*a*) *Wir haben uns schon davon überzeugt, dass Kugeln in normierten linearen Räumen konvex sind.*

(*b*) *Eine besonders wichtige Klasse konvexer Teilmengen sind lineare Teilräume eines gegebenen Vektorraums. Die Abgeschlossenheit des Unterraums unter den Vektorraumoperationen liefert $\lambda v + (1-\lambda)w \in A$ sogar für beliebige $\lambda \in \mathbb{K}$ und $v, w \in A$.*

Lemma 2.15. *Sind $A \subset V$ nicht leer, konvex und abgeschlossen im Hilbertraum $(V, \langle\cdot|\cdot\rangle)$ und $v \in V$ beliebig, so existiert ein eindeutig bestimmtes $w_0 \in A$ derart, dass*

$$\|v - w_0\| = \operatorname{dist}(v, A) \equiv \inf_{w \in A} \|v - w\|.$$

Wir finden also zu gegebenem $v \in V$ genau ein $w_0 \in A$, das unter allen Punkten in A am nächsten an v liegt.

Beweis. Für $v \in A$ ist einfach $w_0 = v$.
Wir betrachten nun $v \in V \backslash A$. Nach Definition des Infimums finden wir eine Folge $(w_n)_n$ in A derart, dass

$$\|v - w_n\| \to \operatorname{dist}(v, A) \qquad \text{für } n \to \infty$$

gilt. Wir schreiben kurz $d(v) := \operatorname{dist}(v, A)$.
(I) $(w_n)_n$ ist Cauchyfolge:

$$\begin{aligned} 0 \le \|w_n - w_m\|^2 &= \|(w_n - v) + (v - w_m)\|^2 \\ &\overset{(2.1)}{=} 2\|w_n - v\|^2 + 2\|v - w_m\|^2 - \|2v - w_n - w_m\|^2 \\ &= 2\underbrace{\|w_n - v\|^2}_{\overset{n\to\infty}{\longrightarrow} d(v)^2} + 2\underbrace{\|w_m - v\|^2}_{\overset{m\to\infty}{\longrightarrow} d(v)^2} - 4\underbrace{\Big\|v - \underbrace{\tfrac{1}{2}(w_m + w_n)}_{\in A}\Big\|^2}_{\ge d(v)^2} \end{aligned}$$

Für hinreichend große m, n ist die rechte Seite also kleiner als jedes beliebig vorgegebene $\varepsilon > 0$, also ist $(w_n)_n$ tatsächlich eine Cauchyfolge.

(II) Der Grenzwert $w_0 := \lim_{n\to\infty} w_n$ der Cauchyfolge $(w_n)_n$ liegt in der abgeschlossenen Menge A und erfüllt wegen der Stetigkeit der Norm:

$$\|w_0 - v\| = \lim_{n\to\infty} \|w_n - v\| = \operatorname{dist}(v, A),$$

wie behauptet.

(III) Wir haben noch die Eindeutigkeit von w_0 zu zeigen. Sei also u_0 ein weiterer Vektor in A mit $\|v - u_0\| = d(v)$. Dann gilt wieder mittels Parallelogrammgleichung:

$$
\begin{aligned}
0 &\le \|w_0 - v + v - u_0\|^2 \\
&= 2\|w_0 - v\|^2 + 2\|v - u_0\|^2 - 4\left\|v - \tfrac{1}{2}(u_0 + w_0)\right\|^2 \\
&\le 2d(v)^2 + 2d(v)^2 - 4d(v)^2 \;=\; 0,
\end{aligned}
$$

also $\|w_0 - u_0\| = 0$, d.h., $u_0 = w_0$. □

Das folgende Lemma passt hier gut als Übungsaufgabe, wird aber erst später benutzt.

Lemma 2.16. *Ist $C \subset V$ eine konvexe Teilmenge eines normierten linearen Raums V, so sind der Abschluss $\bar{C}$ und für jedes $r > 0$ die r-Umgebung*

$$U_r(C) = \{v \in V : \|v - w\| < r \text{ für ein } w \in C\} \equiv \bigcup_{w \in C} B_r(w)$$

von C auch konvex.

Beweis. Übung. □

Als Folgerung erhalten wir den ersten Satz von Riesz. Zunächst führen wir noch eine Notation ein.

Definition 2.17. *Zur Teilmenge $A \subset V$ eines Hilbertraums $(V, \langle\cdot|\cdot\rangle)$ bezeichnet*

$$A^\perp := \{w \in V : \langle w|v\rangle = 0 \text{ für alle } v \in A\}$$

das orthogonale Komplement von A.

Streng genommen handelt es sich um das orthogonale Komplement des von A aufgespannten linearen Unterraums. Für $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ mit beliebigen $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ und $v_1, \ldots, v_n \in A$ sowie $w \in A^\perp$ gilt nämlich

$$\langle w|v\rangle = \sum_{k=1}^{n} \alpha_k \langle w|v_k\rangle = 0.$$

Lemma 2.8 besagt, dass das orthogonale Komplement einer dichten Teilmenge $M \subset V$ trivial ist: $M^\perp = \{0_V\}$. Überzeugen Sie sich, dass diese Aussage schon in der Einheitssphäre von V dichte Teilmengen M gilt.
Wir formulieren den folgenden Satz nur für abgeschlossene Teilräume, also Unterhilberträume.

Satz 2.18 (1. Satz von Riesz)**.** *Für jeden Unter(hilbert)raum V_1 des Hilbertraums $(V, \langle\cdot|\cdot\rangle)$ ist $V_1^\perp$ ein Unterhilbertraum von V, und es gilt $V_1 \oplus V_1^\perp = V$.*

Die letzte Aussage bedeutet, dass jedes $v \in V$ eindeutig als $v = v_1 + v_2$ mit $v_1 \in V_1$ und $v_2 \in V_1^\perp$ geschrieben werden kann. Spätestens dafür brauchen wir die Vollständigkeit von V_1.

Beweis. (I) Existenz der Zerlegung. V_1 ist abgeschlossen und konvex, also finden wir zu $v \in V$ ein eindeutig bestimmtes

$$v_1 =: \pi_1(v) \in V_1 \qquad \text{mit} \qquad \|v_1 - v\| = \operatorname{dist}(v, V_1).$$

Für beliebige $\alpha \in \mathbb{K}$ und $u_1 \in V_1$ gilt dann

$$\|v - v_1\| = \inf_{w \in V_1} \|v - w\| \leq \|v - (v_1 + \alpha u_1)\|.$$

Daraus folgt (Übung!) $v_2 := v - v_1 \perp u_1$ für beliebige $u_1 \in V_1$, also $v_2 \in V_1^\perp$. Nach Definition von v_2 gilt natürlich auch $v = v_1 + v_2$.

(II) Eindeutigkeit. Sind $\tilde{v}_1 \in V_1, \tilde{v}_2 \in V_1^\perp$ weitere Vektoren mit $v_1 + v_2 = v = \tilde{v}_1 + \tilde{v}_2$, so erhalten wir für

$$u := \underbrace{v_1 - \tilde{v}_1}_{\in V_1} = \underbrace{\tilde{v}_2 - v_2}_{\in V_1^\perp}$$

$\langle u|u\rangle = 0$, also $u = 0$. Das ist nur der Fall, wenn $v_1 = \tilde{v}_1$ und $v_2 = \tilde{v}_2$ sind. □

Einen ganz ähnlichen Satz für allgemeine Banachräume werden wir in einer Übungsaufgabe kennen lernen und daraus eine Verallgemeinerung des Satzes von Heine und Borel auf allgemeine endlichdimensionale normierte Vektorräume schließen.
Die Aussage und der Beweis des Satzes motivieren folgende Notation.

Definition 2.19. *Sind V_1, V_2 Unterräume des Hilbertraums $(V, \langle\cdot|\cdot\rangle)$ mit $V_2 = V_1^\perp$ (also $V = V_1 \oplus V_2$), so bezeichnen wir zu gegebenem $v \in V$ die eindeutig bestimmten $\pi_1(v) := v_1 \in V_1$, $\pi_2(v) := v_2 \in V_2$ mit $v = v_1 + v_2$ als orthogonale Projektionen von v auf V_1 bzw. V_2.*

Lemma 2.20. *Für den Unterraum V_1 des Hilbertraums V sind die orthogonalen Projektionen π_1 auf V_1 und π_2 auf $V_2 = V_1^\perp$ lineare Abbildungen mit $\operatorname{im}\pi_1 = \ker\pi_2 = V_1$ und $\operatorname{im}\pi_2 = \ker\pi_1 = V_2$.*

Beweis. (I) Zur Linearität haben wir zu $v, w \in V$ und $\alpha \in \mathbb{K}$

$$\pi_1(\alpha v + w) = \alpha v_1 + w_1$$

zu zeigen. Wir müssen als dank der Eindeutigkeit nur zeigen, dass $\alpha v_1 + w_1 (\in V_1)$ den Abstand von $\alpha v + w$ zu V_1 minimiert. Sei dazu $u \in V_1$ beliebig. Wegen $\alpha v_2 + w_2 \in V_1^\perp$ liefert der Pythagoras in der Tat:

$$\begin{aligned} \|u - (\alpha v + w)\|^2 &= \|u - (\alpha v_1 + w_1) - (\alpha v_2 + w_2)\|^2 \\ &= \|u - (\alpha v_1 + w_1)\|^2 + \|\alpha v_2 + w_2\|^2 \\ &= \underbrace{\|u - (\alpha v_1 + w_1)\|^2}_{\geq 0} + \|\alpha v + w - (\alpha v_1 + w_1)\|^2. \end{aligned}$$

(II) Die Aussagen zu Kern und Bild sind leicht gezeigt:

$$\ker \pi_1 = \{v \in V : \pi_1(v) = 0\} = \{v \in V : v = 0 + v_2 \text{ für ein } v_2 \in V_1^\perp\} = V_1^\perp.$$

Die anderen Aussagen dieses Typs ergeben sich ganz analog. $\square$

Beispiel 18. *Wir betrachten den Raum $V = L_2([0,\pi];\mathbb{K})$ und als V_1 den Unterraum der linearen Funktionen. V_1 wird also von den Funktionen $\mathbf{1}_{[0,\pi]}$ und $\mathrm{id}_{\mathbb{R}}$ aufgespannt:*

$$V_1 = \left\{\alpha \mathbf{1}_{[0,\pi]} + \beta \mathrm{id} : \alpha, \beta \in \mathbb{K}\right\}.$$

Wir wollen die orthogonale Projektion von $v = \sin$ auf V_1 berechnen. Wir haben also $\| \sin - \alpha \mathbf{1}_{[0,\pi]} - \beta \mathrm{id}\|_2$ bezüglich α und β zu minimieren. Wir minimieren das Quadrat:

$$\begin{aligned} F(\alpha,\beta) &= \| \sin - \alpha \mathbf{1}_{[0,\pi]} - \beta \mathrm{id}\|_2^2 \\ &= \int_0^\pi \sin^2 t + \alpha^2 + \beta^2 t^2 - 2\alpha \sin t - 2\beta t \sin t + 2\alpha\beta t \,\mathrm{d}t \\ &= \frac{\pi}{2} + \alpha^2\pi + \frac{\beta^2}{3}\pi^3 - 4\alpha - 2\beta\pi + \alpha\beta\pi^2. \end{aligned}$$

Wir berechnen

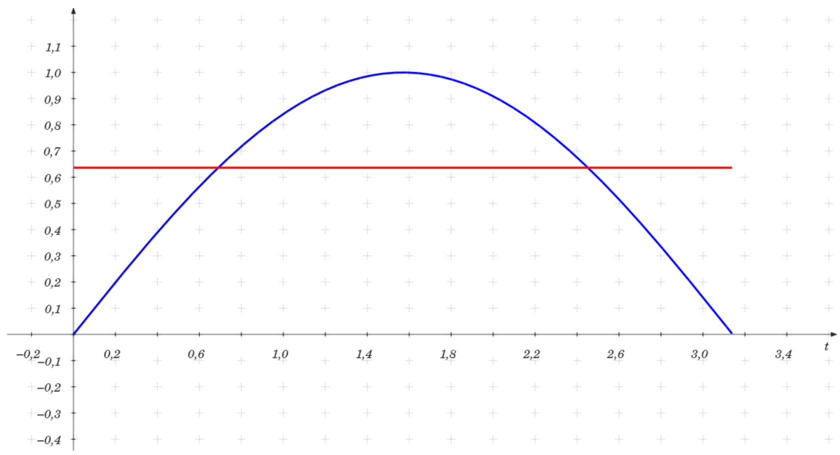
$$DF = \begin{pmatrix} 2\alpha\pi - 4 + \beta\pi^2 \\ \frac{2}{3}\pi^3\beta - 2\pi + \alpha\pi^2 \end{pmatrix},$$

lösen das lineare Gleichungssystem $DF = 0$ nach $\beta = 0$, $\alpha = \frac{2}{\pi}$ auf und überzeugen uns, dass die Hessesche

$$D^2F = \begin{pmatrix} 2\pi & \pi^2 \\ \pi^2 & \frac{2}{3}\pi^3 \end{pmatrix}$$

(unabhängig von α und β) positiv definit ist (da $a_{1,1} = 2\pi > 0$ und $\det A = \frac{1}{3}\pi^4 > 0$). Bei den so berechneten α, β finden wir also ein Minimum vor:

$$v_1(t) = \frac{2}{\pi}.$$



Das war einiges an Rechnung (und man stelle sich nur vor, der Unterraum V_1 wäre von höherer Dimension als 2). Wir werden in Kürze einen deutlich einfacheren Weg, orthogonale Projektionen auf abgeschlossene Unterräume zu berechnen, kennen lernen.

2.3 Stetige lineare Funktionale, zweiter Satz von Riesz

Wir beginnen mit dem Begriff des (linearen) Funktionals, der namensgebend für den Titel unserer Vorlesung ist. Diese definieren wir zunächst auf beliebigen normierten Vektorräumen.

Definition 2.21. 1. *Ein lineares Funktional auf einem $\mathbb{K}$-Vektorraum V ist eine ($\mathbb{K}$-)lineare Abbildung $\phi : V \supset \mathcal{D}_\phi \to \mathbb{K}$:*

$$\phi(\alpha v + w) = \alpha\phi(v) + \phi(w) \qquad \text{für alle } \alpha \in \mathbb{K}, v, w \in \mathcal{D}_\phi.$$

Dabei ist der Definitionsbereich $\mathcal{D}_\phi$ von ϕ ein linearer Unterraum von V.

2. *Ein sublineares Funktional auf V ist eine Abbildung $p : V \supset \mathcal{D}_p \to \mathbb{R}$ derart, dass für alle $v, w \in \mathcal{D}_p$ und $\alpha > 0$*

$$p(v + w) \leq p(v) + p(w) \qquad \text{und} \qquad p(\alpha v) \leq \alpha p(v)$$

gelten.

3. *Ist $(V, \|\cdot\|)$ ein normierter Vektorraum, so nennen wir ein lineares Funktional ϕ auf V stetig, falls es auf ganz V definiert und bezüglich der durch $\|\cdot\|$ erzeugten Topologie (folgen-)stetig ist:*

$$\|v_n - v\| \to 0 \implies |\phi(v_n) - \phi(v)| \to 0.$$

Bemerkung. *Da $(V, \|\cdot\|)$ und $(\mathbb{K}, |\cdot|)$ insbesondere metrische Räume sind, hätten wir völlig äquivalent auch ε-δ-Stetigkeit oder schlicht Stetigkeit bezüglich der von $\|\cdot\|$ auf V und von $|\cdot|$ auf $\mathbb{K}$ induzierten Topologien fordern können.*

Überzeugen Sie sich zur Übung, dass für ein sublineares Funktional sogar $p(\alpha v) = \alpha p(v)$ für $v \in \mathcal{D}_p$, $\alpha \geq 0$ gilt.
Man beachte, dass wir nicht von der Stetigkeit in einem Punkt v_0 gesprochen haben. Das ist kein Zufall.

Lemma 2.22. *Ein lineares Funktional ist entweder in jedem oder in keinem $v_0 \in V$ stetig.*

Beweis. Ist ϕ in irgend einem v_0 stetig, so gilt für jedes $w_0 \in V$ und jede Folge $(w_n)_n$ mit $w_n \to w_0$

$$\begin{aligned} |\phi(w_n) - \phi(w_0)| &= \big|\phi(w_n) - \phi(w_0 - v_0) + \phi(w_0 - v_0) - \phi(w_0)\big| \\ &= \big|\phi(\underbrace{w_n - w_0 + v_0}_{=:v_n \to v_0}) - \phi(v_0)\big| \to 0 \qquad \text{für } n \to \infty. \end{aligned}$$

Dabei haben wir $\|v_n - v_0\| = \|w_n - w_0 + v_0 - v_0\| = \|w_n - w_0\| \to 0$ und zuvor die Linearität von ϕ benutzt. □

Insbesondere genügt uns also die Stetigkeit in der Null, um auf die globale Stetigkeit eines linearen Funktional zu schließen.

Beispiel 19. *Als prototypisches Beispiel eines global definierten sublinearen Funktionals auf* $(V, \|\cdot\|)$ *dient die Norm selbst:* $p(v) = \|v\|$. *Die Stetigkeit ist hier auch klar, und die Bedingungen für die Sublinearität sind einfach die Dreiecksungleichung und die Homogenität der Norm.*
Jede Halbnorm ist ebenfalls ein sublineares Funktional, eine Quasinorm (z.B. $p = \|\cdot\|_1$ *auf* $\ell_\infty \supsetneqq \mathcal{D}_p = \ell_1$*) wäre ein Beispiel für ein nicht global definiertes sublineares Funktional.*
Umgekehrt überzeugen Sie sich schnell, dass ein stetiges lineares Funktional $\phi : V \to \mathbb{K}$ *via* $p(v) := |\phi(v)|$ *ein sublineares Funktional und sogar eine Halbnorm definiert.*

Beispiel 20. *Auf* $V = L_1([0,1])$ *ist*

$$\phi(v) = \int_{[0,1]} v(t)\,\mathrm{d}t$$

ein stetiges lineares Funktional, wie Sie schnell nachrechnen, wenn Ihnen diese Aussage aus der Maßtheorie nicht ohnehin schon geläufig ist.

Die global definierten linearen Funktionale bilden unter punktweisen Operationen ihrerseits wieder einen $\mathbb{K}$-Vektorraum:

$$(\alpha\phi + \psi)(v) := \alpha\phi(v) + \psi(v) \qquad \text{für } v \in V.$$

Definition 2.23. *Der Raum der linearen Funktionale* $\phi : V \to \mathbb{K}$ *auf einem* $\mathbb{K}$-*Vektorraum* V *wird als* V^* *bezeichnet und heißt* *(algebraischer) Dualraum von* V.
Für einen normierten Vektorraum bezeichnen wir mit V' *den Raum der stetigen linearen Funktionale auf* V *und nennen* V' *der* *(topologischen) Dualraum von* V.

Sie überzeugen sich schnell, dass Summen und skalare Vielfache von stetigen linearen Funktionalen stetig sind und folgern, dass V' ein linearer Unterraum von V^* ist.
Wir wissen aus der Analysis 2, dass lineare Abbildungen $\phi : \mathbb{R}^n \to \mathbb{R}$ zwingend stetig sind. Das gilt für beliebige endlich-dimensionale Räume, und zwar unabhängig von der Norm[2]. Damit gibt es für endlichdimensionale normierte Vektorräume keinen Unterschied zwischen algebraischem und topologischem Dualraum. Da drängt sich die Frage auf, wie ein unstetiges lineares Funktional aussehen sollte. Gibt es derartiges überhaupt?

Beispiel 21. *Wir betrachten den Raum* $V = C^1([-\pi,\pi];\mathbb{R})$ *der stetig differenzierbaren*[3] *Funktionen* $v : [-\pi,\pi] \to \mathbb{R}$, *versehen mit der Norm*

$$\|v\|_\infty = \max_{-\pi \le t \le \pi} |v(t)|.$$

Da stetig differenzierbare Funktionen insbesondere stetig und Summen sowie Vielfache von stetig differenzierbaren Funktionen stetig differenzierbar sind, ist V *natürlich ein linearer Unterraum des Banachraums* $(C([-\pi,\pi];\mathbb{R}), \|\cdot\|_\infty)$, *also tatsächlich ein normierter Vektorraum. Auf* V *betrachten wir nun das lineare Funktional*

$$\phi : V \to \mathbb{R}, \quad \phi(v) = v'(0).$$

[2] Wir erinnern uns, dass auf endlichdimensionalen Räumen alle Normen äquivalent sind.
[3] In $t = 0$ bzw. $t = \pi$ genügt uns die einseitige Ableitung.

Das ist auf ganz V definiert und offensichtlich linear.
Nun betrachten wir die durch

$$v_n = \frac{1}{\sqrt{n}} \sin(n\cdot), \qquad \textit{also } v_n(t) = \frac{1}{\sqrt{n}} \sin(nt)$$

gegebene Folge $(v_n)_n$ in V, für die offensichtlich

$$\|v_n - 0_V\|_\infty = \max_{-\pi \le t \le \pi} |v_n(t) - 0| = \frac{1}{\sqrt{n}} \to 0,$$

also gilt $v_n \to v = 0_V$ in V. Allerdings ist

$$\phi(v_n) = \frac{n}{\sqrt{n}} \cos(nt)\big|_{t=0} = \sqrt{n} \to \infty, \qquad \textit{aber } \ \phi(v) = 0.$$

Wir bemerken, dass es sich bei V unter der angegebenen Norm nicht um einen Banachraum handeln kann, da V den Raum der Polynome enthält, welcher seinerseits bereits dicht in $C([-\pi, \pi]; \mathbb{R})$ liegt.
Stetige lineare Funktionale können wir auch anders erkennen.

Lemma 2.24. *Ein lineares Funktional $\phi : V \to \mathbb{K}$ auf einem normierten $\mathbb{K}$-Vektorraum $(V, \|\cdot\|)$ ist genau dann stetig, wenn ein $c \geq 0$ derart existiert, dass für alle $v \in V$*

$$|\phi(v)| \leq c\|v\|$$

gilt.

Beweis. Haben wir ein solches c gefunden, so folgt aus der Bedingung die Stetigkeit in der Null:

$$|\phi(v_n) - \underbrace{\phi(0_V)}_{=0}| \leq c\|v_n\| \to 0,$$

falls $v_n \to 0_V$, also ist ϕ dann überall stetig.
Ist umgekehrt ϕ in 0_V stetig, so existiert zu $\varepsilon = 1 > 0$ ein $\delta > 0$ derart, dass $|\phi(v) - 0| < 1$ für alle $v \in V$ mit $\|v - 0_V\| < \delta$ gilt. Für alle $w \in V \backslash \{0\}$ gilt dann mit $c := \frac{2}{\delta}$:

$$|\phi(w)| = \left|\phi\left(\frac{2\|w\|}{\delta}\frac{\delta}{2\|w\|}w\right)\right| = \left|\frac{2\|w\|}{\delta}\right|\left|\phi\Big(\underbrace{\frac{\delta}{2\|w\|}w}_{=v,\ \|v\|=\frac{\delta}{2}}\Big)\right| < \frac{2}{\delta}\|w\|1 = c\|w\|.$$

□

Man spricht im Rahmen der Funktionalanalysis daher auch gern von beschränkten statt stetigen linearen Funktionalen und meint mit beschränkt, dass ein c wie in der Aussage des Lemmas existiert.

Definition 2.25. *Auf dem Vektorraum V' zu einem normierten Vektorraum $(V, \|\cdot\|_V)$, $V \neq \{0\}$, definieren wir die Operatornorm $\|\cdot\|_{V'}$ durch*

$$\|\phi\|_{V'} = \sup\{|\phi(v)| : v \in V, \|v\|_V = 1\}.$$

Bemerkung. *Dass es sich tatsächlich um eine Norm handelt, rechnen wir schnell nach. Zunächst ist die Norm endlich, denn zu gegebenem ϕ gilt $\|\phi\|_{V'} \leq c$ für jedes c, das die Bedingung des Lemmas erfüllt*[4]*.*
Ist $\phi = 0_{V'}$ das Nullfunktional, so ist $\|\phi\|_{V'}$ das Supremum über viele Nullen, also $\|0_{V'}\| = 0$. Auch sonst nehmen wir stets das Supremum über nicht negative Zahlen, also ist $\|\phi\|_{V'} \geq 0$ für alle ϕ.
Sei umgekehrt $\|\phi\|_{V'} = 0$, also $\phi(v) = 0$ für alle $v \in V$ mit $\|v\|_V = 1$. Zu jedem $w \in V$ finden wir dann ein $v \in V$ mit $\|v\| = 1$ und ein $\alpha \in \mathbb{K}$ mit $w = \alpha v$ und berechnen

$$\phi(w) = \phi(\alpha v) = \alpha\phi(v) = 0,$$

also $\phi = 0_{V'}$.
Ist nun neben $\phi \in V'$ auch $\alpha \in \mathbb{K}$ gegeben, so ergibt sich die Homogenität aus

$$|\phi(\alpha v)| = |\alpha\phi(v)| = |\alpha||\phi(v)|.$$

Nehmen wir nun das Supremum über alle v in der Einheitssphäre und beachten noch, dass beide Faktoren auf der rechten Seite nicht negativ sind, können wir $|\alpha|$ aus dem Supremum herausziehen und erhalten die Behauptung.
Auch die Dreiecksungleichung erbt die Norm vom Betrag in $\mathbb{K}$. Sind nämlich $\phi, \psi \in V'$, so berechnen wir wieder für jedes $v \in V$:

$$|(\phi + \psi)(v)| \leq |\phi(v)| + |\psi(v)|.$$

Wieder können überträgt sich diese Ungleichung auf das Supremum über alle v mit $\|v\|_V = 1$, da sämtliche Summanden nicht negativ sind.

Wir kommen nun zu unseren Hilberträumen zurück und stellen fest, dass jedes Element eines Hilbertraums auf natürliche Art ein stetiges lineares Funktional erzeugt.

Beispiel 22. *Sind $(V, \langle\cdot|\cdot\rangle)$ ein Hilbertraum und $v \in V$ beliebig gegeben, so ist*

$$\phi_v : V \to \mathbb{K}, \qquad \phi_v(w) := \langle v|w\rangle$$

ein stetiges lineares Funktional auf V. Die Linearität folgt sofort aus der Linearität des Skalarprodukts im zweiten Argument, und die Stetigkeit haben wir auch schon gesehen:

$$|\phi_v(w_n) - \phi_v(w)| = |\langle v|\underbrace{w_n - w}_{\to 0}\rangle| \to 0.$$

Tatsächlich gilt sogar die Umkehrung.

Satz 2.26 (2. Satz von Riesz, Darstellungssatz von Riesz-Fréchet)**.** *Ist ϕ ein stetiges lineares Funktional auf dem Hilbertraum $(V, \langle\cdot|\cdot\rangle)$, so existiert genau ein $v \in V$ mit $\phi = \phi_v$, also*

$$\phi(w) = \langle v|w\rangle \qquad \text{für alle } w \in V.$$

[4]Genauer ist nämlich $\|\phi\|_{V'}$ das Infimum über alle solchen c.

Beweis. (I) Existenz. Wir betrachten linearen Unterraum $V_1 := \ker\phi = \{w \in V : \phi(w) = 0\}$. Dieser ist (als Urbild der in $\mathbb{K}$ abgeschlossenen Menge $\{0\}$ unter der nach Voraussetzung stetigen Abbildung ϕ) abgeschlossen[5], also ein Unterhilbertraum von V. Nach Satz 2.18 ist $V = V_1 \oplus V_1^\perp$.
Falls $V_1^\perp = \{0_V\}$, so ist $V_1 = V$, also $\phi(w) = 0$ für alle w. Dann ist $v = 0$ das gesuchte ϕ darstellende Element.
Ist $V_1^\perp \neq \{0\}$, so wählen wir $u \in V_1^\perp \backslash \{0\}$ und haben $\phi(u) \neq 0$, denn sonst wäre auch $u \in \ker\phi = V_1$ und dann $\langle u|u\rangle = 0$, also $u = 0$.
Für beliebige $w \in V$ gilt nun

$$\phi\left(w - \frac{\phi(w)}{\phi(u)}u\right) = \phi(w) - \frac{\phi(w)}{\phi(u)}\phi(u) = 0,$$

also $w - \frac{\phi(w)}{\phi(u)}u \in V_1$. Wegen $u \in V_1^\perp$ folgt dann aber

$$0 = \left\langle u \middle| w - \frac{\phi(w)}{\phi(u)}u \right\rangle = \langle u|w\rangle - \frac{\phi(w)}{\phi(u)}\langle u|u\rangle,$$

also:

$$\phi(w) = \frac{\phi(u)}{\|u\|^2}\langle u|w\rangle = \left\langle \frac{\phi(u)}{\|u\|^2}u \middle| w \right\rangle.$$

Dabei war $w \in V$ beliebig. Mit $v := \frac{\phi(u)}{\|u\|^2}u$ haben wir unser gesuchtes darstellendes Element also gefunden.

(II) Zur Eindeutigkeit nehmen wir an, es gäbe $v, \tilde{v} \in V$ mit der gewünschten Eigenschaft. Dann gälte für alle $w \in V$:

$$0 = \phi(w) - \phi(w) = \langle v|w\rangle - \langle \tilde{v}|w\rangle = \langle v - \tilde{v}|w\rangle,$$

also insbesondere für $w = v - \tilde{v}$: $\langle v - \tilde{v}|v - \tilde{v}\rangle = 0$, also $v = \tilde{v}$. □

Bemerkung. *Tatsächlich gilt umgekehrt: Kann jedes stetige lineare Funktional ϕ auf dem Prähilbertraum $(V, \langle\cdot|\cdot\rangle)$ als $\phi = \langle v|\cdot\rangle$ mit einem $v \in V$ dargestellt werden, so ist V schon ein Hilbertraum. Überzeugen Sie sich davon, indem Sie für $y \in \tilde{V}$ (Vervollständigung von V, versehen mit Skalarprodukt $(\cdot|\cdot)$) zeigen, dass $y \in V$ gilt. Erinnern Sie sich daran, dass y eine Äquivalenzklasse von Cauchyfolgen in V ist.*
Gelingt es uns also, für einen Prähilbertraum ein stetiges lineares Funktional zu finden, das nicht durch ein Element von V darstellbar ist, so haben wir gezeigt, dass der fragliche Raum nicht vollständig ist.

Beispiel 23. *Alle stetigen linearen Funktionale auf $V = L_2(\Omega; \mathbb{K})$ sind von der Form*

$$\phi(w) = \int_\Omega v(x)^*\, w(x)\,\mathrm{d}x \qquad \textit{für alle } w \in L_2$$

mit einem eindeutigen $v \in L_2(\Omega; \mathbb{K})$.

[5] Überzeugen Sie sich zur Übung, dass auch die Umkehrung gilt: Ist für $\phi \in V^*$ der Kern $\ker\phi$ abgeschlossen in V, so ist ϕ stetig. Und andererseits ist $\ker\phi$ genau dann dicht in V, wenn ϕ unbeschränkt ist.

Wir haben also für jeden Hilbertraum V eine Bijektion $J : V \to V'$ gefunden, die durch $J(v) = \langle v|\cdot\rangle$ gegeben ist. Diese ist offensichtlich konjugiert linear: $J(\alpha v + w) = \alpha^* J(v) + J(w)$. In gewissem Sinn ist also jeder Hilbertraum sein eigener Dualraum.

Tatsächlich kann man den $\mathbb{K}$-Vektorraum $V^\dagger$ betrachten, der die gleichen Elemente wie V enthält, dessen Addition mit der in V übereinstimmt, und dessen Multiplikation mit Skalaren durch

$$\alpha \bullet v := \underbrace{\alpha^* v}_{\text{berechnet in } H} \quad , \quad \alpha \in \mathbb{K}, v \in V^\dagger .$$

Damit wird $\tilde{J} : V^\dagger \to V'$ zu einem Vektorraumisomorphismus.
Ist $\mathbb{K} = \mathbb{R}$, braucht man diese Feinheiten natürlich nicht, sondern hat direkt einen Isomorphismus $J : V \to V'$.

Wir können sogar noch mehr aussagen.

Korollar 2.27. *Auf dem Raum V' der stetigen linearen Funktionale auf dem Hilbertraum $(V, \langle\cdot|\cdot\rangle)$ ist durch*

$$(\phi_v|\phi_u) := \langle u|v\rangle$$

ein Skalarprodukt $(\cdot|\cdot)$ gegeben, das die Operatornorm $\|\cdot\|_{V'}$ erzeugt.

Beweis. Übung (man rechne die Skalarprodukteigenschaften nach und zeige $\|\phi_v|_{V'} = \sqrt{(\phi_v|\phi_v)}$ für alle $v \in V$). □

Der oben beschriebene Isomorphismus $\tilde{J}$ ist also sogar ein isometrischer Isomorphismus. Als Hilbertraum ist V also (bis auf die etwas merkwürdige Multiplikation mit Skalaren) nicht von seinem Dualraum V' zu unterscheiden.

2.4 Fourierentwicklung in Hilberträumen

Wir kommen nun zu unseren Orthogonal- bzw. Orthonormalsystemen zurück und erinnern uns, dass wir im $\mathbb{K}^n$ jeden Vektor v als Linearkombination

$$v = \sum_{k=1}^{n} v^k e_k$$

schreiben können, wobei $v^k \in \mathbb{K}$ die Koordinaten des Vektors v in der Basis $(e_1, \ldots, e_n)$ aus Standardeinheitsvektoren sind, welche wir durch $v^k = \langle e_k|v\rangle$ berechnen.
Das versuchen wir nun zu verallgemeinern. Das Konzept von Orthogonal- und Orthonormalsystemen hatten wir schon eingeführt und auch an den Satz des Pythagoras erinnert: Sind $v_1, \ldots, v_n \in V$ paarweise orthogonale Vektoren in einem Hilbertraum V, so gilt

$$\|v_1 + \ldots + v_n\|^2 = \|v_1\|^2 + \ldots + \|v_n\|^2.$$

Satz 2.28. *Ist $\{v_n\}_{n\in\mathbb{N}}$ ein Orthogonalsystem im Hilbertraum V, so konvergiert genau dann die Reihe $\sum_{n=1}^{\infty} v_k$ in V, wenn die Reihe $\sum_{n=1}^{\infty} \|v_n\|^2$ in $\mathbb{R}$ konvergiert.*

Beweis. Konvergenz der Reihe $\sum_{n=1}^{\infty} v_n$ bedeutet natürlich die Konvergenz der Folge der Partialsummen $s_k := \sum_{n=1}^{k} v_n$.
(I) Wir zeigen zunächst die Konvergenz in V. Da es sich bei V um einen Hilbertraum handelt, müssen wir nur zeigen, dass $(s_k)_k$ eine Cauchyfolge ist. Dazu finden wir wegen der Konvergenz der Reihe $\sum_{n=1}^{\infty} \|v_n\|^2$ zu gegebenem $\varepsilon > 0$ ein k_0 derart, dass für alle $k \geq k_0, m \in \mathbb{N}$

$$|\sigma_{k+m} - \sigma_k| = \sum_{n=k+1}^{k+m} \|v_n\|^2 < \varepsilon^2$$

ist. Dabei haben wir mit $\sigma_k := \sum_{n=1}^{k} \|v_n\|^2$ die Partialsummen der reellen Reihe bezeichnet. Dann ist

$$\|s_{k+m} - s_k\|^2 = \left\| \sum_{n=k+1}^{k+m} v_n \right\|^2 = \sum_{n=k+1}^{k+m} \|v_n\|^2 < \varepsilon^2.$$

Dabei durften wir den verallgemeinerten Pythagoras anwenden, weil es sich um eine endliche Summe handelte. Wir haben also für $k \geq k_0$:

$$\|s_{k+m} - s_k\| < \varepsilon,$$

also bilden die Partialsummen eine Cauchyfolge, die notgedrungen in V konvergiert.

(II) Haben wir Konvergenz in V gegeben, so gehen wir umgekehrt vor, finden zu gegebenem $\varepsilon > 0$ ein k_0 derart, dass $\|s_{k+m} - s_k\| < \sqrt{\varepsilon}$ für $k \geq k_0$, $m \in \mathbb{N}$ gilt und erhalten mit den gleichen Rechnungen $|\sigma_{k+m} - \sigma_k| < \varepsilon$ für $k \geq k_0$. □

In Erinnerung an den $\mathbb{K}^n$ oder auch an klassische Fourierreihen führen wir folgende Definition ein.

Definition 2.29. *Ist $\{v_n\}_{n\in\mathbb{N}}$ ein Orthonormalsystem im Hilbertraum V, so bezeichnen wir zu gegebenem $v \in V$ die Zahlen $\langle v_n|v\rangle$, $n \in \mathbb{N}$, als Fourierkoeffizienten von v bzgl. $\{v_n\}_{n\in\mathbb{N}}$.*

Beispiel 24. *Das motivierende Beispiel ist $V = L_2((-\pi,\pi);\mathbb{C})$ mit dem Orthonormalsystem*

$$\left\{ \frac{\mathbf{1}}{\sqrt{2\pi}} \right\} \cup \left\{ \frac{\cos(n\cdot)}{\sqrt{\pi}}, \frac{\sin(n\cdot)}{\sqrt{\pi}} \right\}_{n\in\mathbb{N}}.$$

Die Fourierkoeffizienten sind dann durch

$$a_0 = \left\langle \frac{\mathbf{1}}{\sqrt{2\pi}} \middle| v \right\rangle = \frac{1}{\sqrt{2\pi}} \int_{(-\pi,\pi)} v(t)\,\mathrm{d}t,$$

$$a_n = \left\langle \frac{\cos(n\cdot)}{\sqrt{\pi}} \middle| v \right\rangle = \frac{1}{\sqrt{\pi}} \int_{(-\pi,\pi)} \cos(nt)v(t)\,\mathrm{d}t$$

und

$$b_n = \left\langle \frac{\sin(n\cdot)}{\sqrt{\pi}} \middle| v \right\rangle = \frac{1}{\sqrt{\pi}} \int_{(-\pi,\pi)} \sin(nt)v(t)\,\mathrm{d}t$$

gegeben. Da die v_n reellwertig sind, müssen wir uns hier über komplexe Konjugation keine Gedanken machen.

Bemerkung. *Um deutlich zu machen, dass es sich um einen komplexen Hilbertraum handelt, wird das Orthonormalsystem gern via Eulerscher Formel zu $\{\frac{1}{\sqrt{2\pi}}e^{\imath k\cdot}\}_{k\in\mathbb{Z}}$ umgeschrieben*[6]*. Die Normierungsfaktoren $\frac{1}{\sqrt{2\pi}}$ errechnen wir schnell aus*

$$\langle e^{\imath k\cdot}|e^{\imath k\cdot}\rangle = \int_{-\pi}^{\pi} e^{-\imath kt}e^{\imath kt}\,\mathrm{d}t = 2\pi.$$

Die Fourierkoeffizienten bezüglich eines beliebigen Orthonormalsystems erfüllen folgende Ungleichung.

Satz 2.30 (Besselsche Ungleichung)**.** *Sind $\{v_n\}_{n\in\mathbb{N}}$ ein Orthonormalsystem im Hilbertraum V und $v \in V$ beliebig, so gilt*

$$\sum_{n=1}^{\infty} |\langle v_n|v\rangle|^2 \leq \|v\|^2.$$

Selbstverständlich gilt die Ungleichung auch für endliche Orthonormalsysteme, ist dann aber bei weitem nicht mehr so spannend.

Beweis. Zu $k \in \mathbb{N}$ setzen wir $w_k = v - \sum_{n=1}^{k}\langle v_n|v\rangle v_n$ und berechnen dank der Linearität des Skalarprodukts:

$$\langle v_l|w_k\rangle = \langle v_l|v\rangle - \sum_{n=1}^{k}\langle v_n|v\rangle \underbrace{\langle v_l|v_n\rangle}_{=\delta_{l,n}} = \begin{cases} 0 & \text{für } l \leq k, \\ \langle v_l|v\rangle & \text{sonst.} \end{cases}$$

Damit können wir auf die paarweise orthogonalen $w_k, \langle v_1|v\rangle v_1, \ldots, \langle v_k|v\rangle v_k$ den verallgemeinerten Pythagoras anwenden:

$$\begin{aligned} \sum_{n=1}^{k} |\langle v_n|v\rangle|^2 \underbrace{\|v_n\|^2}_{=1} &\leq \sum_{n=1}^{k} \|\langle v_n|v\rangle v_n\|^2 + \underbrace{\|w_k\|^2}_{\geq 0} \\ &= \left\|\sum_{n=1}^{k}\langle v_n|v\rangle v_n + w_k\right\|^2 = \|v\|^2. \end{aligned}$$

Die Partialsummen der Reihe mit nicht negativen Summanden(!) sind also gegen $\|v\|^2$ beschränkt, und diese Ungleichung bleibt im Grenzübergang erhalten. $\square$

Zusammen mit Satz 2.28 ergibt das:

[6] Man beachte, dass wir komplexe Koeffizienten benutzen dürfen und daher jedes $e^{\imath k\cdot}$ als $\cos(k\cdot) - \imath\sin(k\cdot)$ für $k \geq 0$ bzw. als $\cos(-k\cdot) - \imath\sin(-k\cdot)$ für $k < 0$ schreiben können.

Korollar 2.31. *Sind $\{v_n\}_{n\in\mathbb{N}}$ ein Orthonormalsystem im Hilbertraum V und $v \in V$ beliebig, so konvergiert die Reihe $\sum_{n=1}^{\infty}\langle v_n|v\rangle v_n$ in V.*

Konvergente Reihen sind oft hilfreich, also verdient diese Klasse einen eigenen Namen.

Definition 2.32. *Sind $\{v_n\}_{n\in\mathbb{N}}$ ein Orthonormalsystem im Hilbertraum V und $v \in V$ beliebig, so heißt $\sum_{n=1}^{\infty}\langle v_n|v\rangle v_n$ die Fourierreihe von v bzgl. $\{v_n\}_{n\in\mathbb{N}}$.*

Bislang konnte uns noch niemand garantieren, dass die Fourierreihe von v gegen v selbst konvergiert. Man denke nur an das klassische Beispiel $V = L_2((-\pi,\pi);\mathbb{R})$ und das reduzierte Orthonormalsystem $\{\sin(n\cdot)\}_{n\in\mathbb{N}}$. Damit können wir nur ungerade Funktionen sinnvoll darstellen. Tatsächlich ist für jede gerade Funktion $v \in L_2((-\pi,\pi))$:

$$\int_{(-\pi,\pi)} \sin(kt)v(t)\,\mathrm{d}t = \int_{(-\pi,0)} \underbrace{\sin(kt)v(t)}_{=-\sin(-kt)v(-t)}\,\mathrm{d}t + \int_{(0,\pi)} \sin(kt)v(t)\,\mathrm{d}t = 0.$$

Satz 2.33. *Für ein Orthonormalsystem $\{u_n\}_n$ im Hilbertraum $(V,\langle\cdot|\cdot\rangle)$ sind folgende Aussagen äquivalent.*

(i) *Für alle $v \in V$ gilt $v = \sum_n \langle u_n|v\rangle v_n$.*

(ii) *$\{u_n\}_n$ ist maximal: Gilt $u_n \perp v$ für alle n, so ist $v = 0$ bzw. $\{u_n\}_n^{\perp} = \{0\}$.*

(iii) *Für alle $v \in V$ gilt in der Besselschen Ungleichung Gleichheit:*

$$\sum_n |\langle u_n|v\rangle|^2 = \|v\|^2.$$

(Parsevalsche Vollständigkeitsrelation)

(iv) *Für alle $v, w \in V$ gilt*

$$\langle v|w\rangle = \sum_n \langle v|u_n\rangle\langle u_n|w\rangle.$$

Wir haben offen gelassen, über welchen Bereich n läuft, da die gleiche Aussage auch für den endlichdimensionalen Fall gilt. Wir können es also mit endlichen, abzählbar unendlichen oder sogar überabzählbar unendlichen Orthonormalsystemen zu tun haben. Wir werden uns hier nur den höchstens abzählbar unendlichen Fall beweisen.

Beweis. $(i) \implies (iv)$: Zu beliebigen $v, w \in V$ nutzen wir die Darstellung nach (i) und berechnen

$$
\begin{aligned}
\langle v|w\rangle &\overset{(i)}{=} \left\langle \sum_n \langle u_n|v\rangle u_n \middle| \sum_m \langle u_m|w\rangle u_m \right\rangle \\
(\text{Stetigkeit von } \langle\cdot|\cdot\rangle) \quad &= \sum_n \sum_m \big\langle \langle u_n|v\rangle v_n \big| \langle u_m|w\rangle u_m \big\rangle \\
&= \sum_n \sum_m \underbrace{\langle u_n|v\rangle^*}_{=\langle v|u_n\rangle} \langle u_m|w\rangle \underbrace{\langle u_n|u_m\rangle}_{=\delta_{n,m}} \\
&= \sum_n \langle v|u_n\rangle\langle u_n|w\rangle.
\end{aligned}
$$

$(iv) \implies (iii)$: ergibt sich einfach durch $v = w$ in (iv). Die Summanden auf der linken Seite sind dann $\langle u_n|v\rangle^*\langle u_n|v\rangle = |\langle u_n|v\rangle|^2$.

$(iii) \implies (ii)$: Gilt für $v \perp u_n$ für alle n, so ist nach (iii)

$$
\|v\|^2 = \sum_n \underbrace{|\langle u_n|v\rangle|^2}_{=0} = 0, \qquad \text{also } v = 0_V.
$$

$(ii) \implies (i)$: Sind $v \in V$ beliebig und $w := \sum_n \langle u_n|v\rangle u_n$ ($\in V$ nach Korollar 2.31), so wollen wir zeigen, dass $v = w$ ist. Nach (ii) genügt es dafür $(v - w) \perp u_n$ für alle n zu zeigen. Tatsächlich ist:

$$
\begin{aligned}
\langle u_n|v - w\rangle &= \left\langle u_n \middle| v - \sum_m \langle u_m|v\rangle u_m \right\rangle \\
&= \langle u_n|v\rangle - \sum_m \big\langle u_n \big| \langle u_m|v\rangle u_m \big\rangle \\
&= \langle u_n|v\rangle - \sum_m \underbrace{\langle u_n|u_m\rangle}_{=\delta_{n,m}} \langle u_m|v\rangle \;=\; 0.
\end{aligned}
$$

Dabei haben wir auf dem Weg zur zweiten Zeile die (absolute) Konvergenz der Reihe benutzt, die uns erlaubte, die Linearität des Skalarprodukts für eventuell unendlich viele Summanden anzuwenden. □

Im überabzählbaren Fall müssten wir noch definieren, was wir unter einer überabzählbaren Summe verstehen wollen. In den uns hier interessierenden Fällen ist das sogar möglich. Sind $\mathcal{I}$ eine überabzählbare Indexmenge, $\{\alpha_i\}_{i\in\mathcal{I}}$ eine Familie nicht negativer reeller Zahlen und $\{v_i\}_{i\in\mathcal{I}}$ eine Familie von Vektoren im normierten Vektorraum V, so setzen wir

$$
\sum_{i\in\mathcal{I}} \alpha_i := \sup \left\{ \sum_{i\in\mathcal{J}} \alpha_j : \mathcal{J} \subset \mathcal{I} \text{ endlich} \right\} \in [0, \infty]
$$

(Konvergenz der Summe bedeutet dann natürlich, dass das Supremum endlich ist) und sagen, $\sum_{i\in\mathcal{I}} v_i$ konvergiere in V (gegen $v \in V$), falls zu jedem $\varepsilon > 0$ eine endliche Menge $\mathcal{K} \subset \mathcal{I}$ derart existiert, dass

$$
\left\| v - \sum_{j\in\mathcal{J}} v_j \right\| < \varepsilon \quad \text{für alle endlichen } \mathcal{J} \subset \mathcal{I} \text{ mit } \mathcal{K} \subset \mathcal{J}
$$

ist. Damit können wir den Aussagen (i), (iii) und (iv) des Satzes Sinn verleihen, und der Satz gilt sinngemäß weiterhin. Ebenso sind die Besselungleichung und die Aussage von Satz 2.28 sinnvoll und gelten auch im überabzählbaren Fall.

Haben wir ein Orthonormalsystem, das die Aussagen von Satz 2.33 erfüllt, so können wir nach (i) jeden Vektor $v \in V$ als (im allgemeinen abzählbar unendliche) Linearkombination der u_n darstellen. Dass keine anderen als die angegebenen (Fourier-)Koeffizienten infrage kommen, sehen wir sofort. Bezeichnen wir nämlich die Fourierkoeffizienten von v mit $\alpha_n = \langle u_n | v \rangle$ und nehmen an, ein Vektor w sei durch $w = \sum_n \beta_n u_n$ gegeben. Ist $\alpha_{n_0} \neq \beta_{n_0}$ für ein n_0, so erhalten wir nach Pythagoras

$$\begin{aligned}
\|v - w\|^2 &= \left\| \sum_n (\alpha_n - \beta_n) u_n \right\|^2 \\
&= \left\| (\alpha_{n_0} - \beta_{n_0}) u_{n_0} + \sum_{n \neq n_0} (\alpha_n - \beta_n) u_n \right\|^2 \\
&= |\alpha_{n_0} - \beta_{n_0}|^2 \|u_{n_0}\|^2 + \left\| \sum_{n \neq n_0} (\alpha_n - \beta_n) u_n \right\|^2 \\
&\geq |\alpha_{n_0} - \beta_{n_0}|^2 > 0,
\end{aligned}$$

und zwar unabhängig davon, ob die übrigen α_n und β_n übereinstimmen.
Diese eindeutige Darstellbarkeit eines jeden Vektors durch eine feste Sammlung von speziellen vorgegebenen Vektoren kennen wir schon aus der linearen Algebra. Dort hatten wir das System der speziellen Vektoren als Basis bezeichnet. Folgende Definition sollte also nicht überraschen.

Definition 2.34. *Ein Orthonormalsystem $\{u_n\}_n$ im Hilbertraum $(V, \langle\cdot|\cdot\rangle)$ heißt vollständiges Orthonormalsystem (oder Orthonormalbasis oder Hilbertraumbasis) für V, falls es die Bedinungen von Satz 2.33 erfüllt.*

Man beachte, dass es sich bei einer Orthonormalbasis im allgemeinen nicht um eine Vektorraumbasis handelt! Eine solche wäre ein linear unabhängiges System $\{v_i\}_{i \in \mathcal{I}}$ derart, dass jedes $v \in V$ (wegen der linearen Unabhängigkeit eindeutig) als Linearkombination *endlich vieler* der v_i schreiben lässt: $v = \sum_{k=1}^{m(v)} \alpha_k v_{i_k}$. Mittels Auswahlaxiom kann man zeigen, dass jeder Vektorraum (unabhängig von eventuellen Normen oder Skalarprodukten) eine solche (Hamel-)Basis besitzt. Tatsächlich kann man zeigen, dass jede Hamel-Basis eines Banachraums entweder endlich oder überabzählbar unendlich ist.

Bislang haben wir definiert, was wir unter einer Orthonormalbasis verstehen wollen und uns angesehen, woran wir eine solche erkennen können. Über die Existenz solcher Basen wissen wir aber noch nicht viel. Im endlichdimensionalen Fall ist das nicht schwierig.

Beispiel 25 (Gram-Schmidt-Verfahren)**.** *Im euklidischen $\mathbb{K}^n$ haben wir die kanonische Orthonormalbasis $\{e_1, \ldots, e_n\}$ aus den Standardeinheitsvektoren. Ist $(V, \langle\cdot|\cdot\rangle)$ ein beliebiger Hilbertraum mit (Vektorraum-)Dimension n über $\mathbb{K}$ und ist $\{v_1, \ldots, v_n\}$ eine beliebige*

(Vektorraum-)Basis von V, so erhalten wir eine Orthonormalbasis $\{b_1, \ldots, b_n\}$ wie folgt.

1. Setze $b_1 = \frac{1}{\|v_1\|} v_1$.

2. Setze $w_2 = v_2 - \langle b_1 | v_2 \rangle b_1$. Überzeugen Sie sich, dass w_2 die orthogonale Projektion auf $(\mathbb{K}b_1)^\perp$, das orthogonale Komplement des von b_1 aufgespannten Unterraums, ist (vgl. Beweis von Satz 2.30). Wegen $v_2 \notin \mathbb{K}b_1$ ist $w_2 \neq 0$, und selbstvertändlich ist $w_2 \perp b_1$. Anschließend normieren wir w_2 zu $b_2 = \frac{1}{\|w_2\|} w_2$. 3. Wir wiederholen den zweiten Schritt für $k = 3, \ldots, n$: Setze

$$w_k = v_k - \sum_{l=1}^{k-1} \langle b_l | v_k \rangle b_l \in <\{b_1, \ldots, b_{k-1}\}>^\perp,$$

stelle fest, dass $w_k \neq 0$ und $w_k \perp b_l$ für $l = 1, \ldots, k-1$ gelten. Setze $b_k = \frac{1}{\|w_k\|} w_k$.

Dass der endlichdimensionale Fall einfach ist, kennen wir schon von den Vektorraumbasen. Wie sieht es nun aber im unendlichdimensionalen Fall aus.

Beispiel 26. *Im prototypischen unendlichdimensionalen Hilbertraum ℓ_2 setzen wir einfach unsere Basen des $\mathbb{K}^n$ fort:*

$$b^{(n)} \in \ell_2, \quad b_k^{(n)} = \delta_{k,n}$$

liefert offensichtlich ein Orthonormalsystem, und jedes $v = (v_k)_{k \in \mathbb{N}}$ in ℓ_2 ist durch

$$v = \sum_{k=1}^{\infty} v_k b^{(k)}$$

gegeben. Selbstverständlich ist

$$v_k = \sum_{l=1}^{\infty} \delta_{l,k}^* v_l = \langle b^{(k)} | v \rangle,$$

der richtige Fourierkoeffizient, wie in der allgemeinen Darstellung vorgegeben.
Wir prüfen auch sehr schnell die Maximalität: Gilt zu $v \in \ell_2$ für alle $n \in \mathbb{N}$: $0 = \langle b^{(n)} | v \rangle = v_n$, so ist $v = 0$.

Wie bei Basen üblich, können wir nicht ein einziges Element weg lassen, ohne die Basiseigenschaft zu verlieren.

Korollar 2.35. *Ist $\{v_n\}_{n \in \mathbb{N}}$ eine Orthonormalbasis im Hilbertraum $(V, \langle \cdot | \cdot \rangle)$, dann ist $\{v_n\}_{n \in \mathbb{N}, n \neq n_0}$ für kein $n_0 \in \mathbb{N}$ eine Orthonormalbasis.*

Beweis. $v = v_{n_0}$ ist normiert, also $v \neq 0$, erfüllt aber für alle $n \in \mathbb{N}$, $n \neq n_0$: $v \perp v_n$. Damit ist (ii) in 2.33 nicht erfüllt. □

Beispiel 27. *Das Orthonormalsystem aus Beispiel 24 in $L_2((-\pi,\pi);\mathbb{K})$ ist ebenfalls eine Orthonormalbasis. Durch Skalierung und Translation der reellen Achse können wir daraus Orthonormalbasen für jeden $L_2((a,b);\mathbb{K})$ machen.*
Wir dürfen aber keine der Basisfunktionen weg lassen.

Weitere Beispiele sind Systeme orthogonaler Polynome auf kompakten Intervallen oder (zusammen mit geeigneten Gewichtungsfunktionen) auf der reellen (Halb-)Achse. In quaderförmigen Gebieten $\Omega \subset \mathbb{R}^d$ ergeben Produkte trigonometrischer Funktionen Orthonormalbasen von $L_2(\Omega;\mathbb{K})$, für Kugeln benutzt man Bessel-Funktionen und deren Verwandte.
Konkrete Orthonormalbasen zu finden, ist oft recht schwierig. Allerdings können wir eine große Klasse von Hilberträumen angeben, für die wir die Existenz einer abzählbaren Orthonormalbasis garantieren können.

Definition 2.36. *Ein metrischer Raum (V,d) heißt separabel, wenn es eine abzählbare Teilmenge $B \subset V$ mit $\bar{B} = V$ gibt.*

Beispiel 28. *(a) $\mathbb{R}^n$ ist als normierter Vektorraum separabel, da die abzählbare Teilmenge $\mathbb{Q}^n$ (unter jeder Norm) dicht in $\mathbb{R}^n$ liegt.*

(b) ℓ_2 ist ein separabler Hilbertraum, denn

$$B = \{a = (a_n)_n \in c_{00} : a_n \text{ algebraisch für alle } n\}$$

ist eine abzählbare dichte Teilmenge. Ist nämlich $v \in \ell_2$ beliebig, so finden wir zu $\varepsilon > 0$ zunächst ein N mit

$$\sum_{n=N+1}^{\infty} |v_n|^2 \le \frac{\varepsilon^2}{4}$$

und dann zu $n \in \mathbb{N}$, $n \le N$, eine algebraische Zahl a_n mit $|a_n - v_n|^2 < 2^{-n-2}\varepsilon^2$. Dann gilt (mit $a_n = 0$ für $n \ge N+1$):

$$\|v-a\|_2 = \sqrt{\sum_{n=1}^{\infty} |v_n - a_n|^2} \le \sqrt{\frac{\varepsilon^2}{4}\sum_{n=1}^{N} 2^{-n} + \sum_{n=N+1}^{\infty} |v_n|^2} \le \varepsilon.$$

(c) $C([a,b];\mathbb{R})$ ist unter der Supremumsnorm ein separabler Banachraum. Nach dem Approximationssatz von Stone-Weierstraß ist zunächst der Raum W der Polynome mit reellen Koeffizienten eine dichte Teilmenge. Andererseits finden wir zu jedem Polynom, definiert durch $p(t) = \sum_{j=0}^m \alpha_j t^j$ und gegebenem $\varepsilon > 0$ ein weiteres, durch $q(t) = \sum_{j=0}^m r_j t^j$ gegebenes Polynom mit rationalen Koeffizienten derart, dass

$$\|p-q\|_\infty = \sup_{a\le t\le b} \left|\sum_{j=0}^{m} (\alpha_j - r_j) t^j\right| \le \sup_{a\le t\le b} \sum_{j=0}^{m} |\alpha_j - r_j||t|^j \le \sum_{j=0}^{m} |\alpha_j - r_j|(|a| \vee |b|)^j < \frac{\varepsilon}{2}.$$

Man suche dazu etwa die r_j so, dass $|\alpha_j - r_j| \leq \frac{\varepsilon}{2m}(|a| \vee |b| \vee 1)^m$ gilt. Die Menge

$$B = \left\{ \sum_{j=0}^{m} r_j(\cdot)^j : m \in \mathbb{N}, r_0, \ldots, r_m \in \mathbb{Q} \right\}$$

ist also eine abzählbare dichte Teilmenge.

Wir können einem metrischen Raum auch ansehen, wenn er nicht separabel ist. Ein Kriterium liefert das folgende intuitiv klare Lemma: Können wir in einem metrischen Raum überabzählbar viele Punkte unterbringen, die paarweise um mindestens δ voneinander entfernt sind, so kann der Raum nicht separabel sein.

Lemma 2.37. *Ist $A \subset V$ eine überabzählbare Teilmenge eines metrischen Raums (V, d) derart, dass für ein $\delta > 0$ gilt: $d(v, w) > \delta$ für alle $v, w \in A$ mit $v \neq w$, so ist V nicht separabel.*

Beweis. Wäre $B \subset V$ eine abzählbare dichte Teilmenge von V, so gäbe es zu jedem $v \in A$ und $\varepsilon = \frac{\delta}{2}$ ein $b \in B$ mit $d(v, b) < \varepsilon$. Das müsste aber nach Dreiecksungleichung zu jedem v ein anderes b sein. Sind nämlich $v, w \in A$, $v \neq w$, so gilt für jedes $b \in B \subset V$:

$$\delta < d(v, w) \leq d(v, b) + d(b, w).$$

Das ist aber im Widerspruch zur B abzählbar, A überabzählbar. $\square$

Dieses Lemma liefert nun auch erste Beispiele oft benutzter nicht separabler Banachräume.

Beispiel 29. (*a*) *ℓ_∞ ist nicht separabel. Die Menge $\{0, 1\}^{\mathbb{N}}$ aller Folgen $v = (v_n)_n$ mit $v_n \in \{0, 1\}$ ist nicht abzählbar (ihre Mächtigkeit ist die der Potenzmenge von $\mathbb{N}$), und zu $\delta = \frac{1}{2}$ gilt: Sind $v, w \in \{0, 1\}^{\mathbb{N}}$, $v \neq w$, also $v_{n_0} \neq w_{n_0}$ für mindestens ein n_0:*

$$d(v, w) = \|v - w\|_\infty = \sup_{n \geq 1} |v_n - w_n| \geq |v_{n_0} - w_{n_0}| = 1 > \delta.$$

(*b*) *Ganz analog ist $L_\infty((0, 1); \mathbb{K})$ nicht separabel. Eine überabzählbare Teilmenge mit $\|v - w\|_\infty > \frac{1}{2}$ für $v \neq w$ wäre zum Beispiel $A = \{\mathbf{1}_{(0,t)} : t \in (0, 1)\}$.*

Nicht separable Hilberträume sind schon etwas schwieriger zu finden. Man könnte etwa an $L_2(\mu)$ für ein nicht σ-endliches Maß μ denken. So ist $L_2((\mathbb{R}, \zeta); \mathbb{K})$ auf $\mathbb{R}$, versehen mit dem Zählmaß, nicht separabel.

Satz 2.38. *Der Hilbertraum V besitzt genau dann eine höchstens abzählbar unendliche Orthonormalbasis, wenn V (unter der durch das Skalarprodukt induzierten Metrik) separabel ist.*

Beweis. (I) Ist $\{u_n\}_n$ ein höchstens abzählbares Orthonormalsystem, so ist

$$B = \left\{ \sum_{n=1}^{N} a_n u_n : N \in \mathbb{N}, a_1, \ldots, a_N \text{ algebraisch} \right\}$$

eine abzählbare dichte Teilmenge. Zu $v \in V$ und $\varepsilon > 0$ finden wir nämlich ein $N \in \mathbb{N}$ derart, dass

$$\left\| v - \sum_{n=1}^{N} \langle u_n | v \rangle u_n \right\| < \frac{\varepsilon}{2}$$

gilt. Zu den $\alpha_n = \langle u_n | v \rangle$, $n = 1, \ldots, N$ finden wir nun noch algebraische a_n derart, dass

$$\left\| \sum_{n=1}^{N} a_n u_n - \sum_{n=1}^{N} \alpha_n u_n \right\| < \frac{\varepsilon}{2}.$$

Zusammen liefert die Dreiecksungleichung, dass $\sum_{n=1}^{N} a_n u_n$ das gegebene v auf höchstens ε genau approximiert.

(II) Für die Umkehrung benutzen wir wieder Gram-Schmidt. Ist $\{0_V\} \neq M = \{x_k : k \in \mathbb{N}\}$ eine abzählbare dichte Teilmenge, so wählen wir $k_1 = \min\{k \in \mathbb{N} : x_k \neq 0\}$ und setzen $u_1 = \frac{1}{\|x_{k_1}\|} x_{k_1}$. Dann spannt u_1 den gleichen Unterraum V_1 auf wie $\{x_1, \ldots, x_{k_1}\}$. Ist das schon ganz V, so sind wir fertig. Andernfalls finden wir in $M_1 = \{x_k : k > k_1\}$ ein erstes x_{k_2}, welches nicht in V_1 liegt. Von diesem nehmen wir den zu V_1 orthogonalen Anteil w_2 und normieren diesen zu u_2. Dann ist $\{u_1, u_2\}$ ein Orthonormalsystem, das den gleichen Raum V_2 wie $\{x_1, \ldots, x_{k_2}\}$ aufspannt.
So fahren wir fort, solange nicht $V_m = <\{u_1, \ldots, u_m\}> = V$ ist. Passiert das bei endlichem m, so ist $\{u_1, \ldots, u_m\}$ nicht nur eine Orthonormal-, sondern auch eine Vektorraumbasis. Andernfalls erhalten wir induktiv zu jedem $n \in \mathbb{N}$ ein u_n derart, dass $\{u_n\}_{n\in\mathbb{N}}$ ein Orthonormalsystem ist und die gleiche lineare Hülle hat wie M.
Ist nun $v \in V$ beliebig mit $\langle u_n | v \rangle$ für alle $n \in \mathbb{N}$, so folgt wegen der Linearität des Skalarprodukts $\langle w | v \rangle$ für alle w in der linearen Hülle $<\{u_n\}_{n\in\mathbb{N}}> = <M>$, also insbesondere für alle $w \in M$. Wegen der Stetigkeit des Skalarprodukts gilt dann aber $\langle w | v \rangle$ für alle $w \in V$, also auch für $w = v$ und damit $v = 0$. Nach (ii) in Satz 2.33 ist $\{u_n\}_n$ also eine Orthonormalbasis. □

Als Folgerung erhalten wir eine ganz bemerkenswerte Aussage: Zu jedem $n \in \mathbb{N}$ gibt es bis auf isometrische Isomorphie genau zwei n-dimensionale Hilberträume, nämlich die euklidischen $\mathbb{R}^n$ und $\mathbb{C}^n$. Ebenso gibt es bis auf isometrische Isomorphie genau zwei unendlichdimensionale separable Hilberträume, nämlich $\ell_2^{\mathbb{R}}$ und $\ell_2^{\mathbb{C}}$. Wir formulieren die zweite Aussage als Lemma.

Lemma 2.39. *Ist $(V, \langle \cdot | \cdot \rangle)$ ein unendlichdimensionaler separabler Hilbertraum, so ist V isometrisch isomorph zu ℓ_2.*

Beweis. Satz 2.38 sichert die Existenz einer abzählbaren Orthonormalbasis $\{u_n\}_{n\in\mathbb{N}}$. Zu $v \in V$ bezeichnen wir die Folge der Fourierkoeffizienten von v mit

$$J(v) := (\langle u_n|v\rangle)_{n\in\mathbb{N}} \in \ell_2^{\mathbb{K}}.$$

Dass $J(v)$ quadratintegrierbar ist, folgt dabei aus Satz 2.28.
Wir haben zu zeigen, dass J ein isometrischer Isomorphismus ist.
Die Linearität ist klar, da das Skalarprodukt von V linear im zweiten Argument ist.
Ist ferner $a = (a_n)_n \in \ell_2^{\mathbb{K}}$ beliebig, so ist $v := \sum_{n=1}^{\infty} a_n u_n \in V$ ein Urbild von a unter J, also ist J surjektiv.
J ist isometrisch, denn für $v, w \in V$ ist

$$(J(v)|J(w)) \;=\; \sum_{n=1}^{\infty} \langle v_n|v\rangle^* \langle v_n|w\rangle \;\overset{(iv)\text{ in S.2.33}}{=}\; \langle v|w\rangle.$$

Dabei haben wir das Skalarprodukt des ℓ_2 mit $(\cdot|\cdot)$ bezeichnet.
Daraus folgt schließlich auch die Injektivität. Ist nämlich $J(v) = 0_{\ell_2}$, so ist

$$\langle v|v\rangle = \big(J(v)|J(v)\big) = 0,$$

also $v = 0$. $\square$

Das Vorgehen ist also genau so, wie wir es aus dem endlichdimensionalen Fall kennen. Wir identifizieren $v \in V$ einfach mit dem (Spalten-)vektor seiner Koordinaten (Fourierkoeffizienten) bezüglich einer frei gewählten (Orthonormal-)Basis von V. Da es sich um eine Orthonormalbasis handelt, übertragen sich alle Rechnungen (Vektorraumoperationen, Skalarprodukt, insbesondere Norm und Abstand) eins zu eins von V nach ℓ_2 und vor allem auch wieder zurück.

KAPITEL 3

Funktionalanalytische Grundprinzipien

3.1 Lineare Operatoren

Bevor wir uns mit den allgemeinen Prinzipien der linearen Funktionalanalysis beschäftigen, klären wir noch einige Begrifflichkeiten. Wir erinnern uns an die Definition linearer Funktionale. Diese sind einfach spezielle lineare Abbildungen. Den Begriff linearer Abbildungen kennen wir aus der linearen Algebra. Wir interessieren uns natürlich für lineare Abbildungen zwischen normierten linearen Räumen und geben diesen einen speziellen Namen.

Definition 3.1. *1. Sind* $(V, \|\cdot\|_V)$, $(Y, \|\cdot\|_W)$ *normierte* $\mathbb{K}$*-Vektorräume, so heißt* $F : V \supset \mathcal{D}_F \to Y$ *ein linearer Operator, falls für alle* $v, w \in \mathcal{D}_F$ *und* $\alpha \in \mathbb{K}$

$$F(\alpha v + w) = \alpha F(v)$$

gilt. Der lineare Unterraum $\mathcal{D}_F$ *heißt Definitionsbereich von* F*.*
2. Den Raum der linearen Operatoren F *von* V *nach* Y *nennen wir* $L(V, Y)$*. Dessen Unterraum der global definierten linearen Operatoren bezeichnen wir mit*

$$\mathrm{Hom}(V, Y) = \{F \in L(V, Y) : \mathcal{D}_F = V\}.$$

3. Der lineare Operator $F \in \mathrm{Hom}(V, Y)$ *heißt stetig in* $v \in V$*, falls für alle konvergenten Folgen* $(v_n)_n$ *in* V *mit Grenzwert* v

$$\lim_{n\to\infty} F(v_n) = F(v)$$

gilt. F *heißt stetig, wenn* F *in allen* $v \in V$ *stetig ist. Wir bezeichnen mit*

$$\mathcal{L}(V, Y) := \{F \in \mathrm{Hom}(V, Y) : F \textit{ stetig}\}$$

den Raum der stetigen linearen Operatoren von V *nach* Y*.*

Die Stetigkeitsdefinition über Folgen haben wir wieder von den metrischen Räumen übernommen.

Beispiel 30. *Die (stetigen) linearen Funktionale sind Spezialfälle (stetiger) linearer Operatoren mit* $(Y, \|\cdot\|_Y) = (\mathbb{K}, |\cdot|)$*. Es ist also* $L(V, \mathbb{K}) = V^*$, $\mathcal{L}(V, \mathbb{K}) = V'$*.*

Anhand dieses Beispiels erinnern wir uns auch an die Definition der Beschränktheit und der Operatornorm, die genau dem gleichen Prinzip folgen.

Definition 3.2. *Ein linearer Operator* $F : V \to Y$ *heißt beschränkt, falls ein* $C \geq 0$ *derart existiert, dass*

$$\|F(v)\|_Y \leq C\|v\|_V \qquad \textit{für alle } v \in V$$

gilt.

Wie in Abschnitt 2.3 erhalten wir auch hier die Äquivalenz von Beschränktheit und Stetigkeit. Nach unseren Standardargumenten zu Verschiebungen und Streckungen bzw. Stauchungen erhalten wir sogar noch mehr.

Lemma 3.3. *Für $F \in \mathrm{Hom}(V, Y)$ sind folgende Aussagen äquivalent.*

(a) *Es existiert ein $v \in V$ derart, dass F in v stetig ist*[a].

(b) *F ist stetig in allen $v \in V$.*

(c) *F ist stetig in 0_V.*

(d) *F ist beschränkt.*

(e) *$F(B^V)$ ist beschränkt in Y.*

(f) *$F(S^V)$ ist beschränkt in Y.*

(g) *$F(M)$ ist beschränkt in Y für jede beschränkte Teilmenge $M \subset V$.*

Dabei sind B^V und S^V die Einheitskugel bzw. Einheitssphäre in V.

[a] Erinnern Sie sich: V enthält als Vektorraum mindestens ein Nullelement.

Beweis. Den Beweis haben wir schon weitgehend in Abschnitt 2.3 geführt. Die Anpassung von $Y = \mathbb{K}$ auf allgemeine Y nehmen Sie als Übung selbst vor. $\square$

Definition 3.4. *Für einen beschränkten linearen Operator $F : V \to Y$ definieren wir die Operatornorm*

$$\|F\| := \sup\{\|F(v)\|_Y : v \in V,\ \|v\|_V \leq 1\}.$$

Dass es sich tatsächlich um eine Norm handelt, prüfen Sie wie schon im Fall der linearen Funktionale schnell nach. Ebenso schnell erhalten wir den Zusammenhang zwischen der Norm und der Konstante C aus der Definition der Beschränktheit.

Lemma 3.5. *Für $F \in \mathcal{L}(V, Y)$ ist*

$$\|F\| = \inf\{C \geq 0 : \|F(v)\|_Y \leq C\|v\|_V \text{ für alle } v \in V\}.$$

Aus der linearen Algebra erinnern wir uns, dass $\mathrm{Hom}(V, Y)$ für beliebige Vektorräume V, Y wieder ein Vektorraum ist. Im Fall stetiger linearer Operatoren zwischen normierten linearen Räume haben wir sogar noch mehr.

Satz 3.6. *Sind $(V, \|\cdot\|_V)$ und $(Y, \|\cdot\|_Y)$ normierte $\mathbb{K}$-Vektorräume, so ist auch $\mathcal{L}(V, Y)$ ein normierter Vektorraum unter der Operatornorm, und $\mathcal{L}(V, Y)$ ist ein Banachraum, wenn Y ein Banachraum ist.*

Beweis. Wir haben nur noch die Vollständigkeit zu prüfen, falls Y ein Banachraum ist. Dazu sei $(F_n)_n$ eine Cauchyfolge in $\mathcal{L}(V, Y)$. Für beliebige $v \in V$ ist dann $(F_n(v))_n$ eine Cauchyfolge in Y: Ohne Beschränkung der Allgemeinheit ist $v \neq 0_V$ (sonst ist $F_n(v) = 0_Y$ für alle n), und wir finden zu $\varepsilon > 0$ ein n_0 derart, dass $\|F_n - F_m\| < \frac{\varepsilon}{\|v\|_V}$ für alle $n, m \geq n_0$ gilt. Dann ist

$$\|F_n(v) - F_m(v)\|_Y = \|(F_n - F_m)(v)\|_Y \leq \|F_n - F_m\|\|v\|_V < \varepsilon.$$

Damit können wir für alle $v \in V$

$$F(v) = \lim_{n\to\infty} F_n(v)$$

definieren und erhalten einen linearen Operator $F : V \to Y$. Dieser ist auch stetig, denn für beliebige $v \in V$ gilt (wegen der Stetigkeit der Norm):

$$\|\underbrace{(F - F_n)(v)}_{=F(v)-F_n(v)}\|_Y = \lim_{m\to\infty} \underbrace{\|(F_m - F_n)(v)\|_Y}_{\le\|F_m-F_n\|\|v\|_V} \le \left(\liminf_{m\to\infty} \|F_m - F_n\|\right) \|v\|_V.$$

Dabei ist der Limes inferior auf der rechten Seite gegen ein von n und v unabhängiges $C \geq 0$ beschränkt, da die Cauchyfolge $(F_n)_n$ in $\mathcal{L}(V, Y)$ beschränkt ist. Für

$$\|v\|_V < \delta = \frac{\varepsilon}{1 + \sup_m \|F_m\|}$$

ist nämlich wegen $\|F\| \le \sup_m \|F_m\|$ und damit $\|F(v)\|_Y \le \|F\|\|v\|_V < \varepsilon$, F also stetig in 0_V. □

Wir sehen uns einige Beispiele linearer Operatoren an.

Beispiel 31. *Den Fall $V = \mathbb{K}^n$, $Y = \mathbb{K}^m$, jeweils versehen mit einer beliebigen Norm müssen wir kaum erwähnen. Hier sind alle lineaeren Abbildungen beschränkt, wie wir aus der linearen Algebra wissen.*

Beispiel 32. *Den Dualraum $V' = \mathcal{L}(V; \mathbb{K})$, also den Raum der stetigen linearen Funktionale auf V kennen wir schon. Der Satz liefert nun, dass V' stets vollständig ist, auch wenn V selbst das unter Umständen nicht ist.*
Konkret liefert etwa die Hölderungleichung, dass zu $p \in (1, \infty)$ jedes $w \in L_{p'}(\Omega)$ ($p' = \frac{p}{p-1}$) via

$$\phi_w(v) = \int_\Omega w(x)^* v(x)\,\mathrm{d}x$$

ein stetiges lineares Funktional auf $L_p(\Omega)$ mit $\|\phi_w\| \le \|w\|_{p'}$ definiert. Tatsächlich gilt sogar Gleichheit, und analog zum zweiten Rieszschen Darstellungssatz 2.26 ist mittels dieser Identifikation $L_p'(\Omega) \cong L_{p'}(\Omega)$. Man schreibt in Analogie zum Skalarprodukt auch

$$\phi_w(v) = \langle w, v\rangle$$

und bezeichnet dieses Konstrukt als duale Paarung. Im Fall $\mathbb{K} = \mathbb{C}$ ist die duale Paarung wie auch das Skalarprodukt linear im zweiten Argument (aus V) und konjugiert linear im ersten Argument (aus V').
Ähnlich gilt $L_1' \cong L_\infty$, jedoch gilt $L_\infty' \supsetneqq L_1$.

Beachten Sie, dass für die L_p-Räume ($p \in (0, 1)$) wegen $(p')' = p$ gilt:

$$(L_p')' \cong L_p,$$

wobei für einen normierten Vektorraum V mit $V'' = (V')'$ der Raum der stetigen linearen Funktionale auf dem Dualraum V' bezeichnet wird. Diesen nennen wir auch den Bidualraum von V.
Diese Eigenschaft, die wir schon von Hilberträumen kennen, hat einen eigenen Namen.

Definition 3.7. *Ein Banachraum $(V, \|\cdot\|)$ heißt* <u>*reflexiv*</u>*, wenn für seinen Bidualraum $V'' \cong V$ gilt.*
Dabei heißen die normierten $\mathbb{K}$-Vektorräume $(V, \|\cdot\|_V)$ und $(W, \|\cdot\|_W)$ <u>*isometrisch isomorph*</u>*, geschrieben $V \cong W$, wenn es eine bijektive lineare Abbildung $J : V \to W$ mit*

$$\|J(v)\|_W = \|v\|_V \qquad \text{für alle } v \in V$$

gibt. J heißt dann <u>*isometrischer Isomorphismus*</u>*.*

Da V'' nach der obigen Diskussion stets ein Banachraum ist, können nicht vollständige normierte Räume unmöglich reflexiv sein. Deshalb haben wir die Definition nur für Banachräume formuliert.
Für einen reflexiven Raum ist die duale Paarung in Analogie zu (Semi-)Skalarprodukten also auch konjugiert symmetrisch in dem Sinn, dass für alle $v \in V$, $w \in V'$

$$\langle w, v\rangle = \langle J(v), w\rangle^*$$

gilt.
Beachten Sie bitte die etwas gefährlich Dopplung der Bezeichnungen in der Definition. Für Prähilberträume hatten wir für den isometrischen Isomorphismus nicht die Norm- sondern auch die Skalarprodukttreue gefordert. Das liegt in der Etymologie des Begriffs isometrisch begründet. Dieser bedeutet, dass wir in beiden Räumen genau gleich messen. In normierten Räumen können wir kraft der Norm Abstände bzw. Längen messen. Daher sollte eine Isometrie normerhaltend sein. In Hilberträumen haben wir zusätzlich noch das Skalarprodukt, das uns neben der Längenmessung via $\|v\|^2 = \langle v|v\rangle$ auch die Messung von Winkeln und insbesondere die Untersuchung der Orthogonalität erlaubt. Eine Hilbertraum-Isometrie sollte also auch das Skalarprodukt selbst erhalten, damit sich an unseren Messungen nichts ändert.
Wie schon bei (Prä-)Hilberträumen, schließen wir auch für isometrische Isomorphismen zwischen normierten Räumen sofort, dass diese lineare Homöomorphismen mit $\|J\| = \|J^{-1}\| = 1$ sind.

Beispiel 33 (Integraloperatoren)**.** *Eine große Klasse beschränkter Operatoren erhalten wir in der Form*

$$F(v)(x) := \int_\Omega k(x, y)v(y)\,\mathrm{d}y.$$

Dabei sind $\Omega \subset \mathbb{R}^d$ offen, $k : \Omega \times \Omega \to \mathbb{K}$ und ein sogenannter Integralkern. Dann nimmt F eine Funktion $v : \Omega \to \mathbb{K}$ entgegen und bildet diese auf eine Funktion $F(v) : \Omega \to \mathbb{K}$ ab. Sind k und v dergestalt, dass das Integral für alle x existiert, so ist die Linearität klar. Für die Wohldefiniertheit und Stetigkeit von F muss k gewisse Bedingungen erfüllen, und v muss aus einem geeigneten Urbildraum V kommen.
Sind zum Beispiel Ω beschränkt und glatt berandet sowie k stetig und beschränkt, kann also zu $K \in C(\bar\Omega \times \bar\Omega; \mathbb{K})$ fortgesetzt werden (wir nennen dieses wieder k, nehmen also einfach an, k sei von vorn herein auf $\bar\Omega \times \bar\Omega$ gegeben), so liefert F einen stetigen linearen Operator $F : L_1(\Omega) \to C^0(\bar\Omega; \mathbb{K})$ mit Operatornorm $\|F\| \leq \|k\|_\infty$. Die Existenz erhalten

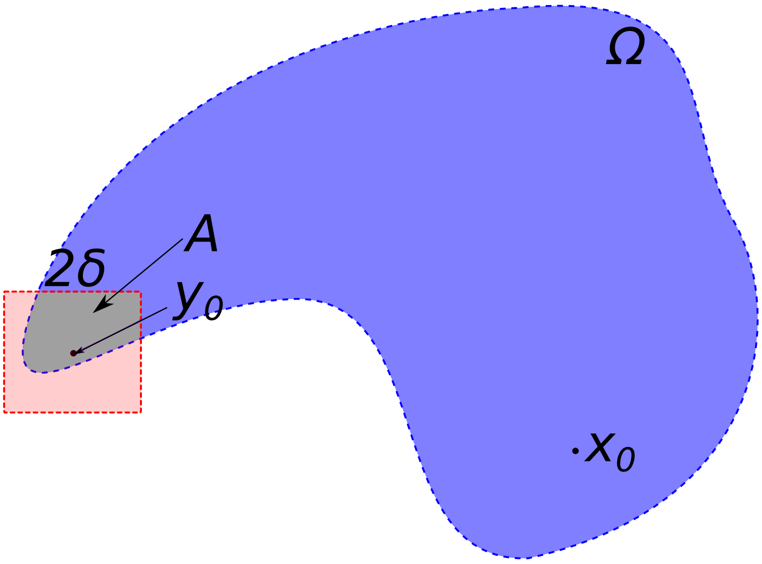
wir aus Satz **??** *über parameterabhängige Integrale (mit* $f(x,y) = k(x,y)u(y)$*) und der Abschätzung*

$$|F(v)(x)| \leq \int_\Omega \underbrace{|k(x,y)|}_{\leq \|k\|_\infty} |v(y)| \, \mathrm{d}y \leq \|k\|_\infty \|v\|_1.$$

Diese liefert auch eine obere Schranke für $\|F\|$*. Diese ist sogar scharf. Da* $|k|$ *sein Maximum auf der kompakten Menge* $\bar{\Omega} \times \bar{\Omega}$ *annimmt, finden wir ein* $(x_0, y_0) \in \bar{\Omega}$ *mit* $|k(x_0, y_0)| = \|k\|_\infty =: a$*. Ist* $a = 0$*, so ist* $k \equiv 0$ *und damit* F *trivial, also* $\|F\| = 0$*. Andernfalls finden wir zu* $\varepsilon > 0$ *ein* $\delta > 0$ *derart, dass* $|k(x_0, y)| \geq (1-\varepsilon)\alpha$ *für alle* $y \in \Omega$ *mit* $|y - y_0| < \delta$ *gilt, wobei* $|\cdot|$ *die Maximumsnorm auf* $\mathbb{R}^d$ *bezeichne. Nun wählen wir* $v = \mathbf{1}_A$*, wobei*

$$A = \{y \in \Omega : |y - y_0| < \delta\} = \Omega \cap \underbrace{(y_0 - \delta\mathbb{1}, y_0 + \delta\mathbb{1})}_{\textit{Würfel um } y_0}$$

ist.



Dann ist $\|v\|_1 = \int_A \, \mathrm{d}y = \mathrm{vol}(A) > 0$[1] *und*

$$\begin{aligned} \|F(v)\|_\infty &\geq |F(v)(x_0)| \\ &= \left| \int_\Omega k(x_0, y) v(y) \, \mathrm{d}y \right| = \left| \int_A k(x_0, y) 1 \, \mathrm{d}y \right| \\ &\geq \inf_{y \in A} |k(x_0, y)| \int_A \mathrm{d}y \\ &\geq a(1-\varepsilon) \qquad \|v\|_1 \end{aligned}$$

[1] Das ist das d-dimensionale Lebesgue-Maß von A. Dass wir Ω einen glatten Rand spendiert haben, liegt daran, dass wir uns auch im Fall $y_0 \in \partial\Omega$ nicht darum sorgen wollten, dass A positives Maß hat.

Also ist für jedes $\varepsilon > 0$:

$$\|k\|_\infty(1-\varepsilon) \le \|F\| \le \|k\|_\infty$$

und damit $\|F\| = \|k\|_\infty$.

Analog überzeugen Sie sich, dass wir mit dem gleichen k auch einen beschränkten linearen Operator $\tilde{F} : C(\bar{\Omega};\mathbb{K}) \to C(\bar{\Omega};\mathbb{K})$, nun aber mit der Normabschätzung

$$\|\tilde{F}\| \le \|k\|_\infty \mathrm{vol}(\Omega).$$

Beispiel 34. *Analog liefert $k \in L_\infty(\Omega \times \Omega)$ einen beschränkten linearen Operator von $L_1(\Omega)$ nach $L_\infty(\Omega)$. Hier liefert der Satz von Fubini die Existenz und Stetigkeit. Konkret betrachten wir die Fouriertransformation: Zu $\Omega = \mathbb{R}$ und $k(s,t) = \frac{1}{\sqrt{2\pi}} e^{-\imath st}$ definieren für $v \in L_1(\mathbb{R})$:*

$$F(v)(s) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\imath st} v(t)\,\mathrm{d}t.$$

Das ist ein stetiger linearer Operator $F : L_1(\mathbb{R}) \to L_\infty(\mathbb{R})$, denn

$$\|F(v)\|_\infty \le \frac{1}{\sqrt{2\pi}} \|v\|_1.$$

Beispiel 35 (Ableitungsoperatoren)**.** *(a) Wir betrachten $V = C^1([0,1];\mathbb{K})$, versehen mit der Norm*

$$\|v\|_V = \|v\|_\infty + \|v'\|_\infty$$

und $Y = C^0([0,1];\mathbb{K})$ (der Raum der stetigen Funktionen versehen mit der Maximumsnorm). Die Linearität des Operators

$$D : V \to Y, \qquad v \mapsto v'$$

steht außer Frage, und die Beschränktheit liefert eine ganz einfache Rechnung:

$$\|D(v)(t)\|_\infty = \max_{t\in[0,1]} |v'(t)| = \|v'\|_\infty \le \|v\|_V,$$

also $\|D\| \le 1$. Dass tatsächlich $\|D\| = 1$ ist, rechnen Sie schnell nach, indem Sie die Funktionen $v_n : [0,1] \to \mathbb{K}$, $v_n(t) = t^n$, betrachten. Für diese ist

$$\|D(v_n)\|_\infty = n = \frac{n}{n+1} \underbrace{\left(\|v_n\|_\infty + \|v_n'\|_\infty\right)}_{n+1} = \frac{n}{n+1}\|v_n\|_V.$$

Allerdings können Sie sich überzeugen, dass es kein $v \in V$ mit $\|D(v)\|_\infty = \|D\|\|v\|_V$ geben kann.

(b) Wir betrachten den gleichen Urbildraum $W = C^1([0,1];\mathbb{K})$, aber nun mit der Norm $\|w\|_W = \|w\|_\infty$. Der Operator D ist weiterhin wohldefiniert und linear, aber nicht mehr beschränkt, wie die Folge v_n aus (a) zeigt:

$$\|D(v_n)\| = n\|v_n\|_W.$$

(c) Verallgemeinerungen ergeben sich schnell. Auf V, Y wie in (a) können wir zu gegebenem $k \in C^1([0,1];\mathbb{K})$ den Operator $F : V \to Y$ durch $F(v) := (kv)'$ definieren und erhalten (nach Produktregel) einen beschränkten Operator.

Beispiel 36 (Multiplikationsoperatoren). *Zu den Räumen $V = C(X;\mathbb{K})$, $Y = C(X;\mathbb{K})$ der stetigen Funktionen auf dem kompakten metrischen Raum X, jeweils versehen mit der Maximumsnorm, und gegebenem $k \in C(X;\mathbb{K})$ definieren wir ganz ähnlich wie in (c) des vorigen Beispiels den Multiplikationsoperator $M : V \to Y$ durch $M(v) = k(\cdot)v(\cdot)$. Dies ist ein stetiger linearer Operator mit Operatornorm $\|k\|_\infty$.*

Beispiel 37 (Shift-Operatoren). *Zu $V = \ell_p$, versehen mit seiner Standardnorm, definieren wir den Linksshift $L : \ell_p \to \ell_p$ durch $L(v) = (v_{n+1})_n \in \mathbb{N}$ sowie den Rechtsshift $R : \ell_p \to \ell_p$ durch*

$$R : (v_n)_n \mapsto (0, v_1, v_2, \dots).$$

Wir erkennen sofort, dass $\|R(v)\|_p = \|v\|_p$, also $\|R\| = 1$ ist. Analog berechnen wir auch für den Linksshift $\|L\| = 1$. Nun finden wir tatsächlich $v \in \ell_p$ mit $\|L(v)\|_p = \|v\|_p$. Das sind mindestens jene, für die $v_1 = 0$ ist.

3.2 Der Satz von Baire

Wir wollen uns nun mit Familien beschränkter Operatoren beschäftigen. Dazu benötigen wir ein Hilfsmittel.

Definition 3.8. *1. Eine Teilmenge $A \subset V$ eines metrischen Raums (V, d) heißt nirgends dicht, falls ihr Abschluss keine inneren Punkte enthält:* $\mathrm{int}(\bar{A}) = \emptyset$.

2. Eine Teilmenge $B \subset V$ eines metrischen Raums heißt von erster Kategorie (oder mager), falls sie als abzählbare Vereinigung nirgends dichter Mengen geschrieben werden kann. Andernfalls heißt B von zweiter Kategorie (oder fett).

Die Bedingung, $\bar{A}$ enthalte keine inneren Punkte, lässt sich alternativ auch anders formulieren: $\bar{A}$ enthält keine offene Kugel von positivem Radius (um einen inneren Punkt gäbe es nämlich eine solche) bzw. $\bar{A}$ enthält keine nicht leere offene Menge.

Beispiel 38. *Endliche Teilmengen normierter Vektorräume sind nirgends dicht. Ebenso ist $\mathbb{N} \subset \mathbb{R}$ unter dem üblichen Betrag nirgends dicht. Es folgt, dass $\mathbb{Q}$ als abzählbare Vereinung einelementiger Mengen eine Menge erster Kategorie in $\mathbb{R}$ ist.*

Beispiel 39. $\mathbb{N}$ *ist nicht nirgends dicht in $\{1\} \cup [2, \infty)$ mit dem Standardbetrag, da*

$$B_{\frac{1}{2}}(1) = \{1\} \subset \mathbb{N} \subset \bar{\mathbb{N}}$$

eine nicht triviale offene Teilmenge ist. Analog ist in jedem diskreten metrischen Raum keine nicht leere Menge nirgends dicht.

Satz 3.9 (Das Prinzip geschachtelter Intervalle). *Ist $(K_n)_n$ eine absteigende Folge nicht leerer, abgeschlossener Teilmengen eines vollständigen metrischen Raums (V, d) mit*

$$\lim_{n\to\infty} \operatorname{diam} K_n = 0 \tag{3.1}$$

so existiert genau ein $v \in V$ derart, dass $\{v\} = \bigcap_{n\geq 1} K_n$.

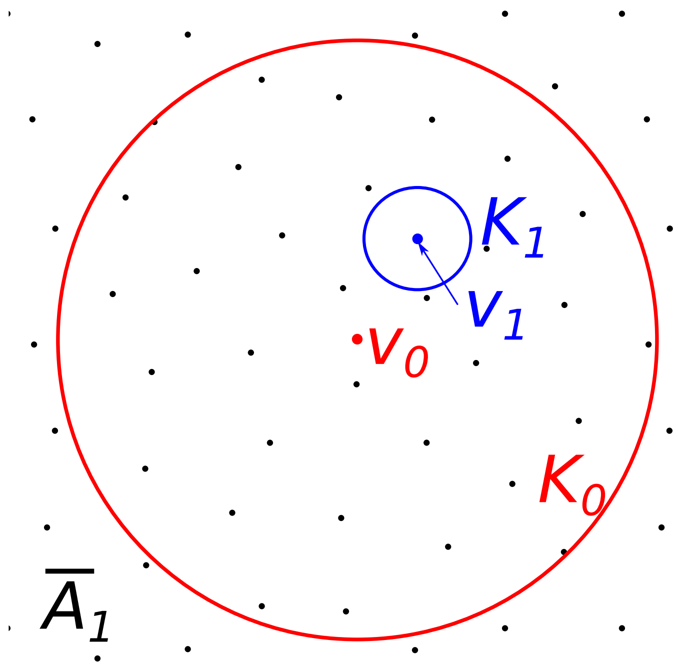
Beweis. Wir wählen für jedes n ein $v_n \in K_n$. Wegen der Bedingung (3.1) handelt es sich dabei um eine Cauchyfolge, die einen eindeutigen Grenzwert v hat. Dieser liegt als Grenzwert der Folge $(v_n)_{n\geq k}$ in jedem der (abgeschlossenen) K_k, also auch im Durchschnitt.
Enthielte der Durchschnitt einen weiteren Punkt w, so wäre wegen (3.1) $d(v,w) < \varepsilon$ für jedes $\varepsilon > 0$, also $d(v,w) = 0$. □

Satz 3.10 (Bairesches Kategorientheorem)**.** *Jede nicht leere offene Teilmenge U eines Banachraums $(V, \|\cdot\|)$ ist von zweiter Kategorie.*

Beweis. Angenommen, es gäbe eine abzählbare Familie $\{A_n\}_{n\in\mathbb{N}}$ nirgends dichter Mengen derart, dass $U = \bigcup_{n\geq 1} A_n$ ist.
Da U offen und nicht leer ist, finden wir ein $v_0 \in U$ und ein $r_0 > 0$ derart, dass $K_0 = \bar{B}_{r_0}(v_0) \subset U$ gilt.
In K_0 finden wir nun ein v_1 mit $\operatorname{dist}(v_1, \bar{A}_1) > 0$. Sonst wäre nämlich $v \in \bar{A}_1$ für alle $v \in K_0$, also $K_0 \subset \bar{A}_1$. Dann enthielte $\bar{A}_1$ aber die offene Kugel $B_{r_0}(v_0)$ im Widerspruch dazu, dass A_1 nirgends dicht ist.
Dann finden wir ein $r_1 \in (0, \frac{1}{2}r_0)$ derart, dass

$$\bar{A}_1 \cap \bar{B}_{r_1}(v_1) = \emptyset, \qquad K_1 = \bar{B}_{r_1}(v_1) \subset \bar{B}_{r_0}(v_0) = K_0$$

ist.



Konstruktion der (ersten beiden) K_n. Der Abschluss der nirgends dichten Menge A_1 ist durch schwarze Punkte angedeutet.

Iterativ erhalten wir eine absteigende Folge abgeschlossener Kugeln K_n mit Radien $r_n \to 0$, deren Durchschnitt einen einzelnen Punkt v enthält. Allerdings liegt v in keinem der $\bar{A}_n$, also insbesondere in keinem der A_n, im Widerspruch zu

$$v \in B_{r_0}(a_0) \subset U = \bigcup_{n \geq 1} A_n.$$

□

Insbesondere sind natürlich Banachräume selbst von zweiter Kategorie.

Bemerkung. *Man zeigt ganz analog, dass jeder vollständige, nicht leere metrische Raum von zweiter Kategorie (in sich selbst) ist.*

Korollar 3.11. *Ist $A \subset V$ von erster Kategorie im Banachraum $(V, \| \cdot \|)$, so ist $V \backslash A$ nicht leer und von zweiter Kategorie in V.*

Beweis. Da $V = A \cup (V \backslash A)$ von zweiter Kategorie ist, muss eine der beiden Teilmenge von zweiter Kategorie sein – das kann nur $V \backslash A$ sein. Insbesondere kann nicht $V = A$ sein, also ist $V \backslash A$ nicht leer. □

Eine klassische Anwendung des Kategorientheorems liegt in der Existenz stetiger, aber nirgends differenzierbarer Funktionen $f : [0, 1] \to \mathbb{R}$. Die Konstruktion einer solchen Funktion geht auf Karl Weierstraß zurück, der explizit ein Beispiel folgender Art angab:

$$f(t) = \sum_{n=0}^{\infty} \frac{\{10^n x\}}{10^n},$$

wobei $\{t\} = \min\{|t - k| : k \in \mathbb{Z}\}$ den Abstand von t zur nächstgelegenen ganzen Zahl bezeichnet. Der Beweis, dass die so definierte Funktion f stetig, aber nirgends differenzierbar ist, erfordert einiges an Rechenarbeit. Das folgende Beispiel beweist die Existenz solcher Funktionen nicht konstruktiv.

Beispiel 40. *Wir betrachten den Banachraum $V = C^0([0, 1]; \mathbb{R})$, versehen mit der Maximumsnorm. Darin definieren wir*

$$A := \{v \in V : v'(t_0) \text{ existiert für mind. ein } t_0 \in [0, 1]\}.$$

Ist $t_0 = 0$ oder $t_0 = 1$, handelt es sich dabei um die einseitige Ableitung.
Können wir zeigen, dass A von erster Kategorie ist. Daraus folgt, dass die Menge $V \backslash A$ aller stetigen Funktionen $v : [0, 1] \to \mathbb{R}$, die nirgends in $[0, 1]$ differenzierbar sind, nicht leer und von zweiter Kategorie in $C^0([0, 1]; \mathbb{R})$ ist.
Dazu betrachten wir die Mengen A_n, bestehend aus jenen Funktionen v, für die ein $t_0 \in [0, 1]$ derart existiert, dass für alle $h \in [-t_0, 1 - t_0]$

$$|v(t_0 + h) - v(t_0)| \leq n|h|$$

gilt. Ist $v \in A$*, so gehört* v *(wegen der Differenzierbarkeit in* t_0 *und der gleichmäßigen Stetigkeit auf* $[0, t_0]$ *sowie auf* $[t_0, 1]$*) zu einem der* A_n*, also* $A = \bigcup_{n \in N} A_n$.
Wir haben also noch zu zeigen, dass alle A_n *nirgends dicht sind. Das geschieht in zwei Schritten.*

(I) Die A_n *sind abgeschlossen. Zur Folge* $(v_k)_k$ *in* A_n *mit Grenzwert* $v \in V$ *(also* $v_k \to v$ *gleichmäßig auf* $[0,1]$*) existieren nämlich Punkte* $t_k \in [0, 1)$ *derart, dass für alle* k

$$|v_k(t_k + h) - v_k(t_k)| \leq n|h| \qquad \text{für } h \in [-t_k, 1 - t_k]$$

gilt. Die Folge $(t_k)_k$ *enthält eine im Kompaktum* $[0, 1]$ *konvergente Teilfolge* $(t_{k_l})_l$ *mit Grenzwert* $t_0 \in [0, 1]$*. Die Teilfolge* $(v_{k_l})_l$ *konvergiert immer noch gleichmäßig gegen die Grenzfunktion* v *der ursprünglichen Folge, also gilt für diese (auf* $[0, 1]$ *stetige, also gleichmäßig stetige) Funktion:*

$$|v(t_0 + h) - v(t_0)| \leq n|h| \qquad \text{für } h \in [-t_0, 1 - t_0],$$

also ist $v \in A_n$.

(II) $\mathrm{int}(\bar{A}_n) = \emptyset$*: Zu jedem* $v \in A_n = \bar{A}_n$ *und* $\varepsilon > 0$ *gibt es eine stetige, stückweise lineare Funktion* $w : [0, 1] \to \mathbb{R}$ *mit* $\|v - w\|_\infty < \varepsilon$ *und* $|w'_+(t)| > n$ *für alle* $t \in [0, 1)$ *und* $|w'_-(t)| > n$ *für* $t \in (0, 1]$*. Dieses* w *liegt also nicht in* A_n*, und damit enthält* A_n *tatsächlich keine offene Kugel.*

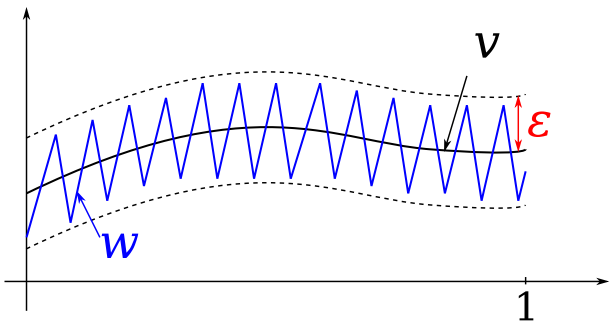


Illustration der Funktion $v \in A_n$ *und der stückweise linearen Approximation g mit überall großer (einseitiger) Ableitung. Man kann* w *sogar so wählen, dass* $|w'_+(t)| = |w'_-(t)| = C > n$ *konstant ist.*

Analog haben E. Corominas und F. Sunyer Balaguer 1954 folgenden Satz gezeigt.

Korollar 3.12. *Sind* $I \subset \mathbb{R}$ *ein nicht triviales Intervall*[a] *und* $f \in C^\infty(I; \mathbb{R})$ *derart, dass für jedes* $t \in I$ *ein (von* t *abhängiges)* $n \in \mathbb{N}$ *derart existiert, dass* $f^{(n)}(t) = 0$ *ist, dann ist* f *ein Polynom.*

[a] d.h., I soll mehr als einen Punkt enthalten

Anders ausgedrückt: Für nicht polynomiale glatte Funktionen gibt es Stellen, an denen alle Ableitungen von Null verschieden sind.
Man beachte, dass für ein Polynom f sogar ein $n_0 = \deg f + 1$ derart existiert, dass $f^{(n_0)}(t) = 0$ für alle t gilt.

Beweis. Wir beschränken uns auf den Fall eines offenen Intervalls I, sonst schränken wir einfach f auf das Innere von I ein.
Wir definieren zu gegebenem f die abgeschlossenen Mengen $A_n = \{t \in I : f^{(n)}(t) = 0\}$ und die offene Menge

$$B = \big\{t \in I : f|_{(a,b)} \text{ ist ein Polynom für gewisse } a, b \in I, a < t < b\big\}.$$

Nach Voraussetzung ist $I = \bigcup_n A_n$, also ist $\{B^c \cap A_n\}_n$ eine Überdeckung des (in I abgeschlossenen) Komplements $B^c = I \backslash B$ von B. Es gibt also zwei Möglichkeiten:

(a) B^c ist nicht leer.
(b) B^c ist leer. Dann ist f ein Polynom wie behauptet.

Wir wollen Möglichkeit (a) ausschließen. Dazu stellen wir fest:
(I) B^c enthält keine isolierten Punkte. Wäre nämlich $t_0 \in B^c$ ein isolierter Punkt in B^c, so wäre f eingeschränkt auf (a, t_0) ein Polynom für ein $t_0 > a \in I$, aber auch eingeschränkt auf (t_0, b) für ein $t_0 < b \in I$. Auf beiden dieser Intervalle wäre f dann jeweils als Polynom durch die Taylor-Entwicklung von f um t_0 gegeben (da f auf ganz (a, b) beliebig oft differenzierbar ist) – also wären beide Polynome gleich und f stimmt mit diesem Polynom auf ganz $(a, b) \ni t_0$ überein.

(II) Nach (I) ist B^c entweder leer oder hat ein nicht leeres Inneres, ist also von zweiter Kategorie. Dann müsste eines der (abgeschlossenen) $A_n \cap B^c$, sagen wir für $n = n_0$, ebenfalls ein nicht leeres Inneres (in B^c) haben, da B^c nicht als abzählbare Vereinigung nirgends dichter Mengen geschrieben werden kann. Also existiert ein Intervall (c, d) positiver Länge derart, dass $[c, d] \cap B^c \subset A_{n_0} \cap B^c$ und $f^{(n_0)}(t) = 0$ für alle $t \in (c, d) \cap B^c$. Da $(c, d) \cap B^c$ keine isolierten Punkte enthält, verschwinden dann auch alle höheren Ableitungen von f auf dieser Menge: $f^{(n)} \equiv 0$ auf $(c, d) \cap B^c$
Nun kann $(c, d) \cap B^c$ nicht ganz (c, d) sein, da sonst nach Konstruktion $(c, d) \subset B$ gälte. Daher ist $B \cap (c, d) \neq \emptyset$, enthält also (wegen der Offenheit von B) ein offenes Intervall. Ist nun $(r, s) \subset B \cap (c, d)$ ein maximales solches (d.h., $r, s \notin B$), auf dem f ein Polynom ist, dann verschwinden dessen Ableitungen $f^{(n)}$, $n \geq n_0$, in den Punkten r und s. Also ist $f^{(n_0)}(t) = 0$ auf (r, s), und damit auch alle höheren Ableitungen. Das klappt für alle derartigen Intervalle (r, s), also gilt auch $f^{(n)} \equiv 0$ auf $B \cap (c, d)$ für $n \geq n_0$ und zusammen $f^{(n)} \equiv 0$ auf $[c, d]$ für $n \geq n_0$. Dann ist aber f auf (c, d) ein Polynom und damit doch $(c, d) \subset B$.
Damit ist also B^c leer. □

3.3 Das Prinzip der gleichmäßigen Beschränktheit

Wir wenden das Bairesche Kategorientheorem nun auf Familien beschränkter Operatoren an.

Satz 3.13 (gleichmäßige Beschränktheit)**.** *Ist $\mathcal{F} \subset \mathcal{L}(V, Y)$ eine beliebige Familie beschränkter linearer Operatoren vom Banachraum V in den normierten Vektorraum Y derart, dass für alle $v \in V$*

$$\sup_{F \in \mathcal{F}} \|F(v)\|_Y < \infty$$

gilt, dann existiert eine abgeschlossene Kugel $B \subset V$ mit positivem Radius derart, dass

$$\sup_{v \in B} \sup_{F \in \mathcal{F}} \|F(v)\|_Y < \infty.$$

Kurz: Ist die Familie $\mathcal{F}$ punktweise beschränkt, so ist sie auch gleichmäßig beschränkt (auf einer Kugel). Die Behauptung gilt trivialerweise für jede endliche (insbesondere die leere) Familie $\mathcal{F}$, ist für unendliche Familien aber ein sehr mächtiges Werkzeug.

Beweis. Zu $k \in \mathbb{N}$ setzen wir

$$A_k := \bigcap_{F \in \mathcal{F}} \{v \in V : \|F(v)\|_Y \leq k\}.$$

Die A_k bilden eine Ausschöpfung von V, also eine aufsteigende Folge von Mengen mit $V = \bigcup_k A_k$. Ferner ist wegen der Stetigkeit für jedes F und k die Menge

$$\{v \in V : \|F(v)\|_Y \leq k\} = (\|\cdot\|_Y \circ F)^{-1}([0,k])$$

abgeschlossen. Damit sind auch die A_k als Durchschnitte abgeschlossener Mengen abgeschlossen. Nach dem Baireschen Kategorientheorem muss eines der A_k also einen inneren Punkt, und damit eine abgeschlossene Kugel B von positivem Radius enthalten. Nach Konstruktion der A_k gilt dann aber

$$\|F(v)\|_Y \leq k \text{ für alle } F \in \mathcal{F},\ v \in B.$$

Das ist gerade die Behauptung. □

Aus diesem Satz folgern wir sofort einen der Kernsätze der linearen Funktionalanalysis.

Theorem 3.14 (Banach-Steinhaus)**.** *Ist $\mathcal{F} \subset \mathcal{L}(V,Y)$ eine beliebige Familie beschränkter linearer Operatoren vom Banachraum V in den normierten Vektorraum Y derart, dass für alle $v \in V$*

$$\sup_{F \in \mathcal{F}} \|F(v)\|_Y < \infty$$

gilt, dann ist $\mathcal{F}$ eine beschränkte Teilmenge von $\mathcal{L}(V,Y)$:

$$\sup_{F \in \mathcal{F}} \|F\| < \infty.$$

Beweis. Sei $B = \bar{B}_r(v_0)$ die nach dem Satz über gleichmäßige Beschränktheit existierende Kugel, auf der die $F \in \mathcal{F}$ gleichmäßig beschränkt sind:

$$C := \sup_{v \in B} \sup_{F \in \mathcal{F}} \|Fv\|_Y < \infty.$$

Wegen der Linearität der F gilt

$$\|F(\varrho(v - v_0))\|_Y \leq \varrho(\|F(v)\|_Y + \|F(v_0)\|_Y) \quad \text{für alle } v \in V, \varrho > 0.$$

Die Beschränktheit überträgt sich also direkt von B auf die abgeschlossene Einheitskugel, denn $v \in \bar{B}_1(0_V) \iff rv + v_0 \in B$, also:

$$\|Fv\|_Y = \left\| F\left(\frac{1}{r}(\underbrace{rv + v_0}_{=:w \in B} - v_0)\right)\right\| \leq \frac{1}{r}\underbrace{\|F(w)\|_Y}_{C} + \frac{1}{r}\underbrace{\|F(v_0)\|}_{=:c} \quad \text{für alle } v \in \bar{B}_1(0_V)$$

und damit

$$\sup_{F\in\mathcal{F}}\|F\| = \sup_{F\in\mathcal{F}}\sup_{\|v\|_V\le 1}\underbrace{\|F(v)\|_Y}_{\le\frac{1}{r}(C+c)} \le \frac{1}{r}(C+c).$$

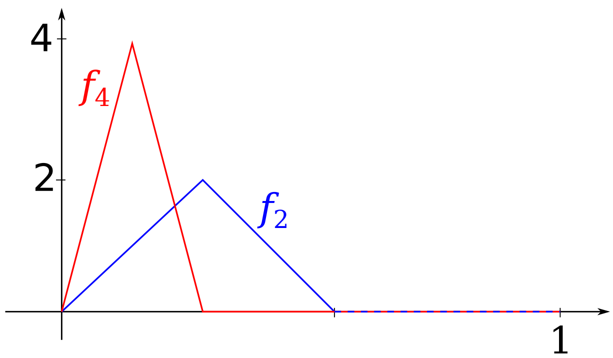
$\square$

Das folgende Beispiel zeigt, dass die Linearität eine wesentliche Rolle spielt. Für allgemeine Familien stetiger Funktionen folgt aus punktweiser Beschränktheit keineswegs die Beschränktheit im betrachteten Raum.

Beispiel 41. *Wir betrachten den (vollständigen und kompakten metrischen) Raum $V = [0,1]$, versehen mit der Standardmetrik, als Bildraum $Y = \mathbb{R}$, ebenfalls versehen mit dem Betrag als Norm, sowie im Raum $C^0([0,1])$ der stetigen (aber nicht notwendigerweise linearen) Abbildungen von V nach Y die Familie $\mathcal{F} = \{f_n\}_{n\in\mathbb{N}}$ mit*

$$f_n(x) = 2n^2x\mathbf{1}_{[0.\frac{1}{2n}]}(x) + 2n(1-nx)\mathbf{1}_{(\frac{1}{2n},\frac{1}{n}]}(x),$$

also $f_n(\frac{1}{2n}) = n$, $f_n(0) = 0 = f_n(\frac{1}{n})$ und dazwischen linear interpoliert. Diese Familie ist punktweise beschränkt, denn $f_n(0) = 0$ für alle n, und zu $x > 0$ finden wir ein $n_0 > \frac{1}{x}$. Für $n \ge n_0$ ist dann $f_n(x) = 0$, und für $n < n_0$ ist $f_n(x) \le n < n_0$. Allerdings ist $\|f_n\|_\infty = n$, die Familie $\{f_n\}_n$ in $C^0([0,1])$ also keineswegs beschränkt.



Und dabei sind die f_n sogar stückweise linear, und wir hätten als Urbildraum natürlich auch den ganzen Banachraum $\mathbb{R}$ statt $V = [0,1]$ nehmen können, ohne den Effekt zu verlieren. Der relevante Raum wäre dann der Raum $C_b^0(\mathbb{R})$ der beschränkten stetigen Funktionen, auf dem wir ebenfalls die Supremumsnorm definieren können.

Als Folgerung aus dem Satz von Banach und Steinhaus erkennen wir, dass die Untersuchung dichter Teilmengen des Urbildraums genügt.

Satz 3.15. *Sind $(V, \|\cdot\|_V)$, $(Y, \|\cdot\|_Y)$ Banachräume und $(F_n)_n$ eine Folge beschränkter linearer Operatoren $F_n \in \mathcal{L}(V, Y)$, so sind folgende Aussagen äquivalent.*

(i) *Es existiert ein $F \in \mathcal{L}(V, Y)$ derart, dass*
$$F(v) = \lim_{n\to\infty} F_n(v) \qquad \text{für alle } v \in V$$
gilt.

(ii) $\sup_n \|F_n\| < \infty$ *und es existiert eine dichte Teilmenge $D \subset V$ derart, dass $\lim_{n\to\infty} F_n(v)$ für alle $v \in D$ existiert.*

Beweis. $(i) \implies (ii)$ ist einfach der Satz von Banach und Steinhaus.

$(ii) \implies (i)$ Wir setzen $M := \sup_n \|F_n\|$. Zu gegebenem $v \in V$ und $\varepsilon > 0$ finden wir ein $w \in D$ mit $\|w - v\|_V < \frac{\varepsilon}{3(1+M)}$. Für die Cauchyfolge $(F_n(w))_n$ finden wir ferner ein n_0 derart, dass
$$\|F_n(w) - F_m(w)\|_Y < \frac{\varepsilon}{3} \qquad \text{für } n, m \geq n_0$$
gilt. Dann schätzen wir für $n, m \geq n_0$ nach Dreiecksungleichung
$$\begin{aligned}\|F_n(v) - F_m(v)\|_Y &\leq \|F_n(v) - F_n(w)\|_Y + \|F_n(w) - F_m(w)\|_Y + \|F_m(w) - F_m(v)\|_Y \\ &\leq \underbrace{\|F_n\|}_{\leq M} \|v - w\|_V + \frac{\varepsilon}{3} + \underbrace{\|F_m\|}_{\leq M} \|v - w\|_V \\ &\leq M\frac{\varepsilon}{3(1+M)} + \frac{\varepsilon}{3} + M\frac{\varepsilon}{3(1+M)} < \varepsilon.\end{aligned}$$
Also ist auch $(F_n(v))_n$ eine Cauchyfolge, die im Banachraum Y gegen einen Grenzwert konvergiert, den wir $F(v)$ nennen. Die so definierte Abbildung $F : V \to Y$ ist nach den Rechenregeln für konvergente Folgen offensichtlich linear und erfüllt im Grenzübergang
$$\|F(v)\|_Y \leq \sup_{n\in\mathbb{N}} \|F_n\|\|v\|_V,$$
ist also beschränkt. □

Die nächste Folgerung verbessert unsere Abschätzung der Norm des Grenzoperators.

Korollar 3.16. *Sind $(V, \|\cdot\|_V)$ ein Banachraum, $(Y, \|\cdot\|_Y)$ ein normierter Vektorraum und $(F_k)_k$ eine Folge beschränkter linearer Operatoren von V nach Y derart, dass $F(v) := \lim_{k\to\infty} F_k(v)$ für alle $v \in V$ existiert. Dann ist $F \in \mathcal{L}(V, Y)$ und*
$$\|F\| \leq \liminf_{k\to\infty} \|F_k\|.$$

Beweis. Die punktweise Beschränktheit ist klar, da für jedes v die Folge $(F_k(v))_k$ sogar konvergiert. Nach dem Satz von Banach und Steinhaus ist also $M := \sup_k \|F_k\| < \infty$. Dann gilt also für alle $v \in V$
$$\|F(v)\| \overset{k\to\infty}{\longleftarrow} \|F_k(v)\|_Y \leq \|F_k\|\|v\|_V \leq M\|v\|_V,$$

also ist F beschränkt. Wählen wir eine Teilfolge $(F_{k_l})_l$ mit $\|F_{k_l}\| \to \liminf_{k\to\infty} \|F_k\|$, so liefert die gleiche Abschätzung:

$$\|F(v)\| \leq \liminf_{k\to\infty} \|F_{k_l}\| \|v\|_V.$$

$\square$

Beispiel 42. *Wir können im allgemeinen im Satz von Banach und Steinhaus und in der letzten Folgerung nicht auf die Vollständigkeit des Raums V verzichten. Dazu betrachten wir den Raum $V = Y = c_{00}$ der abbrechenden Folgen, versehen mit der Norm $\|\cdot\|_\infty$. Als Operatoren betrachten wir die Multiplikationsoperatoren M_n mit $(1, 2, \dots, n, 0, \dots)$, gegeben durch*

$$M_n(v) = (1v_1, 2v_2, \dots, nv_n, 0, \dots).$$

Sie überzeugen sich schnell, dass
1. $\|M_n\| = n$ *ist, dass*
2. *für alle* $v = (v_1, \dots, v_k, 0, \dots) \in c_{00}$

$$\sup_{n\in\mathbb{N}} \|M_n(v)\|_\infty = \|(v_1, \dots, kv_k, 0, \dots)\|_\infty \leq k\|v\|_\infty < \infty$$

ist (die Familie $\{M_n\}_n$ ist punktweise beschränkt) und dass
3. *der punktweise Grenzwert M durch*

$$\lim_{n\to\infty} M_n(v) =: M(v) = (v_1, \dots, kv_k, 0, \dots)$$

gegeben ist.

Allerdings ist M nicht beschränkt, denn für alle $k \in \mathbb{N}$ liefert die Anwendung auf den k-ten Standardeinheitsvektor $e_k = (0, \dots, 0, 1, 0, \dots)$:

$$\|M(e_k)\|_\infty = \|(0, \dots, 0, k, 0, \dots)\|_\infty = k = k\|e_k\|_\infty.$$

Auch die Familie $\{M_n\}_n$ selbst ist nicht beschränkt:

$$\sup_{n\in\mathbb{N}} \|M_n\| = \sup_{n\in\mathbb{N}} n = \infty,$$

erfüllt also nicht die Behauptung des Satzes von Banach und Steinhaus (und keine der beiden Aussagen von Satz 3.15).

Eine weitere Anwendung betrifft Bilinearformen.

Definition 3.17. *Eine Bilinearform ist eine Abbildung* $b : V \times W \to \mathbb{K}$, $(v, w) \mapsto b(v, w)$, *vom Produkt der* $\mathbb{K}$*-Vektorräume* V, W *nach* $\mathbb{K}$ *derart, dass für alle* $v, v' \in V$, $w, w' \in W$ *und* $\alpha, \beta \in \mathbb{K}$

$$b(\alpha v + v', w) = \alpha b(v, w) + b(v', w), \qquad b(v, \beta w + w') = \beta b(v, w) + b(v, w')$$

gilt. Das Paar $\langle V, W\rangle$, *zusammen mit* b, *wird dann auch als duales Paar bezeichnet. Sind* $(V, \|\cdot\|_V)$, $(W, \|\cdot\|_W)$ *normierte lineare Räume, so heißt die Bilinearform* $b : V \times W \to \mathbb{K}$ *partiell stetig, falls für* $v_k \to v$ *in* V, $w_k \to w$ *in* W *und beliebige* $v_0 \in V$, $w_0 \in W$ *gilt:*

$$\lim_{k\to\infty} b(v_k, w_0) = b(v, w_0), \qquad \lim_{k\to\infty} b(v_0, w_k) = b(v_0, w).$$

Bilinearformen kennen wir natürlich schon. Zum Beispiel sind (Semi-)Skalarprodukte im Fall $\mathbb{K} = \mathbb{R}$ positiv (semi)definite, symmetrische Bilinearformen. Diese hatten sich sogar als stetig erwiesen. Der folgende Satz besagt, dass im Fall von Banachräumen die Stetigkeit einer Bilinearform schon aus der partiellen Stetigkeit folgt.

Korollar 3.18 (Mazur-Orlicz)**.** *Ist* $b : V \times W \to \mathbb{K}$ *eine partiell stetige Bilinearform auf den Banachräumen* V, W, *so ist* b *stetig, d.h.:*

$$b(v_n, w_n) \to b(v, w), \text{ falls } v_n \to v \text{ in } V, w_n \to w \text{ in } W.$$

Beweis. Wegen der Bilinearität genügt es, die Behauptung für $v = 0_V$, $w = 0_W$ zu zeigen. Zu $n \in \mathbb{N}$ definieren wir $\psi_n \in W'$ durch $\psi_n(z) := b(v_n, z)$ für $z \in W$ (man beachte, dass $b(v_n, \cdot)$ nach Voraussetzung stetig ist) und erhalten wegen der partiellen Stetigkeit im ersten Argument

$$\lim_{n\to\infty} \psi_n(z) = b(v, z) =: \psi(z).$$

Die Familie $\{\psi_n\}_n$ ist also punktweise und damit nach Banach-Steinhaus auch gleichmäßig beschränkt:

$$M := \sup_n \|\psi_n\|_{W'} < \infty.$$

Damit erhalten wir

$$|b(v_n, w_n)| = |\psi_n(w_n)| \leq \|\psi_n\|_{W'} \|w_n\|_W \leq M\|w_n\|_W \to 0 \quad \text{f. } n \to \infty,$$

also $b(v_n, w_n) \to 0 = b(0, 0)$. □

Einschub: Satz von Arzelá-Ascoli

Eine verwandte Aussage über Familien stetiger Abbildungen zwischen metrischen Räumen kennen Sie möglicherweise bereits aus der Analysis. Dazu führen wir noch einen Begriff ein.

Definition 3.19. *1. Eine Familie $\mathcal{F}$ stetiger Abbildungen $f : V \to W$ von einem metrischen Raum (V, d_V) in einen metrischen Raum (W, d_W) heißt gleichgradig stetig in $v_0 \in V$, falls zu jedem $\varepsilon > 0$ ein $\delta > 0$ derart existiert, dass*

$$d_W(f(v), f(v_0)) < \varepsilon \quad \textit{für alle } v \in B_\delta(v_0) \subset V, f \in \mathcal{F}$$

gilt.
2. $\mathcal{F}$ heißt punktweise gleichgradig stetig, falls $\mathcal{F}$ in jedem $v_0 \in V$ gleichgradig stetig ist.
3. $\mathcal{F}$ heißt gleichmäßig gleichgradig stetig (oft einfach: gleichgradig stetig), falls zu jedem $\varepsilon > 0$ ein $\delta > 0$ derart existiert, dass

$$d_W(f(v), f(v')) < \varepsilon \quad \textit{für alle } f \in \mathcal{F},\ v, v' \in V \textit{ mit } d_V(v, v') < \delta$$

gilt (das δ also nicht von v_0 abhängt).

Das *Gleichgradige* an der Stetigkeit besteht darin, dass wir in allen Fällen ein einzelnes $\delta > 0$ finden, das für *alle* $f \in \mathcal{F}$ genügt.
Der folgende Satz liefert nun eine Aussage über die gleichmäßige Beschränktheit (übersetzt in relative Kompaktheit im Raum $C(V, Y)$) von Familien stetiger Abbildungen. Erinnern Sie sich an Beispiel 41, in dem wir gesehen hatten, dass punktweise Beschränktheit dafür nicht hinreichend sein kann.

Satz 3.20 (Arzelá, Ascoli)**.** *Zum kompakten metrischen Raum (V, d) und dem Banachraum $(Y, \|\cdot\|_Y)$ versehen wir $C(V; Y)$ mit der Supremumsnorm*[a]

$$\|f\|_\infty := \sup_{v \in V} \|f(v)\|_Y$$

und betrachten eine Familie $\mathcal{F} \subset C(V; Y)$. Dann sind folgende Aussagen äquivalent.

(i) $\mathcal{F}$ ist relativ kompakt in $C(V; Y)$.

(ii) Für alle $v \in V$ ist $A_v := \{f(v) : f \in \mathcal{F}\}$ relativ kompakt in Y, und $\mathcal{F}$ ist punktweise gleichgradig stetig, d.h. für alle $v \in V$ und $\varepsilon > 0$ existiert ein $\delta > 0$ derart, dass $\|f(w) - f(v)\|_Y < \varepsilon$ für alle $w \in B_\delta(v)$, $f \in \mathcal{F}$ gilt.

[a] Wie auch $C^0(V; \mathbb{K})$ ist $C(V; Y)$ unter dieser Norm vollständig.

Beweis. $(i) \implies (ii)$. Wir zeigen beide Teile von (ii) separat.

(I) Zu $v \in V$ und $\varepsilon > 0$ finden wir wegen der relativen Kompaktheit (und damit totalen Beschränktheit) von $\mathcal{F}$ endlich viele $f_1, \dots, f_n \in C(V; Y)$ derart, dass $\mathcal{F} \subset \bigcup_{k=1}^n B_{\frac{\varepsilon}{3}}(f_k)$ gilt. Für jedes $k = 1, \dots, n$ finden wir $\delta_k > 0$ derart, dass

$$\|f_k(w) - f_k(v)\|_Y < \frac{\varepsilon}{3} \qquad \text{für alle } w \in V,\ d(v, w) < \delta_k.$$

Für alle $w \in V$, $d(v, w) < \delta = \min\{\delta_1, \dots, \delta_n\}$ und $f \in \mathcal{F}$ gilt dann mit einem $k \in$

$\{1,\dots,n\}$, für das $\|f-f_k\|_\infty < \frac{\varepsilon}{3}$ ist:

$$\begin{aligned}\|f(v)-f(w)\|_Y &\le \|f(v)-f_k(v)\|_Y + \|f_k(v)-f_k(w)\|_Y + \|f_k(w)-f(w)\|_Y \\ &\le \|f-f_k\|_\infty + \frac{\varepsilon}{3} + \|f_k - f\|_\infty \ < \ \varepsilon,\end{aligned}$$

also $\mathcal{F}$ punktweise gleichgradig stetig.

(II) Zu festem $v \in V$ setzen wir $\phi_v : C(V;Y) \to A_v$, $\phi_v(f) = f(v)$. Wegen

$$\|\phi_v(f_1) - \phi_v(f_2)\|_Y = \|f_1(v) - f_2(v)\|_Y \le \|f_1 - f_2\|_\infty \qquad \text{für } f_1, f_2 \in C(V;Y)$$

ist ϕ_v (sogar gleichmäßig) stetig, bildet also die relativ kompakte Menge $\mathcal{F}$ auf eine relativ kompakte Menge $\phi_v(\mathcal{F}) = A_v$ ab.

$(ii) \implies (i)$.Wir zeigen äquivalent zur relativen Kompaktheit die Präkompaktheit von $\mathcal{F}$ im Banachraum $(C(V;Y), \|\cdot\|_\infty)$.
(III) Zu gegebenem $\varepsilon > 0$ und $v \in V$ finden wir wegen der gleichgradigen Stetigkeit ein $\delta_v > 0$ derart, dass

$$\|f(v) - f(w)\|_Y < \frac{\varepsilon}{4} \qquad \text{für alle } w \in B_{\delta_v}(v),\ f \in \mathcal{F}. \tag{3.2}$$

Von der offenen Überdeckung $\{B_{\delta_v}(v)\}$ des kompakten Raums V genügen endlich viele:

$$V \subset \bigcup_{j=1}^{J} B_{\delta_{v_j}}(v_j). \tag{3.3}$$

Nach Voraussetzung ist jedes A_{v_j}, $j = 1,\dots,J$, relativ kompakt, also präkompakt im Banachraum Y. Wir finden zu unserem fest vorgegebenem $\varepsilon > 0$ also endlich viele

$$y_1,\dots,y_M \in A = \bigcup_{j=1}^{J} A_{v_j} \quad \text{mit} \quad A \subset \bigcup_{m=1}^{M} B_{\frac{\varepsilon}{4}}(y_m). \tag{3.4}$$

Man beachte, dass jedes y_m in einem der A_{v_j} liegt, also von der Form $y_m = f(v_j)$ für ein $f \in \mathcal{F}$, $j \in \{1,\dots,J\}$ ist.

(IV) Wir definieren die Abbildung

$$\begin{aligned}&\Phi : \underbrace{\mathrm{Abb}(\{1,\dots,J\};\{1,\dots,M\})}_{=:\Sigma \text{ (endlich!)}} \to \mathcal{P}(\mathcal{F}),\\ &\Phi(\sigma) = \left\{ f \in \mathcal{F} : \|f(v_j) - y_{\sigma(j)}\|_Y < \frac{\varepsilon}{4} \text{ für alle } j = 1,\dots,J \right\}.\end{aligned}$$

Zu gegebenem $f \in \mathcal{F}$ und $j \in \{j,\dots,J\}$ finden wir laut (3.4) ein y_m mit $\|f(v_j) - y_m\|_Y < \frac{\varepsilon}{4}$, also existiert ein $\sigma \in \Sigma$ mit $f \in \Phi(\sigma)$, also:

$$\mathcal{F} \subset \bigcup_{\sigma \in \Sigma} \Phi(\sigma).$$

Wir haben also eine endliche Überdeckung von $\mathcal{F}$ durch Mengen der Form $\Phi(\sigma)$ gefunden.

(V) Schließlich hat jedes nicht leere $\Phi(\sigma)$ Durchmesser $\operatorname{diam}\Phi(\sigma) \le \varepsilon$, kann also in eine Kugel vom Radius ε eingeschlossen werden, sagen wir $\Phi(\sigma) \subset B_\varepsilon(f_\sigma)$. Dazu seien $f, g \in \Phi(\sigma)$ beliebig. Wir haben $\|f(v) - g(v)\|_Y < \varepsilon$ für alle $v \in V$ zu zeigen. Dann ist $\|f - g\| < \varepsilon$ wie behauptet.
Zu $v \in V$ wählen wir ein v_j mit $d(v, v_j) < \delta_{v_j}$ und berechnen dann nach Dreiecksungleichung

$$\|f(v) - g(v)\| \le \underbrace{\|f(v) - f(v_j)\|}_{<\frac{\varepsilon}{4} \text{ nach } (3.2)} + \underbrace{\|f(v_j) - y_{\sigma(j)}\|}_{<\frac{\varepsilon}{4} \text{ da } f\in\Phi(\sigma)} + \underbrace{\|y_{\sigma(j)} - g(v_j)\|}_{<\frac{\varepsilon}{4} \text{ da } g\in\Phi(\sigma)} + \underbrace{\|g(v_j) - g(v)\|}_{<\frac{\varepsilon}{4} \text{ nach } (3.2)} < \varepsilon,$$

wobei hier alle Normen als $\|\cdot\|_Y$ zu lesen sind.
Wir haben also $\mathcal{F}$ durch endlich (höchstens $\sharp\Sigma$) viele ε-Kugeln überdeckt. Da $\varepsilon > 0$ ganz am Anfang beliebig gewählt war, zeigt das die Prä- (und damit relative) Kompaktheit von $\mathcal{F}$. □

Sie kennen die Aussage des Satzes vermutlich in einer etwas übersichtlicheren Form.

Korollar 3.21. *Eine Familie $\mathcal{F}$ stetiger Funktionen $f : \bar{\Omega} \to \mathbb{K}$ auf einer kompakten Menge $\bar{\Omega} \subset \mathbb{R}^n$ ist genau dann relativ kompakt in $C^0(\bar{\Omega}; \mathbb{K})$, wenn sie punktweise beschränkt und gleichgradig stetig ist.*

Hier ist $V = \bar{\Omega}$, versehen mit der euklidischen Metrik, $Y = \mathbb{K}$ mit dem Betrag als Norm und die Familie $\mathcal{F}$ sogar gleichgradig stetig. Manchmal wird sogar statt der punktweisen die gleichmäßige Beschränktheit gefordert. Das verkürzt den Beweis deutlich. In $Y = \mathbb{K}$ sind die beschränkten Mengen $\{f(v)\}_{f\in\mathcal{F}}$ natürlich relativ kompakt.

Beispiel 43. *Zu beschränktem $\Omega \subset \mathbb{R}^n$, $D \in \mathbb{N}$ und $C > 0$ ist die Familie*

$$\mathcal{F} = \{f : \bar{\Omega} \to \mathbb{K} : f \text{ Polynom mit } \deg f \le D \text{ und Koeffizienten } |c_a| \le C\}$$

relativ kompakt in $C^0(\bar{\Omega}; \mathbb{K})$.
Die Bedingungen des Satzes können Sie im Rahmen einer Übung prüfen. Ein Polynom vom Grad d wird dabei als

$$f(x_1, \ldots, x_n) = \sum_{|a|\le D} c_a x_1^{a_1} \cdots x_n^{a_n}$$

geschrieben, wobei a über alle Multiindizes der Länge $|a| \le D$ läuft.

Die Menge aller Polynome kann natürlich nicht relativ kompakt sein, da deren Abschluss nach einer Verallgemeinerung des Satzes von Stone-Weierstraß der gesamte Raum $C^0(\bar{\Omega}; \mathbb{K})$ ist, der selbstverständlich nicht kompakt ist.

3.4 Offene Abbildungen, abgeschlossene Graphen

In diesem Abschnitt widmen wir uns den Sätzen von der offenen Abbildung und dem abgeschlossenen Graphen. Da beide eng verwandt sind, bringen wir sie im gleichen Abschnitt unter. Außerdem werden wir einige Folgerungen dieser beiden Sätze kennen lernen. Wir beginnen mit der recht einleuchtenden Definition offener Abbildungen.

Definition 3.22. *Eine Abbildung $F : V \to Y$ von einem metrischen Raum (V, d_V) in einen metrischen Raum (Y, d_Y) heißt offen, falls sie offene Mengen auf offene Mengen abbildet: $f(U)$ ist offen in Y für alle offenen $U \subset V$.*

Die Definition ergibt natürlich auch ohne Metriken in allgemeinen topologischen Räumen Sinn.

Theorem 3.23 (Satz von der offenen Abbildung, Banach-Schauder; open mapping theorem (OMT))**.** *Sind $(V, \|\cdot\|_V)$ und $(Y, \|\cdot\|_Y)$ Banachräume, so sind für $F \in \mathcal{L}(V; Y)$ folgende Aussagen äquivalent.*

(i) F ist surjektiv.

(ii) F ist offen.

Beweis. (i) $\implies$ (ii). Wegen der Linearität genügt es zu zeigen, dass das Bild $F(B_1(0_V))$ der Einheitskugel in V eine Kugel $B_\delta(0_Y)$ für ein $\delta > 0$ enthält. Ist $U \subset V$ nämlich eine beliebige offene Menge, so finden wir um jedes $v \in U$ eine Kugel $B_\varepsilon(v) = \varepsilon B_1(0_V) + v$ und damit um $F(v)$ eine Kugel $B_{\varepsilon\delta}(F(v)) \subset F(B_\varepsilon(v)) \subset F(U)$.

(I) Zeige: Zu $R > 0$ existiert ein $r > 0$ derart, dass $B_r(0_Y) \subset \operatorname{int}\overline{F(B_R(0_V))}$.
Wir stellen fest, dass wegen der Surjektivität von F und wegen $V = \bigcup_{n\in\mathbb{N}} B_n(0_V)$ gilt:

$$Y = \bigcup_{n\in\mathbb{N}} \overline{F(B_n(0_V))}.$$

Dabei brauchen wir den Abschluss auf der rechten Seite nur für das Bairesche Kategorientheorem, dass uns nun sichert, dass im Banachraum Y (welcher von zweiter Kategorie ist) eine der abgeschlossenen Mengen $\overline{F(B_n(0_V))}$ einen inneren Punkt y enthalten muss. Die Surjektivität von F sichert nun die Existenz eines $v \in V$ mit $F(v) = y$, also gilt (nach Linearität) $0_Y \in \overline{F(B_n(0_V) - v)}$. Nun wählen wir R so groß, dass $B_n(0_V) - v \subset B_R(0_V)$ gilt und erhalten $0_Y \in \operatorname{int}\overline{F(B_R(0_V))}$ und damit

$$B_r(0_Y) \subset \operatorname{int}\overline{F(B_R(0_V))} \text{ für ein } r > 0.$$

(II) Zu $R = \frac{1}{3}$ finden wir also $r > 0$ mit $B_r(0_Y) \subset \operatorname{int}\overline{F(B_{\frac{1}{3}}(0_V))}$. Wir wollen $B_r(0_Y) \subset F(B_1(0_V))$ zeigen, müssen also zu beliebigem $y \in Y$, $\|y\|_Y < r$ ein $v \in B_1(0_V)$ mit $F(v) = y$ finden.

Sei also $y_0 = y \in B_r(0_Y)$ beliebig. Da $F(B_{\frac{1}{3}}(0_V))$ dicht in seinem Abschluss liegt, finden wir zu $\varepsilon = \frac{r}{2}$ ein $v_0 \in B_{\frac{1}{3}}(0_V)$ mit $\|y_0 - F(v_0)\|_Y < \frac{r}{2}$, also

$$\|\underbrace{2(y_0 - F(v_0))}_{=:y_1}\|_Y \leq r.$$

Zu $y_1 \in B_r(0_Y)$ finden wir dann wieder ein $v_1 \in B_{\frac{1}{3}}(0_V)$ mit

$$\|\underbrace{2(y_1 - F(v_1))}_{=:y_2}\|_Y \leq r.$$

Das setzen wir iterativ fort und erhalten Folgen $(y_n)_n$ in $B_r(0_Y)$ und $(v_n)_n$ in $B_{\frac{1}{3}}(0_V)$ derart, dass für alle n

$$y_{n+1} = 2(y_n - F(v_n)) \quad \overset{:2^{n+1}}{\Longrightarrow} \quad 2^{-(n+1)}y_{n+1} = 2^{-n}y_n - \underbrace{2^{-n}F(v_n)}_{=F(2^{-n}v_n)}$$

gilt. Für alle $N \in \mathbb{N}$ gilt also per Teleskopsummentrick:

$$F\left(\sum_{n=0}^{N} 2^{-n}v_n\right) = y_0 - 2^{-(N+1)}y_{N+1} \to y_0 \quad \text{für } N \to \infty, \tag{3.5}$$

denn

$$\|2^{-(N+1)}y_{N+1}\|_Y = 2^{-N+1}\underbrace{\|Y_{N+1}\|_Y}_{<r} \to 0.$$

Da die Reihe

$$\sum_{n=0}^{\infty} \|2^{-n}v_n\|_V \leq \sum_{n=0}^{\infty} 2^{-n}\frac{1}{3} = \frac{2}{3}$$

konvergiert, handelt es sich bei den Partialsummen $\sum_{n=0}^{N} 2^{-n}v_n$ um eine Cauchyfolge, die im Banachraum V gegen ein $v = \sum_{n=0}^{\infty} 2^{-n}v_n$ konvergiert, und von den Partialsummen erbt v dank der Stetigkeit der Norm auch die Beschränkung:

$$\|v\|_V = \left\|\lim_{N\to\infty}\sum_{n=0}^{N} 2^{-n}v_n\right\|_V = \lim_{N\to\infty}\left\|\sum_{n=0}^{N} 2^{-n}v_n\right\|_V \leq \frac{2}{3} < 1,$$

also $v \in B_1(0_V)$. Und schließlich liefert die Stetigkeit von F noch

$$F(v) = F\left(\lim_{N\to\infty}\sum_{n=0}^{N} 2^{-n}v_n\right) = \lim_{N\to\infty} F\left(\sum_{n=0}^{N} 2^{-n}v_n\right) \overset{(3.5)}{=} y_0.$$

Mit $\delta = r$ haben wir also $B_\delta(0_Y) \subset F(B_1(0_V))$ wie gewünscht.

$(ii) \implies (i)$: Zunächst ist $F(0_V) = 0_Y$, also $F(V) \neq \emptyset$.Da F offen ist, enthält $F(V)$ (offen in Y) einen inneren Punkt, also eine Kugel $B = B_r(y_0)$, insbesondere $y_0 = F(v_0)$

für ein $v_0 \in V$. Wegen $F(0_V) = 0_Y$ müssen wir nur noch Urbilder zu $y \in Y \backslash \{0_Y\}$ finden. Mit $a = \|y\|_Y > 0$ ist dann $\frac{r}{2a}y + y_0 \in B \subset F(Y)$, hat also ein Urbild $w \in V$:

$$F(w) = \frac{r}{2a}y + y_0, \qquad \text{also } F\left(\frac{2a}{r}(w - v_0)\right) = \underbrace{\frac{2a}{r}F(w)}_{=y+\frac{2a}{r}y_0} - \underbrace{\frac{2a}{r}F(v_0)}_{=\frac{2a}{r}y_0} = y.$$

Der gesuchte Urbildpunkt zu y ist also $v = \frac{2a}{r}(w - v_0)$. □

Beispiel 44. *Die folgenden Beispiele nicht nur stetiger, sondern sogar glatter Abbildungen von unserem einfachsten Banachraum $(\mathbb{R}, |\cdot|)$ in sich selbst zeigen, dass keine der Implikationen gesichert ist, wenn wir auf die Linearität von F verzichten.*

$f : \mathbb{R} \to \mathbb{R}$, $f(t) = \frac{1}{3}t^3 - t$ ist surjektiv, aber nicht offen (denn $f((-1.1, 1.1)) = [-\frac{2}{3}, \frac{2}{3}]$ ist nicht offen).

$\exp : \mathbb{R} \to \mathbb{R}$ *ist offen, aber nicht surjektiv.*

Beispiel 45. *Für $V = Y = C([a,b];\mathbb{R})$, versehen mit den Normen $\|\cdot\|_V = \|\cdot\|_\infty$ (Maximumnorm) bzw. $\|\cdot\|_Y = \|\cdot\|_1$ (L_1-Norm) ist die Identitätsabbildung $I : V \to Y$, $v \mapsto v$, stetig und selbstverständlich sogar bijektiv, aber nicht offen. Die Stetigkeit folgt dabei aus der Standardabschätzung für Integrale:*

$$\|I(v)\|_1 \leq (b-a)\|v\|_\infty.$$

I kann aber nicht offen sein, dann wäre nämlich auch eine umgekehrte Abschätzung gültig und dann beide Normen äquivalent. Dann wäre aber auch Y ein Banachraum unter $\|\cdot\|_Y$, was wir in Beispiel 5 ausgeschlossen haben.
Das Beispiel zeigt, dass auf die Vollständigkeit nicht verzichtet werden kann, wenn die Aussage des OMT richtig bleiben soll.

Eine wichtige Folgerung aus dem Satz von der offenen Abbildung ist nun die Aussage, dass stetige lineare Bijektionen zwischen Banachräumen sogar lineare Homöomorphismen sind.

Korollar 3.24 (Satz von der stetigen Umkehrung, Banach)**.** *Ist $F \in \mathcal{L}(V,Y)$ eine bijektive (stetige lineare) Abbildung vom Banachraum V in den Banachraum Y, so ist auch die Umkehrung $F^{-1} : Y \to V$ stetig.*

Beweis. Zur offenen Menge $U \subset V$ ist die Urbildmenge unter F^{-1}

$$(F^{-1})^{-1}(U) \;=\; F(U)$$

nach dem Satz von der offenen Abbildung offen in Y. Das ist gerade die (topologische) Definition der Stetigkeit von F^{-1}. □

In Anwendungen werden Theorem und Korollar gern in folgender Form geschrieben, die Ihnen aus der linearen Algebra bekannt vorkommen sollte (man lese $F(v) = y$ als $Av = y$).

Korollar 3.25. *Für $F \in \mathcal{L}(V, Y)$ vom Banachraum V in den Banachraum Y sind die folgenden Aussagen äquivalent.*

(i) *Die Gleichung $F(v) = y$ hat für jedes $y \in Y$ eine eindeutige Lösung $v \in V$, und diese hängt stetig von y ab.*

(ii) *Die Gleichung $F(v) = y$ hat für jedes $y \in Y$ eine Lösung, und zu $y = 0_Y$ ist $v = 0_V$ die einzige Lösung.*

Für den Satz vom abgeschlossenen Graphen müssen wir klären, was es bedeutet, dass der Graph einer Abbildung (naturgemäß eine Teilmenge des kartesischen Produkts von Urbild- und Bildraum) abgeschlossen ist. Dafür erweitern wir das Konzept des Produktraums, das wir schon als orthogonale Summe für Hilberträume kennen, auf allgemeine normierte Räume. Von Orthogonalität kann dabei natürlich keine Rede mehr sein, dafür sind wir bei der Wahl der Norm auf dem Produktraum etwas flexibler.

Definition 3.26. *Sind $(V, \|\cdot\|_V)$ und $(W, \|\cdot\|_W)$ normierte $\mathbb{K}$-Vektorräume, so definieren wir auf dem kartesischen Produkt $V \times W$ die Vektorraumoperationen*

$$(v, w)+(v', w') := (v+v', w+w'), \quad \alpha(v, w) := (\alpha v, \alpha w) \quad \text{f. } v, v' \in V,\ w, w' \in W,\ \alpha \in \mathbb{K}$$

und versehen den so entstandenen Vektorraum $V \oplus W$ mit der Norm

$$\|(v, w)\| := \|v\|_V + \|w\|_W.$$

Den so definierten normierten Vektorraum bezeichnen wir als (normierten) Produktraum $V \times W$.

Bemerkung. *Jede andere p-Norm der Form*

$$\|(v, w)\|_p = \left(\|v\|_V^p + \|w\|_W^p\right)^{\frac{1}{p}}, \quad p \in (1, \infty) \qquad \text{oder} \qquad \|(v, w)\|_\infty = \max\{\|v\|_V, \|w\|_W\}$$

täte es hier auch. Wir bleiben aber bei der Summennorm.
Erinnern oder überzeugen Sie sich, dass es sich dabei tatsächlich um Normen handelt.

Auch die Vollständigkeit erbt das Produkt von seinen Faktoren – wir kennen das Prinzip vom $\mathbb{K}^n$ (unter der Norm $\|\cdot\|_2$) und von Hilberträumen im allgemeinen.

Lemma 3.27. *Sind $(V_i, \|\cdot\|_i)$, $i = 1, \ldots, n$, Banachräume, so ist auch das Produkt $V_1 \times \ldots \times V_n$ ein Banachraum.*

Beweis. (vgl. auch Lemma 2.10) Wir müssen nur den Fall $n = 2$ betrachten, der Rest folgt induktiv. Dazu nennen wir die Räume wieder V und W und stellen fest, dass für eine Cauchyfolge $((v_k, w_k))_k$ in $V \times W$ zu $\varepsilon > 0$ ein k_0 derart existiert, dass für alle $k, m \geq k_0$

$$\varepsilon > \|(v_k, w_k) - (v_m, w_m)\| = \|v_k - v_m\|_V + \|w_k - w_m\|_W$$

gilt. Also sind $(v_k)_k$ und $(w_k)_k$ Cauchyfolgen, die in (den Banachräumen) V bzw. W Grenzwerte v bzw. w haben. Dann ist $(v, w) \in V \times W$ aber auch Grenzwert der ursprünglichen Folge, denn

$$\|(v_k, w_k) - (v, w)\| = \|v_k - v\|_V + \|w_k - w\|_W \to 0 \quad \text{für } k \to \infty.$$

□

Definition 3.28. *Sind V und Y $\mathbb{K}$-Vektorräume und $F : V \to Y$ eine beliebige Abbildung, so heißt $\Gamma_F := \{(v, F(v)) : v \in V\} \subset V \times Y$ der Graph von F.*

Bemerkung. *Der Graph ist auch für partielle Abbildungen $F : V \supset \mathcal{D}_F \to Y$ als $\{(v, F(v)) : v \in \mathcal{D}_F\}$ sinnvoll definiert.*
In jedem Fall ist der Graph Γ_F eines linearen Operators $F : V \to W$ ein Untervektorraum von $V \times W$.

Bislang ist der Graph einfach eine Teilmenge des kartesischen Produkts $V \times Y$. Im Fall normierter Räume können wir nach topologischen Eigenschaften von Γ_F im normierten Produktraum $V \times Y$ fragen. Eine dieser Eigenschaften liefert uns den Begriff des abgeschlossenen Operators.

Definition 3.29. *Eine Abbilding $F : V \to Y$ vom normierten $\mathbb{K}$-Vektorraum $(V, \|\cdot\|_V)$ in den normierten $\mathbb{K}$-Vektorraum $(Y, \|\cdot\|_Y)$ heißt (Graph-)abgeschlossen, wenn Γ_F in $V \times Y$ abgeschlossen ist.*

Manchmal werden in Analogie zu offenen auch Abbildungen *abgeschlossen* genannt, die abgeschlossene Mengen auf abgeschlossene Mengen abbilden. Dann kann auf den Zusatz “Graph-” nicht verzichtet werden. Das ist aber eine für uns wenig hilfreiche Definition, deshalb nutzen wir die angegebene und verzichten in der Regel auf den Zusatz.

Beispiel 46. *Stetige Operatoren sind abgeschlossen. Sind nämlich $F : V \to Y$ stetig und $(v, y) \in \overline{\Gamma_F}$, so finden wir eine Folge $((v_n, F(v_n)))_n$ in Γ_F derart, dass $(v_n, F(v_n)) \to (v, y)$ in $V \times Y$, also*

$$0 \longleftarrow \|(v, y) - (v_n, F(v_n))\|_{V \times Y} = \|v - v_n\|_V + \|y - F(v_n)\|_Y.$$

Beide Terme auf der rechten Seite müssen gegen Null konvergieren, also gelten

$$v_n \to v \text{ in } V \qquad \text{und} \qquad F(v_n) \to y \text{ in } Y.$$

Andererseits folgt aber wegen der Stetigkeit von F aus der ersten dieser Konvergenzen $F(v_n) \to F(v)$ in Y. Wegen der Eindeutigkeit des Grenzwerts gilt also $(v, y) = (v, F(v)) \in \Gamma_F$. Wir haben also $\overline{\Gamma_F} \subset \Gamma_F$, was nur möglich ist, wenn Γ_F abgeschlossen ist.
Beachten Sie, dass wir an keiner Stelle benutzt haben, dass F linear ist. Die Aussage gilt für beliebige stetige Abbildungen, die auf ganz V definiert sind.

Lemma 3.30. *Der lineare Operator $F : V \to Y$ ist genau dann abgeschlossen, wenn aus $v_n \to 0$ in V und $F(v_n) \to y$ in Y schon $y = 0_Y$ folgt.*

Beweis. " $\Longrightarrow$ ": Ist $(0_V, y) \in \overline{\Gamma_F} = \Gamma_F$, so folgt wegen der Linearität $y = F(0_V) = 0_Y$.
"$\Longleftarrow$": Angenommen, $(v_n, F(v_n)) \to (v, z)$ in $V \times Y$. Dann gelten $v_n - v \to 0$ in V und $F(v_n - v) = F(v_n) - F(v) \to z - F(v)$ in Y und damit nach der Bedingung

$$0 \longleftarrow F(v_n - v) \to z - F(v), \qquad \text{also } z = F(v),$$

also $(v, z) = (v, F(v)) \in \Gamma_F$. □

Wir haben im letzten Beispiel gesehen, dass stetige lineare Operatoren abgeschlossen sind. Das folgende Beispiel zeigt, dass die Umkehrung im allgemeinen falsch ist.

Beispiel 47. *Wir betrachten den Raum $V = c_{00}$ der abbrechenden Folgen, versehen mit der Norm $\|\cdot\|_2$. Zu einer beliebigen Zahlenfolge $a = (a_n)_n$ betrachten wir den Multiplikationsoperator $M_a : V \to V$, $M_a(v) = (a_n v_n)_n$. Selbstverständlich ist M_a linear und bildet V in sich selbst ab.*
M_a ist auch abgeschlossen. Gelten nämlich $v^{(k)} \to 0_V$ und $M_a(v^{(k)}) \to y$ in V (also $\|(a_n v^{(k)})_n - (y_n)_n\|_2 \to 0$ für $k \to \infty$), so gilt insbesondere komponentenweise

$$\lim_{n\to\infty} v_n^{(k)} = 0, \ \textit{also} \ \lim_{n\to\infty} a_n v_n^{(k)} = a_n \lim_{n\to\infty} v_n^{(k)} = 0,$$

aber auch

$$\lim_{k\to\infty} a_n v_n^{(k)} = y_n,$$

also $y_n = 0$ für alle $n \in \mathbb{N}$.
Allerdings ist M_a nur beschränkt, wenn $(a_n)_n$ selbst beschränkt ist. Ein Gegenbeispiel (für $a_n = n$) haben wir schon gesehen.

Bei diesem Beispiel ist V unter der gegebenen Norm nicht abgeschlossen. Tatsächlich gilt nämlich sogar:

Theorem 3.31 (Satz vom abgeschlossenen Graphen, closed graph theorem (CGT))**.** *Sind V und Y Banachräume, so ist jeder abgeschlossene lineare Operator $F : V \to Y$ stetig.*

Beweis. Da F abgeschlossen ist, ist Γ_F ein abgeschlossener linearer Teilraum des Banachraums (!) $V \times Y$ und als solcher selbst ein Banachraum. Wir definieren die natürliche Projektion

$$P : \Gamma_F \to V, \qquad (v, F(v)) \mapsto v.$$

Diese ist linear und beschränkt, denn $\|v\|_V \leq \|v\|_V + \|F(v)\|_Y = \|(v, F(v))\|_{V\times Y}$. Außerdem ist P bijektiv (völlig unabhängig von den Eigenschaften von F).
Nach dem Satz über die stetige Umkehrabbildung (Korollar 3.24) ist $P^{-1} : V \to \Gamma_F$, $v \mapsto (v, F(v))$, ebenfalls stetig.
Gilt also $v_n \to v$ (in V), so folgt

$$(v_n, F(v_n)) = P^{-1}(v_n) \to P^{-1}(v) = (v, F(v)),$$

also $F(v_n) \to F(v)$ (in V). Damit ist F also (folgen-)stetig. □

Beispiel 48. *Für $V = C^1([0,1];\mathbb{R})$, versehen mit Norm $\|v\|_V = \max_t |v(t)|$ und $Y = C^0([0,1];\mathbb{R})$, ebenfalls versehen mit der Maximumnorm, betrachten wir den Ableitungsoperator $D : V \to Y$, $v \mapsto v'$. Dieser ist unbeschränkt (man betrachte etwa die Funktionen $v_n \in V$, $v_n(t) = t^n$ für $n \in \mathbb{N}$). Sie rechnen aber (als Übung) schnell nach, dass D abgeschlossen ist.*
Wieder sehen wir anhand dieses Beispiels, dass auf die Vollständigkeit der beteiligten Räume nicht ohne weiteres verzichtet werden kann.

3.5 Der Fortsetzungssatz von Hahn-Banach

Im wesentlichen werden wir uns in diesem Abschnitt die Frage stellen, ob zu einem gegebenem normierten Raum V "genügend" viele stetige lineare Funktionale existieren, um den Raum zu charakterisieren. Genauer stellt sich die Frage, ob es zu jedem $v \in V\backslash\{0\}$ ein $\phi \in V'$ mit $\phi(v) \neq 0$ gibt. Die Idee dazu ist folgende. Wir setzen $\phi(v) = 1$ und haben damit ϕ auf dem von v aufgespannten Unterraum $<\{v\}>$ eindeutig bestimmt. Nun fragen wir uns, ob wir ϕ linear und *stetig* auf ganz V fortsetzen können.
Wir beginnen zunächst mit den uns schon bekannten Halbnormen und stellen fest, dass wir solche durch stetige lineare Funktionale darstellen können.

Lemma 3.32. *Ist $(V, \|\cdot\|)$ ein normierter $\mathbb{K}$-Vektorraum, und sind $\phi_1, \ldots, \phi_n \in V'$ stetige lineare Funktionale, so definiert*

$$p(v) := \sum_{j=1}^{n} |\phi_j(v)|$$

eine Halbnorm auf V.

Beweis. $p(v) \in [0, \infty)$ ist klar (man beachte, dass die Summe endlich ist).
Sind $\alpha \in \mathbb{K}$ und $v \in V$, so ist

$$p(\alpha v) = \sum_{j=1}^{n} |\phi_j(\alpha v)| = \sum_{j=1}^{n} |\alpha||\phi_j(v)| = |\alpha| \sum_{j=1}^{n} |\phi_j(v)| = |\alpha| p(v).$$

Und schließlich gilt

$$|\phi_j(v + w)| = |\phi_j(v) + \phi_j(w)| \leq |\phi_j(v)| + |\phi_j(w)|$$

für $v, w \in V$ und $j = 1, \ldots, n$, also nach Summation:

$$p(v + w) \leq p(v) + p(w).$$

□

Beispiel 49. (*a*) *Mit $\phi_j(x) = x_j$, $j \in J \subset \{1, \ldots, n\}$ erhalten wir auf $V = \mathbb{K}^n$ die Halbnorm*

$$p(v) = \sum_{j \in J} |x_j|,$$

die genau dann zur (Summen-)Norm wird, wenn $J = \{1, \dots, n\}$ ist.

(b) Auf $V = C^0([0,1])$ definiert $\delta_{t_0}(f) = f(t_0)$ zu gegebenem $t_0 \in [0,1]$ ein stetiges lineares Funktional. Damit erhalten wir einen ganzen Zoo von Halbnormen:

$$p(f) = |f(t_0)|$$

und für $t_1, \dots, t_n \in [0,1]$:

$$q(f) = \sum_{k=1}^{n} |f(t_k)|.$$

Wir können Halbnormen nutzen, um Halbordnungen auf einem Vektorraum V zu definieren.

Definition 3.33. *Eine binäre Relation $\preceq$ auf einer nicht leeren Menge X heißt eine Halbordnung, wenn für alle $x, y, z \in V$*

(i) $x \preceq x$ (Reflexivität)

(ii) $x \preceq y,\ y \preceq x \implies x = y$ (Antisymmetrie)

(iii) $x \preceq y,\ y \preceq z \implies x \preceq z$ (Transitivität)

gelten.
Eine Teilmenge $A \subset X$ einer halbgeordneten Menge X heißt total geordnet, falls für alle $a, b \in A$ mit $a \neq b$ entweder $a \preceq b$ oder $b \preceq a$ gilt. Auf A bildet $\preceq$ dann eine (totale) Ordnung.

Eine totale Ordnung ist also eine Halbordnung, bezüglich der alle Paare von Elementen vergleichbar sind.
Die übliche Schreibweise $v \succeq w$ für $w \preceq v$ dürfte niemanden überraschen.

Beispiel 50. *(a) Die klassische Ordnung $\leq$ auf den reellen Zahlen ist eine totale Ordnung und damit insbesondere eine Halbordnung auf $\mathbb{R}$.*

(b) Auf $\mathbb{R}^n$ können wir eine Halbordnung durch

$$x \preceq y \ :\iff\ x_j \leq y_j \quad \text{für alle } j = 1, \dots, n$$

definieren. In $\mathbb{R}^2$ ist dann zum Beispiel jede Gerade positiver Steigung eine total geordnete Teilmenge.
Lassen wir j durch alle natürlichen Zahlen laufen, so erhalten wir eine Halbordnung auf $\mathbb{R}^{\mathbb{N}}$ oder seinen Unterräumen $\ell_p^{\mathbb{R}}$.
Auf $\mathbb{C}^{\mathbb{N}}$ lautet die analoge Definition

$$z \preceq w :\iff \operatorname{Re} w_j \leq \operatorname{Re} w_j,\ \operatorname{Im} w_j \leq \operatorname{Im} w_j \quad \text{für alle } j \in \mathbb{N}.$$

Beispiel 51. *Analog definiert*

$$f \preceq g :\iff f(t) \leq g(t) \quad \text{für alle } t \in [a,b]$$

eine Halbordnung auf $C([a,b];\mathbb{R})$.

Definition 3.34. *1. Ist $(X, \preceq)$ eine halbgeordnete Menge, so nennen wir $x \in X$ eine obere Schranke für $A \subset X$, wenn $a \preceq x$ für alle $a \in A$ gilt.*

2. $m \in A$ heißt maximales Element von $A \subset X$, falls für alle $a \in A$ aus $m \preceq a$ schon $a = m$ folgt.

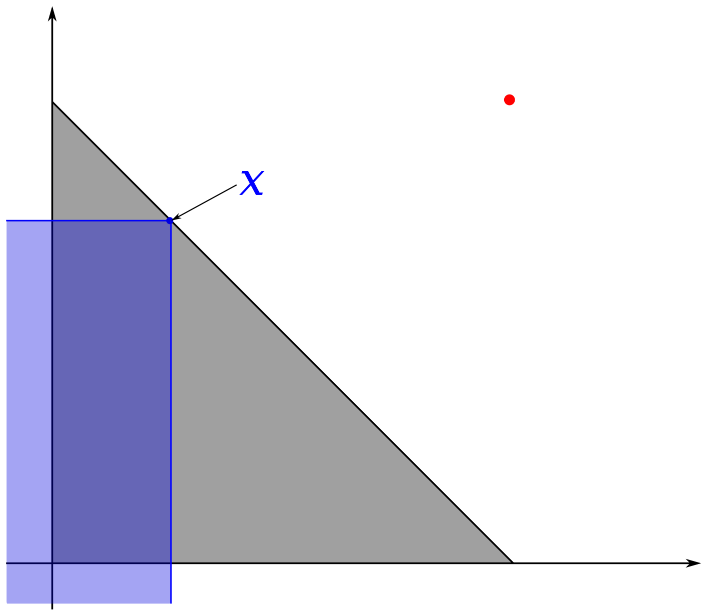
Ein Element $m \in A$ ist also maximal, falls es für jedes gegebene $a \in A$ entweder nicht mit m vergleichbar ist oder $a \preceq m$ erfüllt. Es gibt also in ganz A kein bezüglich $\preceq$ "echt größeres" Element. Beachten Sie, dass wir nie von "dem" maximalen Element in A gesprochen haben. Weder haben wir behauptet, jede Menge enthalte ein maximales Element, noch hatten wir ausgeschlossen, A könne mehrere maximale Elemente enthalten.

Beispiel 52. (a) *In unserer total geordneten Menge $\mathbb{R}$ kennen wir das Prinzip oberer Schranken schon. Maximale Elemente gibt es aber auch dort schon nicht in jeder Teilmenge. Suchen Sie etwa mal ein maximales Element in $(0, 1)$. Aber auch obere Schranken gibt es natürlich nur für nach oben beschränkte Mengen. Die Teilmenge $\mathbb{N} \subset \mathbb{R}$ etwa besitzt keine obere Schranke.*

(b) *Versehen wir $\mathbb{R}^2$ mit unserer komponentenweise definierten Halbordnung $\preceq$ aus Beispiel 50 (b) und betrachten die Teilmenge*

$$N = \{(x_1, x_2) \in \mathbb{R}^2 : x_1, x_2 \geq 0, x_1 + x_2 \leq 1\}$$

(der Standardsimplex in $\mathbb{R}^2$ oder schlicht das Dreieck mit den Eckpunkten $(1, 0)$, $(0, 0)$ und $(0, 1)$), so finden wir $(0, 1)$ und $(1, 0)$ als maximale Elemente.



Simplex N (grau) und zu gegebenem $x = (t, 1-t)$ der Kegel aller $y \in \mathbb{R}^2$ mit $y \preceq x$ (blau).

Allgemeiner sind sogar alle Elemente der Form $x = (t, 1-t)$, $t \in [0, 1]$, maximale Elemente. Haben wir nämlich ein $y = (y_1, y_2) \in N$ mit $y_1 \geq t$ und $y_2 \geq 1-t$, so folgt daraus

$y_1 + y_2 \geq 1$, also wegen der Definition von N sofort $y_1 + y_2 = 1$. Damit liegt y auf der Hypothenuse des Dreiecks. Zwei Punkte auf dieser Strecke können aber nur miteinander verglichen werden, wenn sie gleich sind.
Keines dieser maximalen Elemente ist aber eine obere Schranke von N. Eine offensichtliche solche ist $(1, 1)$ (rot in der Skizze).

ANHANG A

Das Lebesgue-Integral

In diesem Anhang sammeln wir ohne jegliche Beweise einige Aussagen aus der Maßtheorie, die wir für unsere Betrachtungen benötigen.

Definition A.1. *Eine σ-Algebra über einer nicht leeren Menge Ω ist ein System $\mathcal{F}$ von Teilmengen von Ω derart, dass*

- $\emptyset \in \mathcal{F}$
- $A \in \mathcal{F} \implies A^c \in \mathcal{F}$ *($A^c = \Omega \backslash A$ ist das Komplement von A)*
- $A_1, A_2, \dots \in \mathcal{F} \implies \bigcup_{k \geq 1} A_k \in \mathcal{F}$.

Die zu $\mathcal{F}$ gehörenden Mengen nennen wir messbare Mengen, das Paar $(\Omega, \mathcal{F})$ heißt messbarer Raum.

Lemma A.2 (und Definition)**.** *Ist $\mathcal{E}$ eine beliebige Familie von Teilmengen von Ω, so existiert eine kleinste σ-Algebra*

$$\sigma(\mathcal{E}) = \bigcap_{\mathcal{F}\ \sigma\text{-Algebra}, \mathcal{E} \subset \mathcal{F}} \mathcal{F},$$

die $\mathcal{E}$ enthält. Diese wird als von $\mathcal{E}$ erzeugte σ-Algebra bezeichnet.

Die auf dem $\mathbb{R}^n$ vom System der offenen Menge erzeugte σ-Algebra heißt Borelsche σ-Algebra und wird als $\mathcal{B}_n$ bezeichnet[1]. Sie wird auch von der Familie aller halboffenen Quader

$$(a, b] = (a_1, b_1] \times \cdots \times (a_n, b_n], \qquad a = (a_1, \dots, a_n)^T, b = (b_1, \dots, b_n)^T \in \mathbb{R}^n$$

[1]Allgemeiner heißt für jeden topologischen Raum die von den offenen Mengen erzeugte σ-Algebra Borelsche σ-Algebra.

erzeugt.

Definition A.3. *Ein Maß auf einer σ-Algebra $\mathcal{F}$ über der Menge Ω ist eine Abbildung $\mu : \mathcal{F} \to \infty$ derart, dass*

- $\mu(\emptyset) = 0$
- $\mu\left(\bigcup_{k\geq 1} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k)$ *für beliebige paarweise disjunkte $A_1, A_2, \dots \in \mathcal{F}$.*

Zusammen heißt $(\Omega, \mathcal{F}, \mu)$ dann ein Maßraum.

Satz A.4 (und Definition)**.** *Auf der Borelschen σ-Algebra $\mathcal{B}_n$ existiert genau ein Maß β^n, das für alle Quader $(a, b]$ mit $a \leq b$*

$$\beta^n((a,b]) = \mathrm{vol}_n((a,b]) = (b_1 - a_1)\cdots(b_n - a_n)$$

erfüllt.

Allgemeiner wird jedes Maß μ auf der Borelschen σ-Algebra $\mathcal{B}(X)$ über einem topologischen Raum X als Borelmaß bezeichnet. Ein solches heißt regulär, wenn alle kompakten Mengen endliches Maß haben und für alle $A \in \mathcal{B}(X)$

$$\mu(A) = \sup\{\mu(K) : K \subset A \text{ kompakt}\} = \inf\{\mu(U) : U \supset A \text{ offen}\} \tag{A.1}$$

gilt, also jede messbare Menge bezüglich μ beliebig gut von innen durch kompakte und von außen durch offene Mengen approximiert werden kann.

Satz A.5. *Das Maß β^n aus Satz A.4 ist ein reguläres Borelmaß über dem mit seiner (euklidischen) Standardtopologie versehenen $\mathbb{R}^n$.*

Definition A.6. *In einem Maßraum $(\Omega, \mathcal{F}, \mu)$ heißt $N \in \mathcal{F}$ eine (μ-)Nullmenge, wenn $\mu(N) = 0$ ist. Der Maßraum $(\Omega, \mathcal{F}, \mu)$ heißt vollständig, wenn alle Teilmengen von Nullmengen messbar sind (also zu $\mathcal{F}$ gehören):*

$$B \subset N \in \mathcal{F},\ \mu(N) = 0 \quad \Longrightarrow \quad B \in \mathcal{F}.$$

Satz A.7 (und Definition)**.** *Ist $(\Omega, \mathcal{F}, \mu)$ ein Maßraum, so ist durch*

$$\tilde{\mathcal{F}} := \{A \cup B : A \in \mathcal{F},\ B \subset N \text{ für eine Nullmenge } N \in \mathcal{F}\}$$

eine σ-Algebra über Ω gegeben, auf der durch

$$\tilde{\mu}(A \cup B) = \mu(A), \qquad \text{falls } A \in \mathcal{F}, B \subset N \text{ für eine Nullmenge } N$$

ein Maß definiert ist[a]. Für $A \in \mathcal{F}$ gilt $\tilde{\mu}(A) = \mu(A)$, und der vollständige Maßraum $(\Omega, \tilde{\mathcal{F}}, \tilde{\mu})$ heißt die Vervollständigung von $(\Omega, \mathcal{F}, \mu)$.

[a]dessen Definition wohlgemerkt nicht von der Wahl von A und B abhängt. Ist nämlich für ein $A' \in \mathcal{F}$ und ein $B' \subset N'$ mit einer Nullmenge N' $A \cup B = A' \cup B'$, so gilt $\mu(A) = \mu(A')$.

Satz A.8 (und Definition). *Der Maßraum* $(\mathbb{R}^n, \mathcal{B}_n, \beta^n)$ *ist nicht vollständig. Seine Vervollständigung bezeichnen wir mit* $(\mathbb{R}^n, \mathfrak{M}_n, \lambda^n)$ *und nennen* λ^n *das n-dimensionale Lebesgue-Maß. Die Elemente von* $\mathfrak{M}_n$ *heißen (Lebesgue-)messbare Teilmengen des* $\mathbb{R}^n$.

Bislang haben wir nur das Konzept des Volumens von Quadern auf beliebige messbare Mengen erweitert. Da Quader aber eine wesentliche Rolle bei der Definition des Riemannintegrals spielten, sollte uns das bei der Definition eines allgemeineren Integralbegriffs helfen. Dazu brauchen wir zunächst Funktionen, über deren Integral wir uns Gedanken machen können.

Definition A.9. *Sind* $(\Omega, \mathcal{F})$, $(\Omega', \mathcal{F}')$ *messbare Räume, so nennen wir die Abbildung* $f : \Omega \to \Omega'$ *($\mathcal{F}$-$\mathcal{F}'$-)messbar, falls* $f^{-1}(A') \in \mathcal{F}$ *für alle* $A' \in \mathcal{F}'$ *gilt.*

Wir befassen uns vor allem mit reell- oder komplexwertigen Funktionen auf dem $\mathbb{R}^n$ und wählen dabei auf dem Urbildraum die Lebesguesche σ-Algebra $\mathfrak{M}_n$ und auf dem Bildraum die Borelsche. Dabei wollen unseren Funktionen aber auch erlauben, unendliche Werte anzunehmen. Solche Funktionen $f : \Omega \to \bar{\mathbb{R}}(= \mathbb{R} \cup \{-\infty, \infty\})$ bezeichnen wir als numerische Funktione. Dabei versehen wir $\bar{\mathbb{R}}$ mit der aus Mengen der Form $A, A \cup \{\infty\}, A \cup \{-\infty\}, A \cup \{-\infty, \infty\}$ bestehenden σ-Algebra $\bar{\mathcal{B}}_1$.

Definition A.10. *Eine Funktion* $f : \mathbb{R}^n \to \mathbb{R}$ *heißt messbar, wenn sie* $\mathfrak{M}_n$-$\mathcal{B}_1$*-messbar ist.*
Eine numerische Funktion $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ *heißt messbar, wenn sie* $\mathfrak{M}_n$-$\bar{\mathcal{B}}_1$*-messbar ist.*

Nicht nur für auf dem $\mathbb{R}^n$, sondern für auf allgemeinen messbaren Räumen definierte (numerische) Funktionen gilt:

Lemma A.11. *Eine (numerische) Funktion* f *auf einem messbaren Raum* $(\Omega, \mathcal{F})$ *ist genau dann ($\mathcal{F}$-$\mathcal{B}_1$- bzw. $\mathcal{F}$-$\bar{\mathcal{B}}_1$-)messbar, wenn für alle* $c \in \mathbb{R}$

$$\{f \le c\} = \{x \in \Omega : f(x) \le c\} \in \mathcal{F}$$

gilt.

Die messbaren reellwertigen Funktionen bilden einen Vektorraum:

Lemma A.12. *Sind* $f_1, f_2, \dots$ *messbare reellwertige Funktionen auf* $(\Omega, \mathcal{F})$*, so sind auch*

- $\alpha f + g$ *für* $\alpha \in \mathbb{R}$
- $|f|$, f^+, f^- *(wobei* $f^+(x) = \max\{f(x), 0\}$, $f^- = \max\{-f(x), 0\}$*)*
- $\sup_k f_k$, $\inf_k f_k$
- $\limsup_{k\to\infty} f_k$, $\liminf_{k\to\infty} f_k$

(jeweils punktweise verstanden) messbare (numerische) Funktionen.

Insbesondere sind also punktweise Grenzwerte (sofern existent) messbarer Funktionen wieder messbar. Das nutzen wir zur Definition des Integrals aus.

Definition A.13. *Als* <u>*Treppenfunktion*</u> *bezeichnen wir eine messbare Funktion* $f : \Omega \to \mathbb{R}$*, die nur endlich viele Werte annimmt:*

$$f = \sum_{k=1}^{N} c_k \mathbf{1}_{A_k}, \qquad A_1, \dots, A_N \in \mathcal{F},\ c_1, \dots, c_N \in \mathbb{R}.$$

Das <u>*Integral*</u> *einer Treppenfunktion* $f = \sum_{k=1}^{N} c_k \mathbf{1}_{A_k}$ *über* Ω *bezüglich* μ *ist als*

$$\int_\Omega f \,\mathrm{d}\mu := \sum_{k=1}^{N} c_k \mu(A_k)$$

definiert[a]*, falls* $\mu(A_k) < \infty$ *für alle* $k = 1, \dots, N$ *gilt. In diesem Fall heißt* f <u>*(μ-)integrierbar*</u>*.*

[a] Dabei ist $\mathbf{1}_A$ die Indikatorfunktion der Menge A, also $\mathbf{1}_A(x) = 1$ für $x \in A$ und $\mathbf{1}_A(x) = 0$ für $x \notin A$. Der Wert des Integrals hängt nicht von der konkreten Darstellung der Treppenfunktion f ab.

Lemma A.14. *Ist* f *eine nicht negative, messbare, numerische Funktion auf* $(\mathbb{R}^n, \mathfrak{M}_n, \lambda^n)$*, so existiert eine (punktweise) monoton wachsende Folge* $(f_k)_k$ *von Treppenfunktionen, die punktweise gegen* f *konvergiert.*

Definition A.15. *Für eine nicht negative, messbare, numerische Funktion auf* $(\mathbb{R}^n, \mathfrak{M}_n, \lambda^n)$ *definieren wir das (Lebesgue-)Integral als*

$$\int_{\mathbb{R}^n} f \,\mathrm{d}\lambda^n := \lim_{k\to\infty} \int_{\mathbb{R}^n} f_k \,\mathrm{d}\lambda^n,$$

wobei $(f_k)_k$ *eine monoton gegen* f *konvergierende Folge integrierbarer Treppenfunktionen ist.*

In der Regel schreiben wir $\int_{\mathbb{R}^n} f \,\mathrm{d}x$ statt $\int_{\mathbb{R}^n} f \,\mathrm{d}\lambda^n$, unsere Integrale sind also immer, sofern nicht anders gesagt, Lebesgue-Integrale. Wir bemerken, dass diese Definition nicht von der Wahl der konkreten Folge von Treppenfunktionen abhängt und dass das Integral durchaus den Wert ∞ annehmen kann.

Für allgemeine numerische Funktionen zerlegen wir f in $f^+ - f^-$ und definieren:

Definition A.16. *Für eine messbare numerische Funktion auf* $(\mathbb{R}^n, \mathfrak{M}_n, \lambda^n)$ *definieren wir*

$$\int_{\mathbb{R}^n} f \,\mathrm{d}x = \int_{\mathbb{R}^n} f^+ \,\mathrm{d}x - \int_{\mathbb{R}^n} f^- \,\mathrm{d}x,$$

falls mindestens eines der Integrale auf der rechten Seite endlich ist. Sind sogar beide dieser Integrale (und damit auch die linke Seite) endlich, so nennen wir f *(Lebesgue-)integrierbar (über* $\mathbb{R}^n$*).*

Für eine messbare Menge $\Omega \subset \mathbb{R}^n$ *und eine messbare numerische Funktion* f *definieren wir*

$$\int_{\Omega} f \,\mathrm{d}x := \int_{\mathbb{R}^n} f\mathbf{1}_\Omega \,\mathrm{d}x,$$

falls das Integral von $f\mathbf{1}_\Omega$ *im obigen Sinne existiert. Ist dieses Integral endlich, so nennen wir* f *integrierbar über* Ω*.*

Die Definition gilt auch für beliebige Maßräume, allerdings muss man dann schon für Treppenfunktionen den Wert ∞ für das Integral erlauben, da man im allgemeinen keine Folge *integrierbarer* Treppenfunktionen finden wird, die monoton gegen f konvergiert.

Wollen wir Aussagen über die Elemente eines Maßraums treffen, so erweist sich folgende Sprechweise oft als nützlich. Wir sagen, eine Eigenschaft gelte (μ-)fast überall auf Ω, wenn eine μ-Nullmenge N derart existiert, dass die fragliche Eigenschaft für alle $x \in \Omega \setminus N$ gilt.

Satz A.17 (und Definition)**.** *Die numerischen Funktionen* $f, g : \Omega \to \bar{\mathbb{R}}$ *seien messbar. Dann gilt*

- *f ist genau dann integrierbar, wenn $|f|$ integrierbar ist, und es gilt die Dreiecksungleichung für Integrale:*
$$\left|\int_\Omega f \,\mathrm{d}\mu\right| \le \int_\Omega |f| \,\mathrm{d}\mu.$$
- *Monotonie: Gilt $f \le g$ fast überall $\implies$ $\int_\Omega f \,\mathrm{d}\mu \le \int_\Omega g \,\mathrm{d}\mu$, falls beide Integrale (in $\bar{\mathbb{R}}$) existieren.*
- *Sind f und g integrierbar, so ist auch $\alpha f + g$ für alle $\alpha \in \mathbb{R}$ integrierbar und $\int_\Omega \alpha f + g \,\mathrm{d}\mu = \alpha \int_\Omega f \,\mathrm{d}\mu + \int_\Omega g \,\mathrm{d}\mu$. Die integrierbaren reellwertigen Funktionen bilden also einen (reellen) Vektorraum, den wir $\mathcal{L}_1(\mu)$ nennen, und das Integral ist eine lineare Abbildung $\int_\Omega \cdot \,\mathrm{d}\mu : \mathcal{L}_1(\mu) \to \mathbb{R}$.*
- *Gilt $f = g$ fast überall und existiert eines der Integrale $\int_\Omega f \,\mathrm{d}\mu$ oder $\int_\Omega g \,\mathrm{d}\mu$, so existieren beide Integrale und stimmen überein. Ferner gilt dann $\int_\Omega |f - g| \,\mathrm{d}\mu = 0$.*

Die letzte Aussage erlaubt die Definition des Raums L_1.

Definition A.18. *Auf der Menge der messbaren numerischen Funktionen* $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ *definieren wir die Äquivalenzrelation* $\sim$ *durch*

$$f \sim g :\Longleftrightarrow f = g \text{ fast überall.}$$

Mit $L_1(\mathbb{R}^n)$ *bezeichnen wir den Raum all jener Äquivalenzklassen, die einen integrierbaren Repräsentanten enthalten. Analog definieren wir* $L_1(\Omega)$ *für messbare Mengen* $\Omega \subset \mathbb{R}^n$.

Wir bemerken, dass jede Äquivalenzklasse aus $L_1(\mathbb{R}^n)$ auch einen reellwertigen Repräsentanten $f \in \mathcal{L}_1(\mathbb{R}^n)$ enthält. Oft werden wir die Äquivalenzklasse $[f] \in L_1(\mathbb{R}^n)$ kurz als f schreiben und von L_1-Funktionen sprechen. Analog erhalten wir auch die Räume $L_p(\mathbb{R}^n)$, oder allgemeiner $L_p(\mu)$, für $p \geq 1$. Die entsprechenden Definitionen sind

$$\mathcal{L}_p(\mu) = \left\{ f : \Omega \to \bar{R} \ : \ f \text{ messbar} \ , \ \int |f| \, \mathrm{d}\mu < \infty \right\}$$

für $p \in [1, \infty)$ und

$$\mathcal{L}_\infty(\mu) = \left\{ f : \Omega \to \bar{R} \ : \ f \text{ messbar} \ , \ \exists C \geq 0 : |f(\omega)| \leq C \text{ für } \mu\text{-fast alle } \omega \in \Omega \right\}.$$

Dann fassen wir wieder alle fast überall übereinstimmenden Funktionen zu einer Äquivalenzklasse zusammen und erhalten die Räume L_p.
Aus der Definition haben wir sofort die Normen

$$\|f\|_p := \left(\int_\Omega |f|^p \, \mathrm{d}\mu \right)^{\frac{1}{p}}, \qquad \|f\|_\infty := \underbrace{\inf\{C > 0 : |f(\omega)| \leq C \text{ fast überall}\}}_{=:\operatorname{ess\,sup}_{\omega \in \Omega} |f(\omega)|}$$

abgelesen.

Lemma A.19 (Hölder-Ungleichung). *Sind* $p, p' \in (1, \infty)$ *mit* $\frac{1}{p} + \frac{1}{p'} = 1$ *oder* $p = 1$, $p' = \infty$, *so gilt für* $f \in L_p(\mu)$, $g \in L_{p'}(\mu)$:

$$\|fg\|_1 \leq \|f\|_p \|g\|_{p'},$$

insbesondere: $fg \in L_1(\mu)$.

Die Hölder-Ungleichung wird zum Beweis der Minkowski-Ungleichung benutzt.

Lemma A.20 (Minkowski-Ungleichung). *Für* $f, g \in L_p(\mu)$, $p \in [1, \infty)$, *gilt*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

Theorem A.21 (Vollständigkeitssatz von Riesz-Fischer). *Die Räume* $(L_p(\mu), \|\cdot\|_p)$ *sind Banachräume für* $p \in [1, \infty]$.

Satz A.22 (Ausschöpfungslemma)**.** *Im Maßraum* $(\Omega, \mathcal{F}, \mu)$ *seien* $A, A_1, A_2, \ldots$ *messbare Mengen derart, dass*

$$A_1 \subset A_2 \subset \ldots \qquad \text{und} \qquad A = \bigcup_{k \in \mathbb{N}} A_k$$

gelten. Dann ist $f : \Omega \to \bar{\mathbb{R}}$ *genau dann integrierbar über* A*, wenn* $f\mathbf{1}_{A_k}$ *für jedes* $k \in \mathbb{N}$ *integrierbar ist und* $\lim_{k\to\infty} \int_{A_k} |f| \,\mathrm{d}\mu < \infty$ *ist, und in diesem Fall gilt*

$$\lim_{k\to\infty} \int_{A_k} f \,\mathrm{d}\mu = \int_A f \,\mathrm{d}\mu.$$

Satz A.23 (parameterabhängige Integrale)**.** *Für den Maßraum* $(\Omega, \mathcal{F}, \mu)$ *und die offene Menge* $U \subset \mathbb{R}^p$ *sei* $f : \Omega \times U \to \mathbb{R}$ *gegeben. Für* $\theta \in U$ *definieren wir* $f_\theta : \Omega \to \mathbb{R}$ *durch* $f_\theta(\omega) := f(\omega, \theta)$ *und nehmen* $f_\theta \in \mathcal{L}^1(\mu)$ *für alle* $\theta \in U$ *an. Dann können wir* $g : U \to \mathbb{R}$ *durch*

$$g(\theta) := \int_\Omega f_\theta \,\mathrm{d}\mu$$

definieren. Wir nehmen ferner an, es existiere ein $h \in L^1(\mu)$ *mit* $|f(\omega, \theta)| \leq h(\omega)$ *für alle* $\theta \in U$ *und fast alle* $\omega \in \Omega$*. Dann gilt:*

(i) Sind die Funktionen $\theta \mapsto f(\omega, \theta)$ *(von* U *nach* $\mathbb{R}$*) für alle* $\omega \in \Omega$ *stetig bei* $\theta^* \in U$*, so ist* g *stetig bei* θ^**.*

(ii) Sind die Funktionen $\theta \mapsto f(\omega, \theta)$ *für jedes* $\omega \in \Omega$ *stetig differenzierbar und existiert ein* $k \in L^1(\mu)$ *mit* $|\partial_{\theta_i} f(\omega, \theta)| \leq k(\omega)$ *für alle* $\theta \in U$*, fast alle* $\omega \in \Omega$ *und alle* $i = 1, \ldots, p$*, so ist auch* g *stetig differenzierbar auf* U *und:*

$$\partial_{\theta_j} g(\theta) = \int_\Omega \partial_{\theta_j} f(\cdot, \theta) \,\mathrm{d}\mu \qquad \text{für alle } \theta \in U, \; j = 1, \ldots, p.$$