Universität des Saarlandes
Fachrichtung 6.1 — Mathematik

Grundlagen der Variationsrechnung

Eine anwendungsorientierte Einführung
in die lineare Funktionalanalysis

nach einer Vorlesung von

Dr. Dominic Breit

Wintersemester 2009/2010.

Einleitung

Viele Fragestellungen aus Physik, Technik oder den Wirtschaftswissenschaften sowie innermathematischer Disziplinen (Geometrie, partielle Differentialgleichungen, Variationsrechnung etc.) führen auf unendlichdimensionale Extremwertaufgaben, bei denen es darum geht, in einer Klasse von möglichen „Zuständen“ jenen mit minimaler „Energie“ zu bestimmen, wobei die „Energie“ durch ein Funktional repräsentiert wird. [a] Die „Zustände“ werden in der Regel durch Funktionen einer oder mehrerer Veränderlicher repräsentiert, und das Funktional ist ein Integral, das in einer durch das Problem gegebenen Weise Ableitungen bis zu einer gewissen Ordnung dieser Funktionen involviert. Da man in unendlichdimensionalen Räumen („Funktionenräumen“) arbeitet, ist nicht klar, ob und unter welchen „natürlichen“ Bedingungen solche unendlichdimensionalen Extremwertaufgaben lösbar sind. Andererseits können auch partielle Differentialgleichungen (speziell nichlineare) oft nicht mehr ohne tieferliegende funktionalanalytische Hilfsmittel gelöst werden.

Wie sich herausstellt, können derartige Probeme häufig auch nicht mehr in Funktionenräumen klassisch differenzierbarer Funktionen gelöst werden. Tatsächlich gibt es oft auch nur singuläre Lösungen, was beispielsweise in der Materialtheorie durch Loch– bzw. Rissbildung beschrieben wird. Bei Problemen mit geometrischem Hintergrund gibt es dagegen oft topologische Gründe, die gegen die Existenz einer überall differenzierbaren Lösung sprechen.

Unser Ziel ist es, die „natürlichen“ Funktionenräume bereitzustellen, in denen die Existenz verallgemeinerter (distributioneller) Lösungen für vieler solcher Problemstellungen unter gewissen Bedingungen unschwer bewiesen werden kann. Unsere Überlegungen werden uns dabei in die Theorie der Sobolev–Räume führen.

Die Frage, ob bzw. unter welchen Bedingungen die so gewonnenen verallgemeinerten Lösungen — welche a priori noch nicht einmal stetig sind — tatsächlich bessere Eigenschaften haben, oder sogar klassische Lösungen produzieren (sog. „Regularitätstheorie“), ist ungleich aufwendiger und kann im Rahmen dieser Vorlesung nicht angegangen werden.

[a] In der Physik hat man z. B. „Energien“ wie Arbeit, elektrisches Potential etc., während in den Wirtschaftswissenschaften die „Energie“ ein Kostenfunktional ist.

Warnung.
Dieses Skript dient als ergänzendes Begleitmaterial zur Vorlesung. Es kann und soll den Besuch sowie eine Mitschrift der Vorlesung nicht ersetzen, und erhebt keinen Anspruch auf Fehlerfreiheit oder Vollständigkeit.

Inhaltsverzeichnis

§ 1. Einführung und Beispiele

In der allgemeinsten Form liest sich ein Variationsproblem als

$$J[u] := \int_\Omega F(x, u(x), \nabla u(x), \nabla^2 u(x))\, dx \longrightarrow \min$$
$$u = \varphi \quad \text{auf } \partial\Omega$$

mit $\Omega \subset \mathbb{R}^n$ und einer gegebenen Randfunktion $\phi : \Omega \to \mathbb{R}$ (∇u bezeichnet den Gradienten und $\nabla^2 u$ die Hesse-Matrix). Gesucht ist also eine Funktion $u^* : \Omega \to \mathbb{R}$ mit $u^*(x) = \varphi(x)$ für alle $x \in \partial\Omega$ und

$$J[u^*] \leq J[v]$$

für alle $v : \Omega \to \mathbb{R}$ mit $v(x) = \varphi(x)$ für alle $x \in \partial\Omega$. Auch Probleme mit vektorwertigen Funktionen $u : \Omega \to \mathbb{R}^N$ ($N \geq 2$) sind weit verbreitet (und natürlich sind auch Funktionale, die Ableitungen höherer Ordnung enthalten denkbar). Am häufigsten begegnet man Minimierungsaufgaben der Form

$$J[u] := \int_\Omega F(\nabla u)\, dx \longrightarrow \min .$$

Beispiele aus Physik und Ingenieurwissenschaften
Ein Volumen $\Omega \subset \mathbb{R}^3$ wird von einer Newtonschen Flüssigkeit durchflossen (dazu gehören z.B. Wasser und Öle sowie die meisten Flüssigkeiten des Alltagsgebrauchs). Sei $u^* : \Omega \to \mathbb{R}^3$ das Geschwindigkeitsfeld der Teilchenbewegung und $f : \Omega \to \mathbb{R}^3$ die einwirkende Kraft (Gravitation, elektrisches Feld). Dann ist u^* Minimierer des Funktionals ($p \in (1, \infty)$)

$$J[u] = \int_\Omega \left\{ \left(1 + |\epsilon(u)|^2\right)^{\frac{p}{2}} - f \cdot u \right\} dx$$
$$\text{u.d.N. } \operatorname{div} u = 0 \text{ auf } \Omega$$

zu gegebenen Randwerten $\varphi : \partial\Omega \to \mathbb{R}^3$. Dies enspricht einer verallgemeinerung des klassischen Stokes-Problems ($p = 2$). Der Exponent p sowie die Randfunktion φ können experimentell bestimmt werden. Zur Notation:

$$|P| := (P:P)^{\frac{1}{2}}, \quad P:Q := \operatorname{tr}(P^TQ) = \sum_i^n \sum_j^N P_{ij}Q_{ij} \quad \text{für } P,Q \in R^{N\times n}$$

$$\epsilon(u) = \frac{1}{2}\left(\nabla u + \nabla u^T\right), \quad \operatorname{div} u := \sum_{i=1}^n \frac{\partial u_i}{\partial x_i}.$$

Sei $\Omega \subset \mathbb{R}^n$ $(n \in \{2,3\})$ ein elastisch-plastischer Körper (d.h. nach Deformation kehrt er nur sehr langsam in den Ursprungszustand zurück). Sei $x \in \Omega$ ein Punkt vor der Deformation und $\overline{x}$ seine Position nach Deformation. Die Verzerrung wird dann durch

$$u^*(x) := \overline{x} - x$$

beschrieben. Dann ist u^* Minimierer von

$$\int_\Omega \left\{|\epsilon^D(u)| \ln(1 + |\epsilon^D(u)|) + (\operatorname{div} u)^2\right\} dx$$
$$\epsilon^D(u) := \epsilon(u) - \frac{1}{n}(\operatorname{div} u)I$$

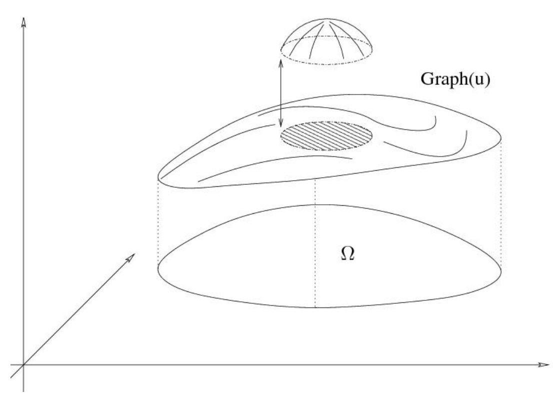
zu gegebenen Randdaten φ.

Geometrie

Minimalflächenproblem: Finde eine Funktion $u^* : \Omega \to \mathbb{R}$ $(\Omega \subset \mathbb{R}^2)$ so dass

$$J[u] = \text{Fläche des Graphen von } u := \int_\Omega \sqrt{1 + |\nabla u|^2}\, dx$$

minimal wird. Dabei soll $u = \varphi$ auf $\partial\Omega$ gelten. Es lässt sich zeigen, dass die



Minimierung des obigen Funktionals äquivalent zur nichtlinearen partiellen Differentialgleichung

$$\operatorname{div}\left(\frac{\nabla u}{\sqrt{1+|\nabla u|^2}}\right) = 0 \text{ auf } \Omega$$

ist.

Wirtschaftswissenschaften

Eine Option (hier der einfachste Fall: europäische Call-Option) ist das Recht eine Aktie an einem ZUKÜNFTIGEN Zeitpunkt ($t = T$) zu einem HEUTE ($t = 0$) festgelegten Preis E zu kaufen. Ist der Wert S der Aktie zum Zeitpunkt T größer als E machen wir einen Gewinn von $S - E$. Die Black-Scholes Theorie (die mit dem Nobelpreis für Wirtschaftswissenschaften ausgezeichnet wurde) beschäftigt sich mit dem Preis $C = C(S,t)$ für dieses Recht auf einem gleichgewichtigen Markt. Hierbei gilt

$$\begin{aligned}\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + rS\frac{\partial C}{\partial S} - rC = 0 \quad \text{auf } [0,\infty)\times[0,T]\\ C(S,T) = \max\{S-E, 0\}, \quad C(0,t) = 0,\\ C(S,t) \approx S \text{ für } S \to \infty.\end{aligned}$$

Hierbei ist $r > 0$ der risikolose Zins. Für den Aktienpreis $S = S(t)$ wird eine Brownsche Bewegung angenommen. Dies ist ein stochastischer Prozess bei dem Aktienrenditen normalverteit sind mit Varianz σ. Obige Gleichung lässt sich (im Gegensatz zu fast allen Variationsproblemen und PDEs) analytsich lösen und erfreut sich daher in der Praxis großer Beliebtheit obwohl die Annahmen des Modells (normalverteilte Renditen) bereits Mitte des letzten Jahrhunderts empirisch wiederlegt worden.

Digitale Bildverarbeitung

Wir betrachten ein gestörtes Schwarz-Weiß-Bild, dass durch eine Funktion $f : \mathbb{R}^2 \supset \Omega \to \mathbb{R}$ beschrieben wird ($\Omega = [0,a]\times[0,b]$). Dabei gibt der Funktionswert $f(x)$ an wie dunkel der Grauton im Bildpunkt x ist (je heller desto größer der Wert). Um die Störung möglichst gut zu beseitigen minimiert man das Funktional ($\alpha > 0$)

$$E_f[u] := \int_\Omega \left\{(u-f)^2 + \alpha|\nabla u|^2\right\} dx.$$

Dabei steht $(f-u)^2$ für den Abstand zum ursprünglichen Bild und $|\nabla u|$ ist ein Maß für die Glattheit des entrauschten Bildes.

Bisher haben wir ausgeblendet in welcher Klasse von Funktionen die Variationsproblemegelöst werden. Dieser Frage ist der größte Teil der Vorlesung gewidmet. Beispiel: Wir suchen Minimerer $u^* : \Omega \to \mathbb{R}$ von

$$\int_\Omega |\nabla u|^2 \, dx$$

zu Randdaten $\varphi : \Omega \to \mathbb{R}$ in einer Klasse $\mathcal{C}$: d.h.

$$\int_\Omega |\nabla u^*|^2 \, dx \leq \int_\Omega |\nabla u|^2 \, dx$$

für alle $u \in \mathcal{C}$ mit $u = \varphi$ auf $\partial\Omega$. Man hat die Vorlesttung das $\mathcal{C} = C^1(\Omega)$ sein sollte. Die Existenz eines C^1-Minimums lässt sich jedoch nicht direkt zeigen. Haben wir eine stetige Funktion $f : K \to [0, \infty)$ und eine kompakte Menge $K \subset \mathbb{R}^n$ so lässt sich einfach die Existenz eines Minimums von f auf K zeigen: Wir betrachten eine Minimialfolge $(x_k)_{k \in \mathbb{N}} \subset K$ [a], d.h.

$$f(x_k) \overset{k \to \infty}{\longrightarrow} \inf_{y \in K} f(y).$$

Eine solche Minimalfolge existiert nach Definiton des Infimums immer. Da K kompakt ist, können wir eine konvergente Teilfolge $(\widetilde{x}_k)_{k \in \mathbb{N}} \subset (x_k)_{k \in \mathbb{N}}$ wählen mit Limes $x \in K$ (hierzu benötigt man Bolzano-Weierstraß). Es folgt

$$f(x) = \lim_{k \to \infty} f(\widetilde{x}_k) = \lim_{k \to \infty} f(x_k) = \inf_{y \in K} f(y)$$

und damit die Minimalität von x. Um unser Argument auf die Variationsrechnung auszudehnen müssen wir $C^1(\Omega)$ mit einer Norm versehen. Die Wahl

$$\|u\| := \|u\|_\infty + \|\nabla u\|_\infty$$

führt auf einen vollständigen Raum. Sie hat jedoch nichts mit obigem Funktional zu tun und ist daher unpassend. Ein zweiter Ansatz ist

$$\|u\| := \|u\|_{L^2} + \|\nabla u\|_{L^2}$$
$$\|f\|_{L^2} := \left[\int_\Omega |f|^2 \, dx \right]^{\frac{1}{2}}.$$

[a] Wir fassen eine Folge (auch) als Teilmenge $\{x_1, x_2, x_3, \ldots\}$ von X auf, daher die Notation $(x_n) \subset X$. (In der Literatur findet man auch die Schreibweise $\{x_n\} \subset X$.) Natürlich ist (x_n) per Definition eine Abbildung von $\mathbb{N} \to X$.

Wir wählen eine Minimalfolge $(u_k) \subset C^1(\Omega)$ mit $u_k = \varphi$ auf $\partial\Omega$, d.h.

$$\int_\Omega |\nabla u_k|^2 \, dx \overset{k\to\infty}{\longrightarrow} \inf_{u \in C^1(\Omega)} \int_\Omega |\nabla u|^2 \, dx.$$

Damit gilt

$$\sup_{k\in\mathbb{N}} \|\nabla u_k\|_{L^2} < \infty.$$

Für Funktionen mit Nullrandwerten gilt die Poincaré-Ungleichung (auf einen Beweis wird an dieser Stelle verzichtet)

$$\|f\|_{L^2} \leq c \, \|\nabla f\|_{L^2} \, .$$

Wir erhalten

$$\begin{aligned}
\sup_{k\in\mathbb{N}} \|u_k\|_{L^2} &\leq \sup_{k\in\mathbb{N}} \|u_k - \varphi\|_{L^2} + \sup_{k\in\mathbb{N}} \|\varphi\|_{L^2} \\
&\leq \sup_{k\in\mathbb{N}} c \, \|\nabla(u_k - \varphi)\|_{L^2} + \sup_{k\in\mathbb{N}} \|\varphi\|_{L^2} \\
&\leq c \sup_{k\in\mathbb{N}} \|\nabla u_k\|_{L^2} + c \, \|\nabla\varphi\|_{L^2} + \|\varphi\|_{L^2} < \infty.
\end{aligned}$$

Die Folge (u_k) ist also im normierten Raum $(C^1(\Omega), \|\cdot\|)$ beschränkt. Der Satz von Bolzano Weierstraß greift jedoch in unendlichdimensionalen Räumen nicht (werden wir später beweisen), so dass wir trotz Beschränktheit keine konvergente Teilfolge wählen können. Im Übrigen gibt es Beispiele von Variationsproblemen, die kein C^1-Minimum besitzen, so dass eine solche Konvergenz i.A. sogar generell ausgeschlossen ist.

Um dieses Problem zu lösen werden wir

- Verschieden Funktionenräume studieren, insbesondere dem Raum

  $$L^p(\Omega) := \left\{ f : \Omega \to \mathbb{R} : \int_\Omega |f|^p \, dx < \infty \right\}$$

  mit $1 \leq p \leq \infty$ kommt eine wichtige Bedeutung zu.

- In unendlichdimensionalen Räumen greift der Satz von Bolzano-Weierstraß nicht, in einer recht großen Klasse (jedoch nicht in $C^1(\Omega)$) gilt aber eine verallgemeinerte Variante (diese Räume heißen reflexiv).

- Wir werden einen passend Raum kennen lernen (Sobolev-Raum), indem wir eine Teilfolge finden können die in einem gewissen Sinn konvergiert (schwache Konvergenz) und so eine Lösung des Variationsproblems erzeugen. Die so gewonnene Lösung hat jedoch analytisch sehr schlechte Eigenschaften (es gibt Beispiele die nirgendwo stetig sind).

Die Aufgabe der Regularitätstheorie (wird in Nachfolgeveranstaltungen behandelt) ist es zu zeigen, dass die schwachen Lösungen (im Sobolev-Raum) bessere Eigenschaften haben, folgende Aussagen sind dabei denkbar:

- Geschwindigkeitsfeld der Fluide: Lösung ist partiell regulär, d.h. es gibt eine „große" Teilmenge Ω_0 von Ω mit $u^* \in C^1(\Omega_0)$. Als Beispiel zur partiellen Regularität betrachte man die Funktion $B_1 \ni x \mapsto x/|x|$, die offenbar überall C^1 ist außer in der Null, d.h. $\Omega_0 = B_1 - \{0\}$. Sie ist Lösung der Gleichung (Übung!)

$$-\Delta u = u|\nabla u|^2 \quad \text{auf } \Omega.$$

  Die Frage ob die Lösung (des Fluid-Problems) sogar in $C^1(\Omega)$ liegt ist bisher ungelöst. Eine bessere Aussage als $C^1(\Omega)$ ist hier nicht zu erwarten.

- Deformationstensor: Im Fall $n = 2$ erhalten wir $u^* \in C^1(\Omega)$, für $n = 3$ nur partielle Regularität. Es ist davon auszugehen, dass sich diese Ergebnisse i.A. nicht verbessern lassen.

- Die Lösung des Minimalflächenproblems gehört zur Klasse $C^\infty(\Omega)$. Das ist das bestmögliche Ergebnis, das jedoch nur in Ausnahmefällen erreicht wird.

- Die Lösung der Black-Scholes Gleichung gehört zur Klasse $C^\infty(\Omega)$. Im Gegensatz zu den anderen Problemen lässt sich hier die Lösung analytisch bestimmen.

Folgende Hilfsmittel werden wir des öfteren benötigen:

- Eulergleichung: Sei u^* Minimierer von

$$J[u] := \int_\Omega |\nabla u|^2 \, dx$$

  in einer Funktionenklasse $\mathcal{C}$. Sei $\varphi \in C_0^\infty(\Omega)$ mit $u^* + t\varphi \in \mathcal{C}$ für $t \in [-\epsilon, \epsilon]$, so hat die Funktion

$$f(t) := J[u + t\varphi]$$

  ein Minimum in $t = 0$. Es folgt

$$0 = f'(0) = \frac{d}{dt}|_{t=0} J[u + t\varphi] = \int_\Omega \frac{\partial}{\partial t}|_{t=0} |\nabla u + t\nabla\varphi|^2 \, dx.$$

Die liefert eine partielle Differentialgleichung.

- Satz von Gauß: Sei $F \in C^1(\overline{\Omega}, \mathbb{R}^n)$ für ein Gaußgebiet $\Omega \subset \mathbb{R}^n$ (d.h. offen, zusammenhängend, beschränkt mit stückweise glattem Rand). Dann gilt

$$\int_\Omega \operatorname{div} F \, d\mathcal{L}^n = \int_{\partial\Omega} F \cdot \mathcal{N} \, d\mathcal{H}^{n-1}.$$

Dabei ist $\mathcal{L}^n$ das n-dimensionale Lebesgue-Maß (=Volumenmaß), $\mathcal{H}^{n-1}$ das $(n-1)$-dimensionale Hausdorff-Maß (=Flächenmaß) und $\mathcal{N}$ die äußere Einheitsnormale an $\partial\Omega$.

- Grundlegene Maßtheorie: Maßintegral, Lebesgue- und Hausdorff-Maß, Satz von Fubini, Transformationssatz.

- Fundamentallemma der Variationsrechnung (werden wir im Abschnitt über Lebesgue-Räume beweisen): Sei v eine über Ω integrierbare Funktion mit

$$\int_\Omega v\varphi \, dx = 0 \quad \forall \varphi \in C_0^\infty(\Omega),$$

so folgt $v \equiv 0$.

§ 2. Klassische Beispiele für Funktionenräume und Grundlegendes

Wir wiederholen zunächst einige grundlegende Tatsachen über normierte Räume. Im Folgenden bezeichne X einen Vektorraum über $\mathbb{R}$, versehem mit einer Norm $\|\cdot\| := \|\cdot\|_X$. Für einen normierten Vektorraum verwenden wir die übliche Notation $\big(X, \|\cdot\|\big)$.
Eine Folge $(x_n) \subset X$ heißt eine *Cauchy–Folge* in X (oder eine Cauchy–Folge bzgl. der Norm $\|\cdot\|$), falls es zu jedem vorgegebenen $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon) \in \mathbb{N}$ gibt, so dass

$$\|x_n - x_m\| < \varepsilon \quad \text{für alle} \quad n, m > n_0$$

ausfällt. Wir schreiben dafür auch kürzer $\|x_n - x_m\| \xrightarrow{n,m} 0$. Wir sagen auch: $\|x_n - x_m\|$ veschwindet für $n, m \gg 1$.
Eine Folge $(x_n) \subset X$ heißt *konvergent* in X, falls ein $x \in X$ existiert mit $\|x_n - x\| \xrightarrow{n} 0$. Wir schreiben dafür üblicherweise $x_n \xrightarrow{n} x$ in X. Bekanntermaßen ist jede in X konvergente Folge auch eine Cauchy-Folge in X; die Umkehrung ist jedoch i. a. nicht richtig.

Definition 2.1 (Banach–Raum)
Sei $\big(X, \|\cdot\|\big)$ ein normierter Raum. Dann heißt X vollständig bzgl. der Norm $\|\cdot\|$, oder ein Banach–Raum, falls jede Cauchy–Folge bzgl. $\|\cdot\|$ in X konvergiert.

Wir erinnern an den Begriff der äquivalenten Normen: Seien zwei Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ auf dem Raum X gegeben. Dann heißen die beiden Normen zueinander *äquivalent*, falls es positive Konstanten c und C gibt, so dass

$$c\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1 \quad \text{für alle} \quad x \in X.$$

Beispielsweise sind bekanntlich im Raum $X = \mathbb{R}^d$ $(d \in \mathbb{N})$ alle Normen zueinander äquivalent, weshalb man sich eine zur Anwendung „passende“ Norm aussuchen kann. In allgemeinen normierten Räumen ist dies natürlich i. a. nicht der Fall (vgl. nachfolgendes Beispiel)).

Bemerkung 2.2

i) Ist X ein endlich–dimensionaler Raum, so sind alle Normen auf X äquivalent, und daher X bzgl. jeder Norm vollständig.

ii) In unendlich–dimensionalen Räumen ist die Aussage aus i) i. a. falsch, wie das folgende Beispiel zeigt.

Beispiel.
Sei $X := C^0[0,1]$ der Raum der auf dem kompakten Intervall $[0,1]$ stetigen, reellwertigen Funktionen. Dann werden durch

$$\|x\|_\infty := \max_{t\in[0,1]} \big|x(t)\big|, \quad \|x\|_1 := \int_0^1 \big|x(t)\big|\,dt$$

zwei Normen auf X erklärt, welche nicht äquivalent zueinander sind. Tatsächlich ist $\big(X, \|\cdot\|_\infty\big)$ vollständig, $\big(X, \|\cdot\|_1\big)$ dagegen nicht (vgl. Aufgabe 3).
Beweis: Nehmen wir an $\big(X, \|\cdot\|_\infty\big)$ und $\big(X, \|\cdot\|_1\big)$ wären äquivalent. Dann existieren positive Konstanten c, C mit

$$c\|x\|_1 \le \|x\|_\infty \le C\|x\|_1 \quad \forall x \in X.$$

(Tatsächlich gilt die erste Ungleichung mit $c = 1$, während die zweite falsch ist). Damit sind die Konvergenzen

$$\begin{aligned} \|x_n - x_m\|_1 &\xrightarrow{n,m} 0 \\ \|x_n - x_m\|_\infty &\xrightarrow{n,m} 0 \end{aligned}$$

äquivalent, also auch die Vollständigkeit beider Räume, ein Widerspruch.

Einleitend behandeln wir nun drei klassische Typen von Funktionenräumen: Räume beschränkter, stetiger und differenzierbarer Funktionen.

A. Räume beschränkter Funktionen

Sei $X \neq \emptyset$ eine beliebige Menge (nicht notwendig enthalten in einem metrischen oder normierten Raum) und sei $\big(Y, \|\cdot\|\big)$ ein normierter Raum. Wir betrachten die Menge

$$\mathcal{B}(X,Y) := \Big\{ u : X \to Y;\ \sup_{x\in X} \big\|u(x)\big\| < \infty \Big\}$$

der beschränkten Funktionen von $X \to Y$. Dann wird $\mathcal{B}(X,Y)$ durch

$$\|u\|_\infty := \|u\|_{\infty;\,X} := \sup_{x\in X} \big\|u(x)\big\|$$

normiert, und $\mathcal{B}(X,Y)$ ist bzgl. dieser Norm ein Banach–Raum, sofern auch $\big(Y, \|\cdot\|\big)$ ein Banach–Raum ist. Ist also Y kein Banach–Raum, so braucht auch $\mathcal{B}(X,Y)$ keiner zu sein. Die Norm $\|\cdot\|_\infty$ wird *Supremumnorm* genannt.

Beweis.
Sei (u_n) eine Cauchy–Folge in $\mathcal{B}(X,Y)$ und sei $\varepsilon > 0$ beliebig. Dann ist für jedes $x \in X$

$$\big\|u_n(x) - u_m(x)\big\| \le \|u_n - u_m\|_\infty < \varepsilon \quad \text{für alle} \quad n, m \gg 1,$$

d. h. $\big(u_n(x)\big)$ ist eine Cauchy–Folge in $\big(Y, \|\cdot\|\big)$. Daher existiert der punktweise Limes $u(x) := \lim_n u_n(x)$ für alle $x \in X$. Bleibt $u \in \mathcal{B}(X,Y)$ nachzuweisen. Sei dazu wieder $\varepsilon > 0$ beliebig und $n_0 = n_0(\varepsilon) \in \mathbb{N}$ so gewählt, dass

$$\|u_n - u_m\|_\infty < \frac{\varepsilon}{2} \quad \text{für alle} \quad n, m > n_0$$

ist. Dann wird für jedes $x \in X$ und alle $n > n_0$

$$\big\|u_n(x) - u(x)\big\| \le \big\|u_n(x) - u_m(x)\big\| + \big\|u_m(x) - u(x)\big\| < \varepsilon,$$

falls $m > n_0$ (in Abhängigkeit von x) so groß gewählt wird, dass

$$\big\|u_m(x) - u(x)\big\| < \frac{\varepsilon}{2}$$

wird. Demzufolge ist $u - u_n \in \mathcal{B}(X,Y)$ mit $\|u - u_n\|_\infty < \varepsilon$ für alle $n > n_0$, insbesondere also $u \in \mathcal{B}(X,Y)$ und $u_m \xrightarrow{m} u$ bzgl. $\|\cdot\|_\infty$, also gleichmäßige Konvergenz. $\square$

B. Räume stetiger Funktionen

Sei $U \neq \emptyset$ Teilmenge eines normierten Raumes X, und sei $\big(Y, \|\cdot\|\big)$ ein Banach–Raum. Wir betrachten den Raum

$$C^0_{\mathcal{B}}(U,Y) := \big\{u \in \mathcal{B}(U,Y);\ u \text{ stetig}\big\} = \mathcal{B}(U,Y) \cap C^0(U,Y)$$

der beschränkten und stetigen Funktionen von $U \to Y$. [a]
Aus dem folgende Lemma ergibt sich zusammen mit A., dass $C^0_{\mathcal{B}}(U,Y)$ ebenfalls ein Banachraum ist.

[a] Man beachte, dass Stetigkeit keine Beschränkheit impliziert. Dies ist i. a. nur dann der Fall, wenn man eine stetige Funktion auf einer kompakten Menge vorliegen hat (Satz von der Annahme von Minimum und Maximum). Allgemein ist also $C^0_{\mathcal{B}}(U,Y) \subsetneqq C^0(U,Y)$ und $C^0_{\mathcal{B}}(U,Y) = C^0(U,Y)$ für kompaktes U.

Lemma 2.2 *a)* $C^0_{\mathcal{B}}(U,Y)$ *versehen mit der Norm von* $\mathcal{B}(U,Y)$ *(also* $\|\cdot\|_\infty$*) ist ein abgeschlossener Unterraum des Banach–Raums* $\mathcal{B}(U,Y)$*.*

b) Abgeschlossene Unterräume von Banachräumen sind wieder Banachräume.

Beweis.
a) Zu zeigen ist folgende Aussage: Sei $(u_n) \subset C^0_{\mathcal{B}}(U,Y)$ eine beliebige Folge mit $u_n \to u \in \mathcal{B}(U,Y)$ bzgl. $\|\cdot\|_\infty$. Dann folgt $u \in C^0_{\mathcal{B}}(U,Y)$. (Das folgt aus dem Prinzip, dass gleichmäßige Limiten stetiger Funktionen stetig sind.)
Dazu seien $\varepsilon > 0$ und $\xi \in U$ beliebig vorgegeben. Es existiert dann ein $n_0 = n_0(\varepsilon) \in \mathbb{N}$ mit

$$\|u_n - u\|_\infty < \frac{\varepsilon}{3} \quad \text{für alle} \quad n > n_0.$$

Insbesondere ist u_{n_0} stetig in ξ, d. h. es gibt ein $\delta = \delta(\varepsilon, n_0) > 0$ derart, dass

$$\big\|u_n(x) - u_{n_0}(x)\big\| < \frac{\varepsilon}{3} \quad \text{für alle } x \in U \text{ mit } \|x - \xi\| < \delta$$

ausfällt. Für diese x folgt

$$\begin{aligned}\big\|u(x) - u(\xi)\big\| &\le \big\|u(x) - u_{n_0}(x)\big\| + \big\|u_{n_0}(x) - u_{n_0}(\xi)\big\| + \big\|u_{n_0}(\xi) - u(\xi)\big\| \\ &\le 2\|u_{n_0} - u\|_\infty + \big\|u_{n_0}(x) - u_{n_0}(\xi)\big\| < \frac{2\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,\end{aligned}$$

und damit die Behauptung. □
b) Sei $(U, \|\cdot\|)$ ein Banachraum und $W \subset U$ ein abgeschlossener Unterraum. Eine Cauchy-Folge $(w_n) \subset (W, \|\cdot\|)$ ist auch eine Cauchy-Folge in $(U, \|\cdot\|)$, die aufgrund der Vollständigekit von $(U, \|\cdot\|)$ einen Limes $w \in U$ besitzt. Wegen der Abgeschlossenheit von W muss aber $w \in W$ gelten, woraus die Vollständigkeit von $(W, \|\cdot\|)$ folgt. □

Bemerkung.
Die Menge U muss nicht notwendig Teilmenge eines normierten Raumes sein, sondern kann auch Teilmenge eines metrischen oder topologischen Raumes sein. Man muss dabei nur wissen, was Stetigkeit in einem metrischen bzw. topologischen Raum bedeutet.

Ist $\Omega \subset \mathbb{R}^d$ ($d \in \mathbb{N}$, $d \ge 2$) offen, so ist der Raum $C^0(\Omega) = C^0(\Omega, \mathbb{R})$ der auf Ω stetigen, reellwertigen Funktionen kein Banach–Raum bzgl. der Supremumnorm $\|\cdot\|_\infty$, weil u als stetige Funktion auf der offenen Menge Ω nicht notwendig beschränkt sein muss — was auch nicht der Fall sein muss, wenn Ω zusätzlich beschränkt ist. Als Spezialfall der obigen Überlegungen ergibt sich aber, dass der Raum

$$C^0(\overline{\Omega}) := \big\{u : \overline{\Omega} \to \mathbb{R};\ u \text{ stetig}\big\},\ ^{\mathsf{b}}$$

versehen mit der *Supremumnorm*

$$\|u\|_\infty := \|u\|_{\infty;\,\Omega} := \sup_{x\in\Omega} \big|u(x)\big|$$

ein Banach–Raum ist. Entsprechend erhält man Banach–Räume differenzierbarer Funktionen.

C. Räume differenzierbarer Funktionen

Sei $\Omega \subset \mathbb{R}^d$ offen und sei $k \in \mathbb{N}_0$. Es bezeichne wie üblich $C^k(\Omega) = C^k(\Omega, \mathbb{R})$ den Raum der k–mal auf Ω stetig differenzierbaren, reellwertigen Funktionen. Man definiert

$$C^k(\overline{\Omega}) := \left\{u \in C^k(\Omega);\ \partial^\gamma u := \frac{\partial^{|\gamma|} u}{\partial x_1^{\gamma_1} \cdots \partial x_d^{\gamma_d}} \in C^0(\overline{\Omega}) \text{ für alle } |\gamma| \le k\right\},\ ^{\mathsf{c}}$$

wobei $\gamma := (\gamma_1, \dots, \gamma_d) \in \mathbb{N}_0^d$ und $|\gamma| := \gamma_1 + \cdots + \gamma_d$ ist. Dann wird durch

$$\|u\|_{C^k} := \|u\|_{C^k(\Omega)} := \sum_{|\gamma|\le k} \|\partial^\gamma u\|_\infty \tag{2.1}$$

eine Norm sowohl auf $C^k(\overline{\Omega})$ erklärt, und der Raum ist bzgl. dieser Norm ein Banach–Raum (vgl. Aufgabe 2). Durch (2.1) wird natürlich auch eine Norm auf $C^k(\Omega)$ definiert, wenn man in der Definition $\|\cdot\|_\infty$ als Supremum und nicht als Maximum auffasst. Allerdings sind diese Räume bzgl. dieser Norm keine Banach–Räume (dies war ja schon für $C^0(\Omega)$ nicht der Fall, wie wir oben gesehen haben).

[b] Stetigkeit von u auf $\overline{\Omega}$ bedeutet, dass u auf Ω stetig ist und eine stetige Fortsetzung nach $\overline{\Omega}$ besitzt, was bedeutet, dass für jeden Randpunkt $\xi \in \partial\Omega$ der Grenzwert

$$\lim_{\Omega\ni x\to\xi} u(x)$$

existiert. Dies ist insbesondere dann der Fall, wenn u auf einer etwas größeren Menge $\widetilde{\Omega} \supsetneqq \Omega$ stetig ist. Allgemeiner besitzt jede beschränkte und gleichmäßig stetige Funktion $u : \Omega \to \mathbb{R}$ eine eindeutig bestimmte ebenfalls beschränkte und stetige Fortsetzung nach $\overline{\Omega}$.

[c] Die Definition besagt, dass eine solche Funktion u also k–mal stetig differenzierbar in Ω ist, und u sowie sämtliche Ableitungen bis zur Ordnung k von u stetig nach $\overline{\Omega}$ fortgesetzt werden können.

D. Funktionen mit kompaktem Träger

Eine für die Variationsrechnung bedeutende Teilklasse von $C^k_{\mathcal{B}}$, sind die Funktionen mit kompaktem Träger. Für $\Omega \subset \mathbb{R}^d$ offen und $k \in \mathbb{N}_0 \cup \{\infty\}$ ist

$$C^k_\circ(\Omega) := \left\{u \in C^k(\Omega);\ \operatorname{spt} u \Subset \Omega\right\}$$

der Raum der k–mal auf Ω stetig differenzierbaren Funktionen mit *kompaktem Träger* in Ω, d. h.

$$\operatorname{spt} u := \overline{\left\{x \in \Omega;\ u(x) \neq 0\right\}}$$

ist kompakte Teilmenge von Ω. [d] Man überlegt sich leicht, dass jede (wenigstens) stetige Funktion mit kompaktem Träger beschränkt sein muss, so dass also $C^k_\circ(\Omega) \subset C^k_{\mathcal{B}}(\Omega)$ ist.

Besondere Bedeutung kommt der Klasse $C^\infty_\circ$ zu, wie wir später noch sehen werden. Die Funktionen dieser Klasse bezeichnet man auch gerne als *Testfunktionen*. Ein Beispiel liefert die Funktion $\eta : \mathbb{R}^d \to \mathbb{R}$,

$$\eta(x) := \begin{cases} \exp\left(\frac{1}{|x|^2-1}\right) & ;\quad |x| < 1 \\ 0 & ;\quad \text{sonst,} \end{cases} \tag{2.2}$$

welche Träger in der abgeschlossenen Einheitskugel $\overline{B}_1(0)$ hat).

Schlußbemerkung.

Natürlich kann man die oben eingeführten Räume $C^k(\Omega)$, $C^k(\overline{\Omega})$, $C^k_{\mathcal{B}}(\Omega)$ sowie $C^k_\circ(\Omega)$ ($\Omega \subset \mathbb{R}^d$ offen und $k \in \mathbb{N}_0$) auch für vektorwertige Funktionen $u : \Omega \to \mathbb{R}^D$ ($D \in \mathbb{N}$) erklären und erhält die entsprechenden Aussagen. Man schreibt $C^k(\Omega, \mathbb{R}^D)$, $C^k(\overline{\Omega}, \mathbb{R}^D)$ etc. für die vektoriellen Räume, oder auch $C^k(\Omega)^D$, $C^k(\overline{\Omega})^D$ etc. [e]

---

[d] Man beachte, dass $\operatorname{spt} u$ für unbeschränktes Ω allgemein nicht kompakt, sondern lediglich abgeschlossen ist. Die Symbolik $A \Subset B$ bedeutet, dass der Abschluss $\overline{A}$ von A eine kompakte Teilmenge von B ist. Man sagt A ist *kompakt enthalten* in B.

[e] In der Literatur findet man für die Räume $C^k_\circ(\Omega)$ bzw. $C^k_\circ(\Omega)^D$ auch die Bezeichnungen $C^k_c(\Omega)$, $\mathcal{D}(\Omega)$ bzw. $C^k_c(\Omega)^D$, $\mathcal{D}(\Omega)^D$ u. ä.

§ 3. Lebesgue–Räume

Wir vereinbaren zunächst einige Sprechweisen, welche aus der elemtaren Maßtheorie bekannt sind.
Sei $X \neq \emptyset$ eine beliebige Menge und sei $\mu : \wp(X) \to [0, \infty]$ ein Maß über X (wobei $\wp(X)$ wie üblich die Potenzmenge von X bezeichnet), d. h. μ hat die Eigenschaft:

$$\mu(A) \leq \sum_{n \in \mathbb{N}} \mu(A_n)$$

für alle $A, A_n \in \wp(X)$ mit $A \subset \bigcup_{n \in \mathbb{N}} A_n$. Nach *Carathéodory* heißt eine Menge $A \in \wp(X)$ *μ–messbar*, falls für jedes $B \in \wp(X)$ gilt:

$$\mu(B) = \mu(B \cap A) + \mu(B \setminus A).$$

Eine Funktion $u : X \to \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ heißt *μ–messbar*, falls das Urbild $u^{-1}(I)$ eines jeden Intervalls $I \subset \overline{\mathbb{R}}$ eine *μ–messbare* Menge ist. [a]
Sei E eine Eigenschaft von Funktionen. Wir sagen, u habe *μ–fast–überall* (kurz: μ–f. ü.) auf X die Eigenschaft E, falls die Menge

$$N := \big\{ x \in X;\ u(x) \text{ erfüllt nicht } E \big\}$$

eine μ–Nullmenge, also $\mu(N) = 0$ ist. Wir sagen auch, u habe in *μ–fast–allen* (kurz: μ–f. a.) Punkten $x \in X$ die Eigenschaft E.
Zwei Funktionen $u, w : X \to \overline{\mathbb{R}}$ sind demnach μ–f. ü. identisch auf X, falls

$$\mu\big(\big\{ x \in X;\ u(x) \neq w(x) \big\}\big) = 0$$

ist. Beispielsweise ist die charakteristische Funktion $\mathbb{1}_{\mathbb{Q}}$ von $\mathbb{Q}$ bzgl. dem eindimensionalen Lebesgue–Maß $\mathcal{L}^1$ f. ü. auf $\mathbb{R}$ identisch der Nullfunktion.

In diesem § wird uns die Frage beschäftigen, wie man integrierbare Funktionen $X \to \overline{\mathbb{R}}$ zu einem normierten Raum — welcher sinnvollerweise ein Banach–Raum sein sollte — zusammenfassen kann.

[a] Ein (verallgemeinertes) Intervall in $\overline{\mathbb{R}}$ ist ein Intervall, bei dem auch die unenlich fernen Punkte $\pm\infty$ als Grenzen zugelassen sind (z. B. $(0, \infty]$, $[-\infty, \infty] = \overline{\mathbb{R}}$).

Wir machen zunächst die vorläufige Definition: 1^{ter} **Versuch**

$$\mathscr{L}^1(X;\mu) := \Big\{ u : X \to \overline{\mathbb{R}};\ u\ \mu\text{–messbar mit } \int_X \big|u(x)\big|\, d\mu(x) < \infty \Big\}.$$

Wegen $\int_X \big|u(x)\big|\, d\mu(x) < \infty$ ist $\mu\big(\{x \in X;\ u(x) = \pm\infty\}\big) = 0$, d. h.

$$\|u\|_1 := \int_X \big|u(x)\big|\, d\mu(x)$$

ist eine wohldefinierte Größe.

Bemerkung 3.1

i) Man will Werte in $\overline{\mathbb{R}}$ zulassen um auch Funktionen mit Singularitäten zu betrachten. Beispielsweise überzeugt man sich leicht davon, dass die Abbildung $x \mapsto 1/\sqrt{|x|}$ zur Klasse $\mathscr{L}^1([-1,1];\mathcal{L}^1)$ gehört.

ii) Auch wenn die Funktion u auf X nur endliche Werte annimmt, muss $\|u\|_1$ nicht existieren. Man betrachte etwa $\mathscr{L}^1((0,1);\mathcal{L}^1)$ und die Abbildung $x \mapsto 1/x$.

Probleme.

i) Sind $u, w \in \mathscr{L}^1(X;\mu)$ und $c \in \mathbb{R}$, so ist $u + cw$ eventuell auf einer μ–Nullmenge ein undefinierter Ausdruck wie z. B. „$\infty - \infty$“. Die Ursache dafür ist, dass wir Werte in $\overline{\mathbb{R}}$ zulassen; $\mathscr{L}^1(X;\mu)$ hat also keine Vektorraum–Struktur.

Deshalb sei von nun an

$$\mathscr{L}^1(X;\mu) := \Big\{ u : X \to \mathbb{R};\ u\ \mu\text{–messbar mit } \int_X \big|u(x)\big|\, d\mu(x) < \infty \Big\}.$$

Dann ist $\mathscr{L}^1(X;\mu)$ ein linearer Raum, und nach den Rechenregeln für μ–messbare Funktionen ist

$$\|cu\|_1 = |c|\|u\|_1 \quad \text{und} \quad \|u+w\|_1 \le \|u\|_1 + \|w\|_1$$

für alle $u, w \in \mathscr{L}^1(X;\mu)$ und $c \in \mathbb{R}$. Folgendes Problem bleibt jedoch.

ii) Ist $u \in \mathscr{L}^1(X;\mu)$ mit $\|u\|_1 = 0$, so ist lediglich $u = 0$ μ–f. ü. auf X, d. h. es kann Punkte $x \in X$ geben mit $u(x) \neq 0$. (Man betrachte beispielsweise wieder die charakteristische Funktion $\mathbb{1}_{\mathbb{Q}}$ und $X := [0,1]$.)

Durch $\|\cdot\|_1$ *wird demnach keine Norm auf* $\mathscr{L}^1(X;\mu)$ *erklärt, sondern eine sog.* Seminorm.

Ist speziell $X := \Omega \subset \mathbb{R}^d$ *offen und* $\mu := \mathcal{L}^d$ *das d–dimensionale Lebesgue-Maß, so kann man die Definition von* $\mathscr{L}^1$ *nochmals modifizieren durch:*

$$\mathscr{L}^1(\Omega) := \mathscr{L}^1(\Omega;\mathcal{L}^d) := \left\{u \in C^0(\Omega);\ \int_\Omega \big|u(x)\big|\,dx < \infty\right\},$$

wobei wir wie üblich abkürzend dx *statt* $d\mathcal{L}^d(x)$ *geschrieben haben. Dann ist zwar* $\mathscr{L}^1(\Omega)$ *ein linearer Raum und* $\|\cdot\|_1$ *eine Norm auf diesem Raum, aber* $\big(\mathscr{L}^1(\Omega), \|\cdot\|_1\big)$ *ist nicht vollständig.*

Zur Begründung betrachten wir das folgende Beispiel:

Sei $\Omega := B_1(0)$ *die offene Einheitskugel im* $\mathbb{R}^d$. *Für* $x \in \Omega$ *und* $n \in \mathbb{N}$ *betrachten wir die Funktionenfolge* (u_n), *welche gegeben wird durch*

$$u_n(x) := \frac{x_1}{\frac{1}{n} + |x|}.$$

Jedes u_n *ist offenbar stetig in* Ω *mit* $|u_n(x)| \le 1$ *für alle* $x \in \Omega$. *Ferner strebt* (u_n) *punktweise auf* Ω *gegen die Funktion* $u : \Omega \to \mathbb{R}$,

$$u(x) := \begin{cases} \frac{x_1}{|x|} & ;\quad x \neq 0 \\ 0 & ;\quad x = 0 \end{cases}. \tag{3.1}$$

Nach dem Satz von Lebesgue über die dominierte Konvergenz gilt daher

$$\int_\Omega \big|u_n(x) - u(x)\big|\,dx = \|u_n - u\|_1 \xrightarrow{n} 0. \tag{3.2}$$

Daraus folgt, dass (u_n) *eine Cauchy–Folge in* $C^0(\Omega)$ *bzgl. der Norm* $\|\cdot\|_1$ *ist (warum?), die aber nicht konvergiert.*

Denn angenommen, es existiert ein $w \in C^0(\Omega)$ *mit* $\|u_n - w\|_1 \xrightarrow{n} 0$. *Wegen* (3.2) *würde dann aber*

$$\int_\Omega \big|u(x) - w(x)\big|\,dx = 0 \quad \Longleftrightarrow \quad w = u \quad \mathcal{L}^d\text{–f. ü. auf } \Omega$$

folgen, was wegen der Stetigkeit von w *in* Ω *und wegen* $u \in C^0(\Omega\setminus\{0\})$ *zu dem Widerspruch* $w = u$ *in* $\Omega\setminus\{0\}$ *führt. (Das hieße, dass* w *eine stetige Fortsetzung von* u *auf ganz* Ω *ist. Der Grenzwert* $\lim_{x\to 0} u(x)$ *existiert jedoch nicht.)*

Unsere Beobachtungen führen also zu dem Schluß, dass schwache *Normen (d. h. solche, die durch Integrale definiert werden) mit* klassischen *Funktionenräumen*

(wie z. B. C^0) nicht verträglich sind.
Das liegt daran, dass der Wert eines Integrals unverändert bleibt, wenn man aus dem Integrationsbereich eine Nullmenge herausschneidet. Mit anderen Worten: F. ü. identische Funktionen haben die gleiche Integral–Norm.
Gerade die Variationsrechnung *zwingt aber dazu, mit integralen Normen zu arbeiten. Wir kehren daher zum Ausgangspunkt unserer Überlegungen zurück und erweitern den Begriff der μ–messbaren Funktion in der Art, dass das Problem ii) der Indefinitheit von $\|\cdot\|_1$ nicht mehr auftritt. Die naheliegende Idee um dies zu erreichen ist, die f. ü. identischen Funktionen zu einer Einheit zusammenzufassen, also Äquivalenzklassen solcher Funktionen zu bilden:*
Sei (X,μ) ein Maßraum und für eine μ–messbare Funktion $u : X \to \mathbb{R}$ sei

$$[u] := \big\{ w : X \to \overline{\mathbb{R}};\ u \sim w \big\}$$

die Äquivalenzklasse von u bzgl. der Relation

$$u \sim w \quad :\Longleftrightarrow \quad u = w \quad \mu\text{–f. ü. auf } X.$$

Es gilt:

*i) u ist μ–*integrierbar*, d. h. es ist $\int_X \big|u(x)\big|\, d\mu(x) < \infty$, falls jedes $w \in [u]$ μ–*integrierbar *ist. In diesem Fall ist offenbar*

$$\int_X \big|u(x)\big|\, d\mu(x) = \int_X \big|w(x)\big|\, d\mu(x) \quad \text{für alle} \quad w \in [u].$$

ii) $[0] = \big\{ w : X \to \overline{\mathbb{R}};\ w = 0\ \mu\text{–f. ü. auf } X \big\}$.

iii) Sind $u, w : X \to \overline{\mathbb{R}}$ μ–messbar und μ–f. ü. endlich (d. h. μ–f ü. reell) auf X, so machen $[u+w]$ und $[cu]$ ($c \in \mathbb{R}$) Sinn.

Definition 3.1 (Fast überall definierte Funktion)
Eine μ–f.ü. (eindeutig) definierte Funktion von $X \to \overline{\mathbb{R}}$ ist die Äquivalenzklasse $[u]$ einer μ–messbaren und μ–f. ü. endlichen Funktion $u : X \to \overline{\mathbb{R}}$.

Als Beispiel betrachte man die Äquivalenzklasse der Funktion $u(x) = \frac{x}{|x|}$: für jeden Vertreter kann in $x = 0$ ein beliebiger Wert vorgegeben werden.

Wir vergessen also die Äquivalenzklasse $[u]$ und reden von einer μ–f. ü. eindeutig definierten Funktion u (später werden wir auch wieder nur von einer Funktion reden, wohlwissend, dass es sich dabei um eine Äquivalenzklasse von Funktionen handelt). Diese kann natürlich nicht mehr punktweise ausgewertet werden;

es gibt lediglich ein eindeutig bestimmtes Integral (sofern dieses existiert). Eine punktweise Auswertung ist demnach nur nach Wahl eines Vertreters bzw. Repräsentanten für $[u]$ möglich.

Bemerkung 3.2

i) *Eine μ–f. ü. definierte Funktion $u : X \to \overline{\mathbb{R}}$ ist p. d. $\leq, =, \geq 0$, falls entsprechendes für jeden Vertreter der zugeh. Äquivalenzklasse $[u]$ zutrifft. Beispielsweise bedeutet $\left[\mathbb{1}_{\mathbb{Q}}\right] = 0$, dass jede mit $\mathbb{1}_{\mathbb{Q}}$ $\mathcal{L}^1$–f. ü. auf $\mathbb{R}$ übereinstimmende Funktion $\mathcal{L}^1$–f. ü. identisch der Nullfunktion ist.*

ii) *Sei $\Omega \subset \mathbb{R}^d$ offen und $u : \Omega \to \mathbb{R}$ sei eine $\mathcal{L}^d$–messbare Funktion. Dann gibt es in $[u]$ höchstens einen stetigen Vertreter (vgl. A. 2.**??**). Allgemein braucht eine solche Funktion also nicht einmal stetig zu sein.*
Gibt es für die fast überall eindeutig definierte Funktion u genau einen stetigen Vertreter (bzw. genau einen Vertreter der Klasse C^k mit einem $k \in [1, \infty]$), so nennt man u selbst wieder stetig (bzw. von der Klasse C^k) und schreibt wie üblich wieder $u \in C^0(\Omega)$ (bzw. $u \in C^k(\Omega)$). (Man beachte, dass diese Sprechweise nur dann Sinn macht, wenn es nur genau einen solchen Vertreter gibt!)

Definition 3.3 (Lebesgue–Raum)
Sei (X, μ) ein Maßraum und sei $1 \leq p < \infty$. Dann heißt der durch

$$L^p(X;\mu) := \left\{u : X \to \overline{\mathbb{R}};\ u\ \mu\text{–f. ü. definiert mit } \|u\|_p < \infty\right\}$$

mit

$$\|u\|_p := \|u\|_{p;X} := \left(\int_X \left|u(x)\right|^p d\mu(x)\right)^{\frac{1}{p}} \in [0, \infty]$$

erklärte normierte Raum $\left(L^p(X;\mu), \|\cdot\|_p\right)$ der Lebesgue–Raum der auf X bzgl. dem Maß μ p–summierbaren Funktionen.

Beispiel 3.4

i) *Sei $\mathbb{N}$ versehen mit dem Zählmaß $\mu_{\#}$. Dann sind die bzgl. $\mu_{\#}$ messbaren Funktionen Folgen $u := (u_n) \subset \mathbb{R}$, und es gibt nur eine $\mu_{\#}$–Nullmenge, nämlich die leere Menge. Man erhält hier*

$$\int_{\mathbb{N}} u(x)\, d\mu_{\#}(x) = \sum_{n=1}^{\infty} u_n$$

und schreibt auch ℓ^p *statt* $L^p(\mathbb{N};\mu_\#)$. *Dann ist*

$$u \in \ell^p \quad \Longleftrightarrow \quad \|u\|_p = \left(\sum_{n=1}^{\infty} \left|u_n(x)\right|^p\right)^{\frac{1}{p}} < \infty$$

ist. Der Raum ℓ^p *heißt der Raum der* p*–summierbaren Folgen. Insbesondere ist* ℓ^1 *genau der Raum der absolut konvergenten Zahlenreihen.*

ii) Für $\Omega \subset \mathbb{R}^d$ *schreiben wir üblicherweise* $L^p(\Omega)$ *statt* $L^p(\Omega;\mathcal{L}^d)$. *Weiter unten werden wir auch* L^p*–Räume* $L^p(\Omega)^D = L^p(\Omega,\mathbb{R}^D)$ *für Funktionen* $\Omega \to \mathbb{R}^D$ ($D \in \mathbb{N}$ *mit* $D \geq 2$) *erklären.*

Satz 3.5 (Vollständigkeit der Lebesgue–Räume)
Sei (X,μ) *ein Maßraum und sei* $1 \leq p < \infty$. *Dann ist* $L^p(X;\mu)$ *ein linearer Raum, welcher vermöge* $\|\cdot\|_p$ *zu einem Banach–Raum wird.*
Insbesondere sind also $\big(L^p(\Omega), \|\cdot\|_p\big)$ *und* $\big(\ell^p, \|\cdot\|_p\big)$ *Banach–Räume.*

Zum Beweis dieser Aussage benötigt man das folgende. [b]

Lemma 3.6

i) **(Hölder–Ungleichung)**
Seien $1 < p, q < \infty$ *konjugierte Exponenten, d. h. es gelte* $\frac{1}{p} + \frac{1}{q} = 1$ (*also* $q = \frac{p}{p-1}$), *und seien* $u \in L^p(X;\mu)$ *sowie* $w \in L^q(X;\mu)$. *Dann ist* $uw \in L^1(X;\mu)$ *und es gilt*

$$\int_X |uw| \, d\mu(x) \leq \|u\|_p \, \|w\|_q.$$

ii) **(Minkowski–Ungleichung)**
Sei $1 \leq p < \infty$ *und seien* $u, w \in L^p(X;\mu)$. *Dann ist auch* $u + w \in L^p(X;\mu)$ *und es gilt*

$$\|u + w\|_p \leq \|u\|_p + \|w\|_p.$$

In der Hölder–Ungleichung gilt Gleichheit, falls $u = cw$ *(f. ü.) mit einem* $c \in \mathbb{R}$ *ist.*

[b] Die Hölder–Ungleichung gilt auch mit den Wahlen $p = \infty$ und $q = 1$ (mit der Konvention $\frac{1}{\infty} := 0$). Dies wird später klar, wenn wir den Raum $L^\infty(X;\mu)$ erklärt haben. Entsprechendes gilt für die Minkowski–Ungleichung.

Beweis.

i) Seien $s, t, \alpha, \beta > 0$ *mit* $\alpha + \beta = 1$. *Wegen der Konkavität des Logarithmus ist dann (vgl. Fig. 2)*

$$\log(\alpha s + \beta t) \geq \alpha \log s + \beta \log t,$$

also $\alpha s + \beta t \geq s^\alpha\, t^\beta$.

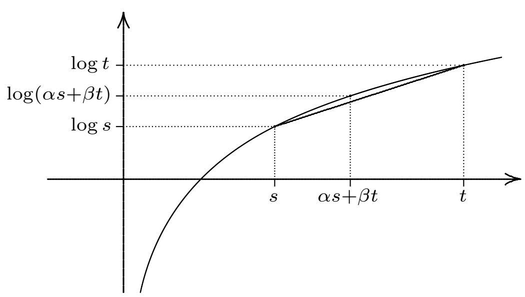


Fig. 2

Sind nun die konjugierten Exponenten $\alpha := \frac{1}{p}$ *und* $\beta := \frac{1}{q}$ *und wählen wir* $s := x^{1/\alpha}$, $t := y^{1/\beta}$ *mit* $x, y > 0$, *so erhalten wir die Ungleichung:*

$$xy \leq \frac{1}{p}\, x^p + \frac{1}{q}\, y^q. \qquad (3.3)$$

Natürlich gilt (3.3) *trivialerweise, wenn* $x = 0$ *oder* $y = 0$ *ist. Seien nun* $\|u\|_p$, $\|w\|_q > 0$ *(sonst ist die Behauptung trivial). Wir setzen nun* $x := \frac{|u|}{\|u\|_p}$ *und* $y := \frac{|w|}{\|w\|_q}$ *und gelangen mit* (3.3) *zu*

$$\frac{|u|}{\|u\|_p}\, \frac{|w|}{\|w\|_q} \leq \frac{1}{p}\, \frac{|u|^p}{\|u\|_p^p} + \frac{1}{q}\, \frac{|w|^q}{\|w\|_q^q} \quad \mu\text{–}f.\,ü.\ auf\ X,$$

und Integration über X *liefert*

$$\frac{1}{\|u\|_p}\, \frac{1}{\|w\|_q} \int_X |uw|\, d\mu \leq \frac{1}{p}\, \frac{1}{\|u\|_p^p} \int_X |u|^p\, d\mu + \frac{1}{q}\, \frac{1}{\|w\|_q^q} \int\limits_X |w|^q\, d\mu = 1,$$

und damit die Behauptung.

ii) Der Fall $p = 1$ *ist offensichtlich. Sei also* $p > 1$ *und* $q := \frac{p}{p-1}$ *(d. h.* p *und* q *sind konjugierte Exponenten). Wir zeigen zunächst, dass* $|u + w|^p$ *integrierbar, also* $u + w \in L^p(X; \mu)$ *ist. Es ist*

$$\begin{aligned} |u + w|^p &\leq \big(|u| + |w|\big)^p \leq \big(2 \max\big\{|u|, |w|\big\}\big)^p \\ &= 2^p \max\big\{|u|^p, |w|^p\big\} \leq 2^p\big(|u|^p + |w|^p\big), \end{aligned}$$

wobei Repräsentanten für u *und* w *gewählt, und das Maximum punktweise* *gebildet wurde, d. h. die Ungleichungen gelten μ–f. ü. auf X. Damit ist*

$$\int_X |u+w|^p \, d\mu \le 2^p \left(\int_X |u|^p \, d\mu + \int_X |w|^p \, d\mu \right),$$

also $u + w \in L^p(X;\mu)$ gezeigt. Nun wird mit i)

$$\begin{aligned}
\|u+w\|_p^p &= \int_X |u+w|^p \, d\mu = \int_X |u+w|^{p-1}|u+w| \, d\mu \\
&\le \int_X |u||u+w|^{p-1} \, d\mu + \int_X |w||u+w|^{p-1} \, d\mu \\
&\le \|u\|_p \left(\int_X |u+w|^{q(p-1)} \, d\mu \right)^{\frac{1}{q}} + \|w\|_p \left(\int_X |u+w|^{(p-1)q} \, d\mu \right)^{\frac{1}{q}} \\
&= \big(\|u\|_p + \|w\|_p\big)\|u+w\|_p^{\frac{p}{q}},
\end{aligned}$$

woraus die Behauptung für $\|u+w\|_p > 0$ unmittelbar folgt (sonst gilt das Behauptete trivialerweise). □

Beweis von Satz 3.5.
Sei $u \in L^p(X;\mu)$. Ist $\|u\|_p = 0$, so ist $u \equiv 0$ (genauer: $[u] = 0$). Ferner ist $\|cu\|_p = |c|\|u\|_p$ für alle $c \in \mathbb{R}$ und wegen Lemma 3.6 ii) gilt die Dreiecksungleichung, d. h. $L^p(X;\mu)$ wird vermöge $\|\cdot\|_p$ zu einem normierten Raum.
Sei dazu $(u_n) \subset L^p(X;\mu)$ eine Cauchy–Folge bzgl. $\|\cdot\|_p$. Dann genügt es zu zeigen, dass eine Teilfolge $\big(u_{n_k}\big)_k$ von (u_n) gegen eine Funktion $u \in L^p(X;\mu)$ konvergiert. Denn eine Cauchy–Folge mit konvergenter Teilfolge ist bereits selbst konvergent:

$$\|u_n - u\|_p \le \|u_{n_k} - u\|_p + \|u_{n_k} - u_n\|_p \xrightarrow{k} 0.$$

Wegen der Cauchy–Bedingung für (u_n) gibt es zu jedem $k \in \mathbb{N}$ ein $N_k \in \mathbb{N}, N_k \ge k$ derart, dass

$$\sum_{k=1}^{\infty} \left\|u_{N_{k+1}} - u_{N_k}\right\|_p \le \sum_{k=1}^{\infty} 2^{-k} = 1. \tag{3.4}$$

Setzen wir $\widetilde{u}_k := u_{N_k}$ und $\omega_\nu := \sum_{k=1}^{\nu} |\widetilde{u}_{k+1} - \widetilde{u}_k|$ für $\nu \in \mathbb{N}$, so ist wegen (3.4) $\|\omega_\nu\|_p \le 1$. Nach dem Lemma von Fatou ist dann

$$\int_X |\lim_\nu \omega_\nu|^p \, d\mu \le \liminf_\nu \int_X |\omega_\nu|^p \, d\mu = \liminf_\nu \|\omega_\nu\|_p^p \le 1,$$

d. h. $\lim_\nu \omega_\nu$ existiert für μ–f. a. $x \in X$. Damit haben wir

$$\left|\widetilde{u}_k(x) - \widetilde{u}_l(x)\right| \le \sum_{\nu=l}^{k} \left|\widetilde{u}_{\nu+1}(x) - \widetilde{u}_\nu(x)\right| \xrightarrow{k,l} 0 \quad \text{für } \mu\text{–f. a. } x \in X,$$

weshalb also $(\widetilde{u}_k)$ für μ–f. a. $x \in X$ eine Cauchy–Folge in $\mathbb{R}$ ist. Die punktweise Grenzfunktion $u(x) := \lim_k \widetilde{u}_k(x)$ existiert für μ–f. a. $x \in X$ ($[u]$ ist dadurch wohldefiniert). Bleibt zu zeigen, dass $u \in L^p(X;\mu)$ ist. Wieder unter Verwendung des Lemmas von Fatou erhalten wir aus der punktweisen Konvergenz von (w_k):

$$\begin{aligned}
\int_X |u - \widetilde{u}_\nu|^p \, d\mu &= \int_X \lim_k |\widetilde{u}_k - \widetilde{u}_\nu|^p \, d\mu \le \liminf_k \int_X |\widetilde{u}_k - \widetilde{u}_\nu|^p \, d\mu \\
&= \liminf_k \|\widetilde{u}_k - \widetilde{u}_\nu\|_p^p = \left(\liminf_k \|\widetilde{u}_k - \widetilde{u}_\nu\|_p\right)^p \\
&\le \left(\liminf_k \sum_{j=\nu}^{k-1} \|\widetilde{u}_{j+1} - \widetilde{u}_j\|_p\right)^p \overset{(3.4)}{\le} \left(\sum_{j=\nu}^{\infty} 2^{-j}\right)^p \xrightarrow{\nu} 0,
\end{aligned}$$

ergo $u - \widetilde{u}_\nu \in L^p(X;\mu)$ für alle $\nu \in \mathbb{N}$, und damit auch $u \in L^p(X;\mu)$. □

Unmittelbar aus der soeben durchgeführten Konstruktion ergibt sich:

Korollar 3.7 (Punktweise Konvergenz fast überall)
Seien $1 \le p < \infty$ und $(u_n) \subset L^p(X;\mu)$ eine Folge mit $u_n \xrightarrow{n}: u$ in $L^p(X;\mu)$. Dann gibt es eine Teilfolge von (u_n) (o. E. (u_n) selbst) derart, dass (nach Wahl eines Vertreters) gilt: $u_n(x) \xrightarrow{n} u(x)$ für μ–f. a. $x \in X$.

L^p–Konvergenz impliziert also punktweise Konvergenz fast überall (wenigstens) für eine Teilfolge für „geeignete“ Vetreter. Dabei bedeutet „geeignet“, dass man für jedes Folgenglied und auch für die Grenzfunktion einen beliebigen Repräsentanten zu wählen hat. Es ist demnach allgemein falsch, dass die Folge selbst punktweise f. ü. konvergiert.

Beispiele.

i) *ℓ^p ist ein Banach–Raum.*

ii) *$L^p(\Omega)$ ist ein Banach–Raum für $\Omega \subset \mathbb{R}^d$.*

iii) *$L^p(\omega;\mathcal{H}^s)$ für eine s–dimensionale (genügend glatte) Untermannigfaltigkeit $\omega \subset \mathbb{R}^d$. Dabei bezeichnet $\mathcal{H}^s$ das s–dimensionale Hausdorff–Maß.*

In unserer Skala fehlt noch der Raum $L^\infty(X;\mu)$, dessen Definition durch die Äquivalenzklassenbildung etwas erschwert wird.

Definition 3.8 (Essentielles Supremum/Infimum)
Sei (X,μ) ein Maßraum und sei u eine μ–f. ü. eindeutig definierte Funktion. Dann heißt die Größe

$$\begin{aligned}\operatorname*{ess\,sup}_{x\in X} u(x) &:= \inf\big\{c\in\mathbb{R};\ w\le c\ \mu\text{–f.ü. auf } X \text{ für jedes } w\in[u]\big\}\\ &= \inf_{\mu(N)=0}\left\{\sup_{x\in X-N} u(x)\right\}\in(-\infty,\infty]\end{aligned}$$

das essentielle (oder wesentliche) Supremum von u auf X. Entsprechend ist das essentielle (oder wesentliche) Infimum $\operatorname{ess\,inf}_{x\in X} u(x)$ von u auf X erklärt.

Mit dem essentiellen Supremum lässt sich nun eine Norm erklären:

$$\|u\|_\infty := \|u\|_{\infty;\,X} := \operatorname*{ess\,sup}_{x\in X}\big|u(x)\big|,$$

Tatsächlich ist $\|\cdot\|_\infty$ eine Norm, welche den Raum $L^\infty(X;\mu)$ der μ–f. ü. auf X eindeutig definierten, beschränkten Funktionen, den wir gleich definieren, zu einem Banach–Raum macht. Wir nennen die Norm $\|\cdot\|_\infty$ auch wieder Supremum–Norm.

Bemerkung 3.9

i) Es ist $\|u\|_\infty<\infty$, falls es einen Vertreter w von u gibt, der außerhalb einer μ–Nullmenge beschränkt ist. In diesem Fall ist $\|u\|_\infty$ die kleinste dabei vorkommende Schranke.

ii) Seien $X:=\mathbb{R}$ und $\mu:=\mathcal{L}^1$. Dann ist $\sup_{\mathbb{R}}\mathbb{1}_{\mathbb{Q}}=1$, aber $\operatorname{ess\,sup}_{\mathbb{R}}\mathbb{1}_{\mathbb{Q}}=0$. Bei der Berechnung des „gewöhnlichen Supremums“ spielt jeder Funktionswert eine Rolle, beim essentiellen Supremum werden dagegen Werte auf Nullmengen ignoriert.

Satz 3.10 (Vollständigkeit von L^∞)
Sei (X,μ) ein Maßraum. Dann ist der Raum

$$L^\infty(X;\mu) := \big\{u: X\to\overline{\mathbb{R}};\ u\ \mu\text{–f. ü. definiert mit } \|u\|_\infty<\infty\big\}$$

der μ–f. ü. eindeutig definierten, beschränkten Funktionen vollständig bzgl. der Supremum–Norm $\|\cdot\|_\infty$.

Beispiele.

i) ℓ^∞ *ist ein Banach–Raum, der Raum der beschränkten Folgen in* $\mathbb{R}$ *(vgl. Bsp. 3.4).*

ii) $L^\infty(\Omega)$ *ist ein Banach–Raum für* $\Omega \subset \mathbb{R}^d$.

iii) $L^\infty(\omega; \mathcal{H}^k)$ *für eine* k*–dimensionale Untermannigfaltigkeit* $\omega \subset \mathbb{R}^d$.

Beweis von Satz 3.10.
Zunächst überlegt man sich leicht, dass durch $\|\cdot\|_\infty$ *tatsächlich eine Norm auf* $L^\infty(X;\mu)$ *erklärt wird. Dies sei dem Leser als Übung überlassen (vgl. A. 2.??). Wir zeigen die Vollständigkeit: Sei dazu* (u_n) *eine Cauchy–Folge in* $L^\infty(X;\mu)$. *Wir wählen Vertreter für* u_n *(wieder mit* u_n *bezeichnet), und wissen, dass*

$$\big|u_n(x) - u_m(x)\big| \le \|u_n - u_m\|_\infty \xrightarrow{n,m} 0 \quad \text{für } \mu\text{–f. a. } x \in X,$$

d. h. für alle $x \in X$ *außerhalb einer* μ*–Nullmenge* N.[c] *Ferner existiert eine positive Konstante* c *mit* $\big|u_n(x)\big| \le \|u\|_\infty \le c$ *für* μ*–f. a.* $x \in X$ *und alle* $n \in \mathbb{N}$. *Also ist*

$$u(x) := \begin{cases} \lim\limits_n u_n(x) & ; \quad x \in X \setminus N \\ 0 & ; \quad x \in N \end{cases}$$

für alle $x \in X$ *wohldefiniert und* μ*–messbar (als punktweise Grenzfunktion* μ*–messbarer Funktionen).*[d] *Für* $x \notin N$ *ist*

$$\big|u(x) - u_m(x)\big| = \lim_n \big|u_n(x) - u_m(x)\big| \le \liminf_n \|u_n - u_m\|_\infty,$$

also $\|u - u_m\|_\infty \le \liminf_n \|u_n - u_m\|_\infty \xrightarrow{m} 0$. *Daher ist* $u - u_m \in L^\infty(X;\mu)$ *und die Behauptung folgt.* □

Beobachtungen.

i) Sind $u_n, u \in L^\infty(X;\mu)$ *mit* $u_n \xrightarrow{n} u$ *in* $L^\infty(X;\mu)$, *so strebt für alle Vertreter* $u_n \xrightarrow{n} u$ *gleichmäßig außerhalb einer* μ*–Nullmenge. Mit anderen Worten ist* $\|\cdot\|_\infty$ *die Norm der* μ*–f. ü. gleichmäßigen Konvergenz.*

ii) Ist $\mu(X) < \infty$, *so ist* $L^p(X;\mu) \supset L^q(X;\mu)$ *für alle* $1 \le p \le q \le \infty$ *und es gilt in diesem Fall*

$$\|u\|_p \le \|u\|_q \, \mu(X)^{1/p-1/q}.$$

[c] Hier geht ein, dass abzählbare Vereinigungen von Nullmengen wieder Nullmengen sind.
[d] Statt u auf N identisch 0 zu setzen, hätte man natürlich auch jeden anderen Wert nehmen können.

> *Man überlege sich etwa für $X = \mathbb{R}$ und $\mu = \mathcal{L}^1$ Beispiele die zeigen, dass $\mu(X) < \infty$ keine unnötige Voraussetzung ist (vgl. auch A. 2.??).*

Unser nächstes Ziel ist die Charakterisierung der zu den Räumen L^p gehörigen dualen Räume linearer, stetiger Abbildungen von $L^p \to \mathbb{R}$. Dazu beschäftigen wir uns vorweg etwas allgemeiner mit linearen Abbildungen zwischen normierten Räumen.
Ist $\big(X, \|\cdot\|\big)$ ein normierter Raum, so bezeichnen wir mit $\mathbb{B}$ die offene Einheitskugel in X, also

$$\mathbb{B} := \big\{x \in X;\ \|x\| < 1\big\}.$$

Definition 3.11 (Stetige lineare Abbildung)
Seien X, Y normierte $\mathbb{R}$–Vektorräume. Eine lineare Abbildung $T : X \to Y$ heißt stetig (oder auch beschränkt), falls die Größe

$$\|T\|_\infty := \|T\|_{\infty;\,X,Y} = \sup_{x \in \overline{\mathbb{B}}} \|Tx\| = \sup_{x \in X\setminus\{0\}} \frac{\|Tx\|}{\|x\|}$$

endlich ist.

Vorsicht! *Der Terminus „beschränkt“ bedeutet hier nicht $\|Tx\| \leq const$ für alle $x \in X$, sondern $\|Tx\| \leq \|T\|_\infty \|x\|$ für alle $x \in X$. Um solche Verwechslungen auszuschließen, sprechen wir im Folgenden stets von einer „stetigen“, linearen Abbildung. Das folgende Lemma rechtfertigt diese Nomenklatur.*

Vorweg erinnern wir an den Begriff der Lipschitz–Stetigkeit. Sind X und Y normierte $\mathbb{R}$–Vektorräume, so heißt eine Abbildung $T : X \to Y$ Lipschitz–stetig *(oder* dehnungsbeschränkt*) mit Lipschitz–Konstante $L \in [0, \infty)$, falls*

$$\|Tx - Ty\| \leq L\|x - y\| \quad \text{für alle} \quad x, y \in X$$

gilt, wobei L die kleinste Zahl mit dieser Eigenschaft ist. Genauer ist:

$$L := \sup_{\substack{x,y \in X \\ x \neq y}} \frac{\|Tx - Ty\|}{\|x - y\|} = \|T\|_\infty$$

(sofern dieses Supremum existiert).

Lemma 3.12
Sei $T : X \to Y$ eine lineare Abbildung. Dann sind äquivalent:

i) $\|T\|_\infty < \infty$.

ii) T *ist stetig in* 0.

iii) T *ist überall in* X *stetig.*

iv) T *ist Lipschitz–stetig.*

Beweis.

Sei $L := \|T\|_\infty < \infty$. *Per Definition ist* $\|Tx\| \leq L$ *für alle* $x \in X$ *mit* $\|x\| \leq 1$. *Für beliebiges* $z \neq 0$ *sei* $x := \frac{z}{\|z\|}$. *Dann ist* $\|Tz\| \leq L\|z\|$, *und mit* $z := x - y$ *folgt*

$$\|Tx - Ty\| \leq L\|x - y\| \quad \text{für alle} \quad x, y \in X,$$

also T *Lipschitz–stetig mit Lipschitz–Konstante* L. *Insbesondere ist* T *überall stetig, also auch im Ursprung.*
Sei nun umgekehrt T *Lipschitz–stetig. Für* $y = 0$ *wird dann* $\|Tx\| \leq L\|x\|$, *also* $\|Tx\| \leq L$ *für alle* $x \in X$ *mit* $\|x\| \leq 1$.
Sei T *stetig in 0, dann ist* $\|Tx\| \leq 1$ *falls* $\|x\| \leq \epsilon$ *ist für ein* $\epsilon > 0$. *Für* $x \in \mathbb{B}$ *ist* $\|\epsilon x\| \leq \epsilon$ *und daher*

$$\|T(x\epsilon)\| \leq 1 \Leftrightarrow \|Tx\| \leq \frac{1}{\epsilon}$$

und es folgt $\|T\| \leq \frac{1}{\epsilon}$, *also (i).* □

Bemerkung 3.13

i) *Die Größe* $\|T\|_\infty$ *heißt (falls sie endlich ist) die Operator–Norm von* T. *Stetige, lineare Abbildungen* T *nennt man auch lineare Operatoren — daher der Name der Norm.*

ii) *Die Menge der stetigen, linearen Abbildungen*

$$\mathscr{L}(X,Y) := \big\{T : X \to Y \ \textit{linear};\ \|T\|_\infty < \infty\big\}$$

ist ein linearer Raum. Wir schreiben $\mathscr{L}(X)$ *statt* $\mathscr{L}(X,X)$.
Sind $S, T \in \mathscr{L}(X,Y)$, *so ist für alle* $x \in X$ *mit* $\|x\| \leq 1$

$$\big\|(S+T)(x)\big\| = \|Sx + Tx\| \leq \|Sx\| + \|Tx\| \leq \|S\|_\infty + \|T\|_\infty$$

oder also $\|S+T\|_\infty \leq \|S\|_\infty + \|T\|_\infty$, *d. h. durch* $\|\cdot\|_\infty$ *wird tatsächlich eine Norm auf* $\mathscr{L}(X,Y)$ *erklärt.*

iii) *Für* $\dim X, \dim Y < \infty$ *sind alle linearen Abbildungen stetig, ganz unabhängig davon, welche Normen* X *und* Y *tragen. Dies ist i. a. falsch für unendlich–dimensionales* X *(vgl. Übung).*

iv) Sei (X, μ) ein Maßraum, wobei $\mu(X) < \infty$ sei. Wir setzen

$$T : L^\infty(X;\mu) \to L^1(X;\mu),\ Tu := u$$

(T ist der sog. Einbettungsoperator). Dann ist

$$\|Tu\|_1 = \int_X |u|\, d\mu \le \mu(X)\, \|u\|_\infty,$$

so dass $\|T\|_\infty = \mu(X)$ ist.

v) Man nennt $X^* := \mathscr{L}(X, \mathbb{R})$ den zu X gehörigen Dualraum*, bestehend aus den stetigen linearen Funktionalen von $X \to \mathbb{R}$.* [e]

Satz 3.14 (Vollständigkeit des Raums der stetigen linearen Abbildungen)
Ist Y ein Banach–Raum, so ist $\mathscr{L}(X, Y)$ vollständig bzgl. der Operator–Norm. Insbesondere ist X^* ein Banach–Raum (X selbst muss aber kein Banach–Raum sein).

Beweis.
Sei $(T_n) \subset \mathscr{L}(X, Y)$ eine Cauchy–Folge. Dann strebt

$$\|T_n x - T_m x\| \le \|T_n - T_m\|_\infty \|x\| \xrightarrow{n,m} 0 \quad \textit{für alle} \quad x \in X,$$

und wegen der Vollständigkeit von Y existiert daher $Tx := \lim_n T_n x$ für jedes $x \in X$. Da (T_n) eine Cauchy–Folge ist, ist (T_n) beschränkt, also $\|T_n\|_\infty \le c$ für alle $n \in \mathbb{N}$ mit einer positiven Konstante c. Wegen der Linearität von T ist für alle $x \in X$

$$\|Tx\| \le \|T_n x - Tx\| + \|T_n x\| \le c\|x\| + \|T_n x - Tx\| \xrightarrow{n} c\|x\|,$$

und dies zeigt $T \in \mathscr{L}(X, Y)$.
Bleibt zu zeigen, dass $\|T_n - T\|_\infty \xrightarrow{n} 0$. Sei dazu $\varepsilon > 0$ beliebig vorgegeben, und $x \in X$ mit $\|x\| \le 1$ fixiert. Da (T_n) punktweise gegen T konvergiert, existiert ein $n_0 = n_0(\varepsilon, x) \in \mathbb{N}$ mit

$$\|Tx - T_{n_0} x\| < \frac{\varepsilon}{2}, \tag{3.5}$$

[e] Auf unendlich–dimensionalen Räumen X sind die linearen Abbildungen $X \to \mathbb{R}$ nicht automatisch stetig, wie dies für endlich–dimensionale X der Fall ist. Daher spricht man hier auch vom *stetigen* oder *topologischen* Dualraum X^*, welcher in diesem Fall ein echter Teilraum des *algebraischen* Dualraums *aller* linearen Abbildungen $X \to \mathbb{R}$ ist. In der Literatur findet man auch die Bezeichnung X' für den stetigen Dualraum.

und wegen der Cauchy–Bedingung existiert ein $n_1 = n_1(\varepsilon) \in \mathbb{N}$*, so dass*

$$\|T_n - T_m\|_\infty < \frac{\varepsilon}{2} \quad \text{für alle} \quad n, m > n_1. \tag{3.6}$$

Nehmen wir o. E. $n_0 > n_1$ *an, so wird mit* (3.5) *und* (3.6) *für alle* $n > n_1$

$$\|Tx - T_n x\| \le \|Tx - T_{n_0} x\| + \|T_n x - T_{n_0} x\| < \varepsilon,$$

womit alles bewiesen ist. □

Nach diesen allgemeinen Betrachtungen wenden wir uns wieder den Lebesgue-Räumen zu, deren Dualräume $L^p(X;\mu)^* = \mathscr{L}\big(L^p(X;\mu), \mathbb{R}\big)$ *wir näher studieren wollen. Wir betrachten vorweg ein Beispiel.*

Beispiel.
Sei (X, μ) *ein Maßraum. Dann wird durch* $T : L^1(X;\mu) \to \mathbb{R}$,

$$Tu := \int_X u \, d\mu$$

ein stetiges, lineares Funktional auf $L^1(X;\mu)$ *definiert. Mit anderen Worten ist* $T \in L^1(X;\mu)^*$. *Denn für jede Funktion* $u \in L^1(X;\mu)$ *ist*

$$|Tu| = \left| \int_X u \, d\mu \right| \le \int_X |u| \, d\mu = \|u\|_1.$$

Für $\mu(X) < \infty$ *liegt* T *sogar in allen Dualräumen* $L^p(X;\mu)^*$ *mit* $1 \le p \le \infty$ *(warum?).*

Bemerkung 3.16
Seien (X, μ) *ein Maßraum,* $1 \le p \le \infty$ *und*

$$q := p' := \begin{cases} \infty & ; \quad p = 1 \\ \frac{p}{p-1} & ; \quad 1 < p < \infty \\ 1 & ; \quad p = \infty. \end{cases} \tag{3.7}$$

Dann ist für jedes fixierte $w \in L^q(X;\mu)$ *das Funktional* $T_w : L^p(X;\mu) \to \mathbb{R}$,

$$T_w u := \int_X uw \, d\mu \tag{3.8}$$

wohldefiniert. Nach der Hölder–Ungleichung (Lemma 3.6 i)) ist

$$|T_w u| \le \|u\|_p \|w\|_q,$$

oder also $\|T_w\|_\infty \le \|w\|_q$*, d. h. es ist* $T_w \in L^p(X;\mu)^*$*. Es gilt sogar*

$$\|T_w\|_\infty = \|w\|_q \quad \textit{für alle} \quad w \in L^q(X;\mu). \tag{3.9}$$

Letzteres sieht man wie folgt: Für $1 \le q < \infty$ *und ein* $w \in L^q(X;\mu)$ *sei* $u := |w|^{q-1}\operatorname{sgn}(w)$*, wobei*

$$\operatorname{sgn}(w) := \begin{cases} \frac{w}{|w|} & ; \quad [w \neq 0] \\ 0 & ; \quad [w = 0] \end{cases} .$$ [f]

Im Fall $q = 1$ *ist dann* $u = \operatorname{sgn}(w)$ *und*

$$T_w u = \int_X |w| \, d\mu = \|w\|_1 \in L^\infty(X;\mu).$$

Im Fall $1 < q < \infty$ *ist*

$$\int_X |u|^p \, d\mu = \int_X |w|^q \, d\mu = \int_X uw \, d\mu = T_w u,$$

und wegen $\|u\|_p = \|w\|_q^{q/p}$ *ist (o. E. sei* $\|u\|_p \neq 0$*, sonst ist* $w = 0$*)*

$$T_w\left(\frac{u}{\|u\|_p}\right) = \|w\|_q.$$

Der Fall $q = \infty$ *ist etwas aufwendiger.*

Halten wir also fest: Die Abbildung $\mathfrak{J} : L^q(X;\mu) \to L^p(X;\mu)^*$*, gegeben durch*

$$\mathfrak{J}w := T_w$$

ist eine lineare Isometrie, d. h. $\mathfrak{J}$ *ist linear und längentreu,*

$$\|T_w\|_\infty = \|w\|_q \quad \textit{für alle} \quad w \in L^q(X;\mu),$$

wobei die letzte Identität nach (3.9) *gilt. Insbesondere ist* $\mathfrak{J}$ *injektiv, denn* $\mathfrak{J}(w) = 0$ *heißt* $T_w = 0$*, was nach* (3.9) $w = 0$ *ergibt. Schwieriger ist der Beweis der Surjektivität (vgl. [Al] Satz 4.12, S. 159), der jedoch nur für* $p < \infty$

[f] $[w \neq 0]$ steht für die Teilmenge des Definitionsbereichs (hier X) der Funktion w, wo $w \neq 0$ ist.

(und $q > 1$) gelingt. Insgesamt folgt:

Satz 3.17 (Riesz)
Sei $1 \leq p < \infty$ und sei q wie in (3.7). Dann ist die Abbildung

$$\mathfrak{J} : L^q(X;\mu) \to L^p(X;\mu)^*, \quad w \mapsto T_w,$$

mit T_w wie in (3.8), ein (isometrischer) Isomorphismus, d. h. es ist

$$L^p(X;\mu)^* \cong L^q(X;\mu)\,^{\mathrm{g}}.$$

Wir schreiben auch kürzer $(L^p)^* = L^q$. Man beachte, dass diese Aussage im Fall $p = \infty$ falsch ist, wie wir weiter unten noch sehen werden.

Beispiel 3.18

i) Für $p = 1$ ist $L^1(X;\mu)^* \cong L^\infty(X;\mu)$, d. h. ist $T : L^1(X;\mu) \to \mathbb{R}$ ein lineares und stetiges Funktional, so gibt es eine Funktion $w \in L^\infty(X;\mu)$, so dass $T = T_w$.

ii) Für $p = 2$ ist $L^2(X;\mu)^* \cong L^2(X;\mu)$, d. h. ist $T \in L^2(X;\mu)^*$, so gibt es eine Funktion $w \in L^2(X;\mu)$, so dass $T = T_w$. Die Isomorphie von L^2 und L^{2^*} lässt sich auch ganz ohne Maßtheorie beweisen, wie wir in § 4 feststellen werden.

iii) Speziell ist $(\ell^2)^* \cong \ell^2$ (vgl. Beispiel 3.4 i)), d. h. ist $T \in \left(\ell^2\right)^*$, so existiert eine Folge $w := (w_n) \in \ell^2$ mit $Tw_n = \sum_{n\in\mathbb{N}} u_n w_n = T_w u$ für alle $u := (u_n) \in \ell^2$.

Es gilt tatsächlich

Lemma 3.7 *Für einen Maßraum (X, μ) gilt im Allgemeinen*

$$L^1(X;\mu) \subsetneq L^\infty(X;\mu)^*.$$

Eine Begründung liefert der folgende Satz von Hahn–Banach, den wir an dieser Stelle nicht beweisen wollen (vgl. [Al] Satz 4.14 und Satz 4.15).

Satz 3.19 (Hahn–Banach)
Seien $\left(X, \|\cdot\|\right)$ ein normierter Raum, $U \subset X$ ein Unterraum und $\tau \in U^*$. Dann gibt es eine Fortsetzung $T \in X^*$ von τ, d. h. $T|_U = \tau$ und $\|T\|_\infty = \|\tau\|_\infty$.

[g] Im Fall $p = 1$ ist zusätzlich zu fordern, dass X abzählbar μ–messbar ist, d. h. es gibt abzählbar viele μ–messbare Mengen $A_n \subset X$ ($n \in \mathbb{N}$) mit $\mu(A_n) < \infty$ und $X = \bigcup_{n\in\mathbb{N}} A_n$. Falls $X = \Omega \subset \mathbb{R}^d$ und $\mu = \mathcal{L}^d$ ist dies offenbar immer erfüllt

Beweis von Lemma 3.17: *Sei nun $\tau : C^0_{\mathcal{B}}(\mathbb{R}^d) \to \mathbb{R}$, $\tau u := u(0)$. Dann ist τ offenbar linear und es ist $|\tau u| \leq \|u\|_\infty = \sup_{\mathbb{R}^d} |u| < \infty$, da u stetig und beschränkt ist. Ferner ist $C^0_{\mathcal{B}}(\mathbb{R}^d) \subset L^\infty(\mathbb{R}^d)$.*
Nach dem Satz von Hahn–Banach 3.19 gibt es eine Fortsetzung $T \in L^\infty(\mathbb{R}^d)^*$ von τ. Angenommen, es existiert eine Funktion $w \in L^1(\mathbb{R}^d)$ mit

$$Tu = \int_{\mathbb{R}^d} uw \, dx = T_w u \quad \textit{für alle} \quad u \in L^\infty(\mathbb{R}^d).$$

Dann ist

$$0 = u(0) = \tau u = Tu = T_w u \quad \textit{für alle} \quad u \in C^0_{\mathcal{B}}(\mathbb{R}^d) \textit{ mit } u(0) = 0.$$

Sei nun $\Omega \subset \mathbb{R}^d$ offen mit $0 \notin \Omega$. Dann ist $\int_\Omega uw \, dx = 0$ für jede Funktion $u \in C^0_\circ(\Omega) \subset C^0_{\mathcal{B}}(\Omega)$, und daher $w = 0$ μ–f. ü. auf Ω, also $w = 0$ μ–f. ü. auf $\mathbb{R}^d$. (Dabei haben wir das nachfolgende Fundamentallemma der Variationsrechnung (Lem. 3.22) benutzt.) Damit ist $T \equiv 0$, also auch $\tau \equiv 0$; ein Widerspruch, da es Funktionen u gibt mit $\tau u \neq 0$. Dies zeigt, dass der Dualraum von L^∞ i. a. nicht mit dem Raum L^1 übereinstimmen kann. □

Wir stellen noch eine andere Version des Satzes von Hahn–Banach vor, die wir später benötigen werden:

Satz 3.20
Sei $(X, \|\cdot\|)$ ein normierter Raum. Dann gibt es zu jedem $x \in X \setminus \{0\}$ ein $T \in X^*$ mit $\|T\|_\infty = 1$ und $Tx = \|x\|$.

Beweis.
Wir betrachten zunächst die lineare Hülle $\langle x \rangle$ von x und definieren auf $U := \langle x \rangle$ für $y = \alpha x$ durch $\tau y = \alpha \|x\|$ ein stetig lineares Funktional mit $\|\tau\|_\infty = 1$ und $\tau x = \|x\|$. Mit Satz 3.19 folgt die Behauptung. □
*In § 1 haben wir den Raum $C^\infty_\circ$ der Testfunktionen kennengelernt. Diese Funktionen spielen eine besondere Rolle, wie der folgende Satz zeigt (vgl. auch A. 2.**??**).*

Satz 3.21 (Dichtheit von $C^\infty_\circ$ in L^p)
Sei $\Omega \subset \mathbb{R}^d$ offen und sei $1 \leq p < \infty$. Dann liegt der Raum $C^\infty_\circ(\Omega)$ dicht in $L^p(\Omega)$, d. h. zu jeder Funktion $u \in L^p(\Omega)$ gibt es eine Folge $(\eta_n) \subset C^\infty_\circ(\Omega)$, so dass $\eta_n \xrightarrow{n} u$ in $L^p(\Omega)$ gilt.

Bemerkung.
Für $p = \infty$ ist die Aussage von Satz 3.21 falsch, denn sonst hätte jede Funktion aus $L^\infty(\Omega)$ einen stetigen Vertreter (gleichmäßige Limiten stetiger Funktionen isnd stetig).

Zum Beweis benötigen wir folgende Hilfsausagen

Lemma 3.8 *a) Sei $u \in L^p(\Omega)$, $u \geq 0$ $\mathcal{L}^d$-f.ü. mit $\Omega \subset \mathbb{R}^d$ offen und $1 \leq p < \infty$. Dann gibt es eine Folge einfacher Funktionen (u_n) mit $u_n \leq u_{n+1}$ und $u_n \xrightarrow{n} u$ in $L^p(\Omega)$.*

b) Sei $u : \Omega \to \mathbb{R}$ eine einfache Funktion, dann gibt es eine Folge $\eta_n \subset C_0^\infty(\Omega)$ mit $\eta_n \xrightarrow{n} u$ in $L^p(\Omega)$ für alle $1 \leq p < \infty$.

Beweis.
a) u ist einfache Funktion heißt u ist von der Form

$$u(x) = \sum_{k=1}^{r} \alpha_k \mathbb{1}_{A_k}(x), \quad \alpha_k \geq 0,$$

mit $\mathcal{L}^d$–fast disjunkten, $\mathcal{L}^d$–messbaren Mengen A_k mit

$$\mathcal{L}^d(A_k) < \infty \quad \text{und} \quad \Omega = \bigcup_{k=1}^{r} A_k.$$

Nach Definition des Maßintegrals existiert eine Folge einfacher Funktionen (u_n) mit $u_n \leq u$ $\mathcal{L}^d$-f.ü. und

$$\int_\Omega |u_n|^p \, dx \nearrow \int_\Omega |u|^p \, dx, \quad n \to \infty$$

wobei wir ohne Einschränkung $u_n \leq u_{n+1}$ annehmen können (dies entspricht der Untersumme). Es folgt wegen $|u_n|^p \leq |u|^p$ $\mathcal{L}^d$-f.ü.

$$\int_\Omega ||u|^p - |u_n|^p| \, dx \xrightarrow{n} 0,$$

also $|u_n|^p \xrightarrow{n} |u|^p$ in $L^1(\Omega)$. Nach Teilfolgenwahl erhalten wir Konvergenz punktweise $\mathcal{L}^d$-f.ü. (vgl. Korollar 3.7) und daher $u_n \xrightarrow{n} u$ $\mathcal{L}^d$-f.ü. (wobei wir die Teilfolge wieder mit (u_n) bezeichnen). Nach dem Satz von Levi über monotone Konvergenz strebt dann bereits $u_n \xrightarrow{n} u$ in $L^p(\Omega)$.

b) Sei u einfach, d.h.

$$u(x) = \sum_{k=1}^{r} \alpha_k \mathbb{1}_{A_k}(x).$$

Wir wählen eine Folge $\eta_n^{A_k} \in C_0^\infty(A_k)$ mit

$$\eta_n^{A_k}(x) = \begin{cases} 1, & \textit{wenn } \operatorname{dist}(x, \partial\Omega) \geq 2/n \\ 0, & \textit{wenn } \operatorname{dist}(x, \partial\Omega) \leq 1/n \end{cases}$$

und $0 \leq \eta_n^{A_k} \leq 1$. Definieren wir $A_k^{2/n}$ als Menge der Punkte in A_k deren Abstand zum Rand von A_k kleiner oder gleich $2/n$ ist, folgt

$$\int_{A_k} |\mathbb{1}_{A_k} - \eta_n^{A_k}|^p \, dx = \int_{A_k^{2/n}} |\mathbb{1}_{A_k} - \eta_n^{A_k}|^p \, dx \leq \mathcal{L}^d(A_k^{2/n}) \xrightarrow{n} 0.$$

Nun definiert man $\eta_n \in C_0^\infty(\Omega)$ durch

$$\eta_n(x) = \sum_{k=1}^{r} \alpha_k \eta_n^{A_k}(x).$$

Es folgt

$$\int_\Omega |\eta_n - u|^p \, dx = \sum_{k=1}^{r} \alpha_k^p \int_\Omega |\mathbb{1}_{A_k} - \eta_n^{A_k}|^p \, dx \xrightarrow{n} 0,$$

also die gewünschte Konvergenz. □

Beweis von Satz 3.21.

Sei eine Funktion $u \in L^p(\Omega)$ vorgegeben. Wir schreiben $u = u^+ + u^-$, wobei $u^+ := \max(u, 0)$ und $u^- := \min(u, 0)$, und approximieren die Funktionen $u^\pm$ separat. (Man überlegt sich leicht, dass mit u auch die Funktionen $u^\pm$ zur Klasse $L^p(\Omega)$ gehören). Nach Lemma 3.8 a) existiert zu $n \in \mathbb{N}$ eine einfache Funktion u_n mit

$$\|u_n - u\|_p \leq \frac{1}{2n}.$$

Zu u_n finden wir nach Lemma 3.8 b) eine Funktion $\eta_n \in C_0^\infty$ mit

$$\|\eta_n - u_n\|_p \leq \frac{1}{2n}.$$

Insgesamt erhalten wir

$$\|\eta_n - u\|_p \leq \|\eta_n - u_n\|_p + \|\eta_n - u_n\|_p \leq \frac{1}{n},$$

d.h. $(\eta_n) \subset C_0^\infty$ *leistet das Gewünschte.* □

Zum Abschluss dieses § führen wir noch lokale sowie vektorielle Versionen der L^p–Räume ein, und beweisen das überaus wichtige Fundamentallemma der Variationsrechnung. Dieses wird uns in § 5 zu einem verallgemeinerten Ableitungsbegriff für L^p–Funktionen führen.

Definition 3.12 *Seien $\Omega \subset \mathbb{R}^d$ offen, und $1 \leq p \leq \infty$. Wir definieren $L^p_{loc}(\Omega)$ als Menge der $\mathcal{L}^d$–f. ü. auf Ω eindeutig definierten Funktionen $u : \Omega \to \overline{\mathbb{R}}$ mit*

$$u \in L^p(\omega) \quad \text{für alle} \quad \omega \Subset \Omega.$$

Zur Erinnerung: $\omega \Subset \Omega$ bedeutet, dass $\overline{\omega}$ kompakte Teilmenge von Ω ist. Als Beispiel betrachte man die Funktion $u(x) = 1/x$ mit $\Omega = (0,1)$. Dann gehört u zur Klasse $L^1_{loc}(\Omega)$ aber nicht zu $L^1(\Omega)$.
(Natürlich lassen sich analog auch lokale Versionen von $L^p(X;\mu)$ definieren, wenn (X,μ) ein Maßraum mit einem normierten (oder auch nur topologischen) Raum X ist.)
Für vektorwertige Funktionen $u := (u^1, \dots, u^D) : \Omega \to \mathbb{R}^D$ $(D \in \mathbb{N})$ definieren wir:

$$L^p(\Omega)^D := L^p(\Omega, \mathbb{R}^D) := \Big\{ u : \Omega \to \mathbb{R}^D;\ u^\nu \in L^p(\Omega) \text{ für alle } \nu = 1, \dots, D \Big\}$$

und versehen diesen Raum mit der Norm:

$$\|u\|_p := \|u\|_{\Omega;p} := \begin{cases} \left(\sum_{\nu=1}^{D} \left\|u^\nu\right\|_p^p \right)^{\frac{1}{p}} & ; \quad p < \infty \\ \max_{\nu=1}^{D} \left\|u^\nu\right\|_\infty & ; \quad p = \infty. \end{cases}$$

Selbstredend übertragen sich alle Eigenschaften der Räume $L^p(\Omega)$ auf die vektoriellen Versionen (Vollständigkeit, Hölder–Ungleichung, ... Dualräume). Schließlich kann man auch hier noch lokale Versionen $L^p_{loc}(\Omega)^D = L^p_{loc}(\Omega, \mathbb{R}^D)$ einführen. [h]

[h] In der Literatur finden sich auch die Bezeichnungen $L_p(\Omega)$, $L_p^{loc}(\Omega)$ sowie entsprechende Notationen für die vektoriellen Versionen.

Lemma 3.22 (Fundamentallemma der Variationsrechnung)
Sei $\Omega \subset \mathbb{R}^d$ offen und $u \in L^1_{loc}(\Omega)$ erfülle

$$\int_\Omega u(x)\eta(x)\,dx = 0 \quad \text{für alle} \quad \eta \in C^\infty_\circ(\Omega). \tag{3.10}$$

Dann ist $u \equiv 0$; genauer ist $u = 0$ f. ü. auf Ω. Ist dagegen

$$\int_\Omega u(x)\eta(x)\,dx \geq 0 \quad \text{für alle} \quad \eta \in C^\infty_\circ(\Omega), \eta \geq 0,$$

so ist $u \geq 0$ f. ü. auf Ω.

Beweis.
Wir zeigen nur den ersten Fall, den zweiten behandelt man analog. Sei $u \in L^1_{loc}(\Omega)$ eine Funktion, welche der Bedingung (3.10) genügt und $\omega \Subset \Omega$ beliebig. Dann gibt es nach Lemma 3.8 eine Folge $(\eta_n) \subset C^\infty_0(\omega)$ mit $0 \leq \eta_n \leq 1$ so dass

$$\eta_n \to \mathbb{1}_\omega \text{ in } L^p(\omega) \text{ bei } n \to \infty.$$

und nach Teilfolgenwahl (vgl. Korollar 3.7)

$$\eta_n \to \mathbb{1}_\omega \text{ punktweise bei } n \to \infty.$$

Mit dem Satz von Lebesgue über dominierte Konvergenz ($|u|$ ist die Majorante) und (3.10) ergibt sich

$$0 = \int_\Omega u(x)\eta_n(x)\,dx \xrightarrow{n} \int_\Omega u(x)\mathbb{1}_\omega(x)\,dx = \int_\omega u(x)\,dx. \tag{3.11}$$

Wir wählen einen Vertreter für u (ebenfalls wieder mit u bezeichnet) und betrachten die $\mathcal{L}^d$–messbaren Mengen

$$\Omega^\pm := \big\{x \in \Omega;\ u^\pm(x) = u(x)\big\}.$$

Sei $\mathcal{L}^d(\Omega^+) > 0$. Wählen wir $\omega \Subset \Omega^+$ mit $\mathcal{L}^d(\omega) > 0$, so erhalten wir wegen (3.11) und $u > 0$ auf ω, dass $\mathcal{L}^d(\omega) = 0$ sein muss; Widerspruch. Es folgt, da $\mathcal{L}^d$ ein Radon-Maß ist, [i]

$$\mathcal{L}^d(\Omega^+) = \sup_{\omega \Subset \Omega^+} \mathcal{L}^d(\omega) = 0$$

[i] Ein Maß $\mu : \mathcal{P}(X) \to [0,\infty]$ auf einem normierten (oder auch nur topologischen) Raum X, heißt Radon-Maß, falls jede Borel-Menge in X μ-messbar ist (μ ist ein Borel-Maß) und kompakte Teilmengen von X endliches Maß haben. In diesem Fall ist insbesondere jede μ-messbare Menge $A \subset X$ von innen dem Maße nach durch kompakte Teilmengen approximierbar, d. h. es ist $\mu(A) = \sup \mu(B); B \Subset A$.

und analog $\mathcal{L}^d(\Omega^-) = 0$*. Demnach ist* $u = 0$ *f.ü. auf* Ω*.* □

Wir erwähnen schließlich die folgende Variante des Fundamentallemmas, die als Lemma von Du Bois–Reymonds *bekannt ist.*

Lemma 3.23 (Du Bois–Reymond)
Seien $\Omega \subset \mathbb{R}^d$ *ein Gebiet und* $u \in L^1_{loc}(\Omega)$ *erfülle*

$$\int_\Omega u(x)\, \nabla\eta(x)\, dx = 0 \quad \text{für alle} \quad \eta \in C^\infty_\circ(\Omega).$$

Dann ist $u \equiv const$*; genauer ist* $u = const$ *f. ü. auf* Ω*.*

Bemerkung.
Für die erste Aussage in Lem. 3.22 sowie die Aussage von Lem. 3.23 gibt es Pendants für Funktionen $u \in L^p_{loc}(\Omega)^D$*.*

§ 4. Die Geometrie von Hilbert–Räumen und die Darstellungssätze von Riesz und Lax–Milgram

Wir wiederholen zunächst einige grundlegende Tatsachen über Skalarprodukte und Hilbert–Räume. Ziel dieses § sind die — insbesondere für die Theorie der partiellen Differentialgleichungen — wichtigen Darstellungssätze von Riesz und Lax–Milgram. Insbesondere wird sich zeigen, dass Hilbert–Räume selbstdual *sind. Wenn nichts anderes gesagt wird, bezeichnet im Folgenden X einen Vektorraum (kurz Raum) über dem Körper $\mathbb{R}$.*

Definition 4.1 (Skalarprodukt)
Ein Skalarprodukt (oder auch inneres Produkt) auf einem $\mathbb{R}$–Vektorraum X ist eine bilineare Abbildung $\langle\cdot\,,\cdot\rangle : X \times X \to \mathbb{R}$ mit folgenden Eigenschaften.

i) $\langle x, y\rangle = \langle y, x\rangle$ *für alle* $x,\, y \in X$ *(Symmetrie).*

ii) $\langle x, x\rangle > 0$ *für alle* $x \in X \setminus \{0\}$ *(positive Definitheit).*

Beispiel 4.2

i) Das kanonische Skalarprodukt auf dem $\mathbb{R}^d$ ist gegeben durch die Vorschrift

$$\langle x, y\rangle := x \cdot y := \sum_{k=1}^{d} x_k y_k,$$

wobei $x := (x_1, \dots, x_d), y := (y_1, \dots, y_d) \in \mathbb{R}^d$. Allgemeiner wird auf dem Raum $\mathbb{R}^{d\times D} \cong \mathbb{R}^{dD}$ $(D \in \mathbb{N})$ ein Skalarprodukt erklärt durch:

$$x : y := \sum_{k,l=1}^{d,D} x_k^l y_k^l := \sum_{k=1}^{d} \sum_{l=1}^{D} x_k^l y_k^l,$$

wobei $x := \left(x_k^l\right)_{\substack{k=1,\dots,d \\ l=1,\dots,D}}$, $y := \left(y_k^l\right)_{\substack{k=1,\dots,d \\ l=1,\dots,D}} \in \mathbb{R}^{dD}$. Für $D = 1$ ist natürlich $x : y = x \cdot y$.

ii) Ist (X, μ) ein Maßraum, so wird auf $L^2(X;\mu)$ ein Skalarprodukt erklärt durch

$$\langle u, w\rangle_2 := \langle u, w\rangle_{2;\,X} := T_w u = \int_X uw\, d\mu,$$

wobei $u, w \in L^2(X;\mu)$. Speziell ist für $u, w \in L^2(\Omega)^D$

$$\langle u, w\rangle_2 := \int_\Omega u(x) : w(x)\, dx.$$

Lemma 4.3
Sei $\langle\cdot\,,\cdot\rangle : X \times X \to \mathbb{R}$ ein Skalarprodukt auf dem $\mathbb{R}$–Vektorraum X, und sei $\|\cdot\| : X \to [0,\infty)$ gegeben durch $\|x\| := \sqrt{\langle x, x\rangle}$. Dann gilt für alle $x, y \in X$:

i) $\big|\langle x, y\rangle\big| \le \|x\|\|y\|$ *(Cauchy–Schwarz–Ungleichung).*

ii) $\|x+y\| \le \|x\| + \|y\|$ *(Dreiecksungleichung).*

iii) $\|x+y\|^2 + \|x-y\|^2 = 2\big(\|x\|^2 + \|y\|^2\big)$ *(Parallelogrammidentität).*

Insbesondere definiert $\|\cdot\|$ eine Norm auf X (die von $\langle\cdot\,,\cdot\rangle$ induzierte Norm*).*

Beweis.
Wir zeigen nur i) und ii). Teil iii) sei dem Leser als Übung überlassen.

i) Seien $x, y \in X$ linear unabhängig. Dann ist $\|y\| > 0$ und für alle $t \in \mathbb{R}$ ist

$$0 < \|x+ty\|^2 = \|x\|^2 + 2t\langle x, y\rangle + t^2\|y\|^2 =: p(t),$$

d. h. p ist ein quadratisches Polynom ohne Nullstellen. Die Diskriminante von p ist daher negativ, also $\|y\|^{-4}\langle x, y\rangle^2 - \|y\|^{-2}\|x\|^2 < 0$, und damit $\langle x, y\rangle^2 < \|x\|^2\|y\|^2$.

ii) Es ist $\|x+y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2$, wegen (i) also

$$\|x+y\|^2 \le \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = \big(\|x\| + \|y\|\big)^2.$$

In i) und ii) gilt genau dann die Gleichheit, wenn $x, y \in X$ linear abhängig sind. □

Bemerkung 4.4
Ist $\big(X, \|\cdot\|\big)$ ein beliebiger normierter Raum, so gibt es i. a. Punkte $x, y \in X$, für welche die Parallelogrammidentität (Lemma 4.3 iii)) falsch ist.
Man kann aber zeigen: Gilt in $\big(X, \|\cdot\|\big)$ die Formel aus Lemma 4.3 iii), so wird $\|\cdot\|$ von einem Skalarprodukt auf X induziert. Das Parallelogrammgesetz ist also notwendig und hinreichend dafür, dass ein normierter Raum ein Skalarprodukt besitzt.

Definition 4.5 (Hilbert–Raum)
Sei H ein $\mathbb{R}$–Vektorraum, versehen mit einem Skalarprodukt $\langle\cdot\,,\cdot\rangle : H\times H \to \mathbb{R}$. Dann heißt $\big(H,\langle\cdot\,,\cdot\rangle\big)$ ein Prä–Hilbert–Raum. Ist H bzgl. der von $\langle\cdot\,,\cdot\rangle$ induzierten Norm $\|\cdot\|$ (vgl. Lemma 4.3) vollständig, so heißt $\big(H,\langle\cdot\,,\cdot\rangle\big)$ ein Hilbert–Raum.

Beispiel 4.6

i) Ist (X,μ) ein Maßraum, so wird der Raum $L^2(X;\mu)$ vermöge dem Skalarprodukt aus Beispiel 4.2 ii) zu einem Hilbert–Raum. Speziell ist ℓ^2 (vgl. Bsp. 3.4) ein Hilbert–Raum vermöge dem Skalarprodukt

$$\langle x,y\rangle_2 = \sum_{n=1}^{\infty} x_n y_n,$$

wobei $x := (x_n), y := (y_n) \in \ell^2$.

ii) Sei $\Omega \subset \mathbb{R}^d$ offen und beschränkt. Dann wird durch

$$\langle u,w\rangle_{\mathcal{B}} := \int_\Omega u : w\,dx$$

ein Skalarprodukt auf $C^0_{\mathcal{B}}(\Omega)^D$ $(D \in \mathbb{N})$ erklärt. Jedoch ist $C^0_{\mathcal{B}}(\Omega)^D$ nicht vollständig bzgl. der von diesem Skalarprodukt induzierten Norm.

Definition 4.7 (Orthogonalkomplement)
Sei $\big(H,\langle\cdot\,,\cdot\rangle\big)$ ein Prä–Hilbert–Raum. Dann heißen $x,y\in H$ orthogonal, i. Z. $x\perp y$, falls $\langle x,y\rangle = 0$ ist. Für eine Teilmenge (nicht notwendig ein Unterraum) $U\subset H$ heißt

$$U^\perp := \big\{x\in H;\ x\perp u \text{ für alle } u\in U\big\}$$

das Orthogonalkomplement von U (mit der Konvention $\emptyset^\perp := H$).

Lemma 4.8
Sei H ein Prä–Hilbert–Raum und sei $U\subset H$ eine Teilmenge (nicht notwendig ein Unterraum). Dann ist $U^\perp$ ein abgeschlossener Unterraum von H.

Beweis.
Der Fall $U=\emptyset$ ist trivial. Sei also $U\neq\emptyset$. Seien $x,y\in U^\perp$ und sei $\lambda\in\mathbb{R}$. Dann ist $\langle x+\lambda y,u\rangle = \langle x,u\rangle + \lambda\langle y,u\rangle = 0$ für alle $u\in U$, also ist $x+\lambda y\in U^\perp$, d. h. $U^\perp$ ist Unterraum von H. Sei nun $(x_n)\subset U^\perp$ eine Folge mit $x_n \xrightarrow{n} x$ für ein $x\in H$. Für jedes $u\in U$ liefert die Cauchy–Schwarz–Ungleichung (Lemma 4.3 i)):

$$\langle u,x\rangle = \langle u,x-x_n\rangle + \langle u,x_n\rangle \le \|u\|\|x-x_n\| \xrightarrow{n} 0,$$

also $\langle u, x\rangle = 0$ *für alle* $u \in U$*. Mithin ist* $x \in U^\perp$*, also* $U^\perp$ *abgeschlossen.* □

Satz 4.9 (Abstraktes Variationsprinzip)
Sei $\big(H, \langle\cdot\,,\cdot\rangle\big)$ *ein Hilbert–Raum und sei* $\emptyset \neq K \subset H$ *abgeschlossen und konvex. Dann gibt es zu jedem* $x \in H$ *ein eindeutig bestimmtes Element* $\xi \in K$*, so dass*

$$\|x - \xi\| = \operatorname{dist}(x, K) := \inf_{z \in K} \|x - z\|$$

ist. Ferner ist ξ *die eindeutige Lösung von* $\langle x - \xi, y - \xi\rangle \leq 0$ *für alle* $y \in K$*. Daher heißt* ξ *auch der Lotfußpunkt von* x *auf* K*. Ferner ist offenbar* $\xi = x$*, falls* $x \in K$ *ist.*

Beweis.
Sei $(\xi_n) \subset K$ *eine* Minimalfolge*, d. h. es gelte*

$$\|x - \xi_n\| \xrightarrow{n} \operatorname{dist}(x, K).$$

Wir zeigen, dass (ξ_n) *konvergiert, indem wir uns überlegen, dass es sich um eine Cauchy–Folge in* H *handelt. Dazu betrachten wir* $\eta^\pm := \frac{1}{2}\big(\xi_n \pm \xi_m\big)$ *für* $n, m \in \mathbb{N}$*. Da* K *konvex ist, ist* $\eta^\pm \in K$*, und wir erhalten mittels der Parallelogrammidentität, (Lemma 4.3 iii)):*

$$\begin{aligned}
\|x - \xi_n\|^2 + \|x - \xi_m\|^2 &= \|x - \eta^+ - \eta^-\|^2 + \|x - \eta^+ + \eta^-\|^2 \\
&= 2\|x - \eta^+\|^2 + 2\|\eta^-\|^2 = 2\|x - \eta^+\|^2 + \frac{1}{2}\|\xi_n - \xi_m\|^2 \\
&\geq 2\operatorname{dist}(x, K)^2 + \frac{1}{2}\|\xi_n - \xi_m\|^2.
\end{aligned}$$

Demnach wird

$$\|\xi_n - \xi_m\|^2 \leq 2\|x - \xi_n\|^2 + 2\|x - \xi_m\|^2 - 4\operatorname{dist}(x, K)^2 \xrightarrow{n,m} 0,$$

womit (ξ_n) *also eine Cauchy–Folge in* H *ist. Wegen der Vollständigkeit von* H *existiert ein* $\xi \in H$ *mit* $\xi = \lim_n \xi_n$*, und da* K *abgeschlossen ist, ist bereits* $\xi \in K$*. Da nach Definition von* ξ

$$\|x - \xi\| = \lim_n \|x - \xi_n\| = \operatorname{dist}(x, K)$$

ist, ist ξ *ein abstandsminimales Element von* K*.*
Bleibt nachzuweisen, dass ξ *mit dieser Eigenschaft eindeutig bestimmt ist. Dazu überlegen wir uns, dass*

$$\langle x - \xi, y - \xi\rangle \leq 0 \quad \textit{für alle} \quad y \in K \tag{4.1}$$

gilt. Sei also $y \in K$ *fixiert. Da* K *konvex ist, ist* $\xi + \theta(y-\xi) \in K$ *für alle* $\theta \in (0,1)$ *und*

$$\begin{aligned}\|x-\xi\|^2 = \operatorname{dist}(x,K)^2 &\leq \big\|x - \big(\xi + \theta(y-\xi)\big)\big\|^2 \\ &= \|x-\xi\|^2 - 2\theta\big\langle x-\xi, y-\xi\big\rangle + \theta^2\|y-\xi\|^2,\end{aligned}$$

woraus nach Division durch θ *und mit* $\theta \downarrow 0$ *die Beziehung* (4.1) *folgt. Nehmen wir an es gebe zwei Lösungen* ξ_1, ξ_2 *von 4.1, dann folgt*

$$\langle x-\xi_1, \xi_2-\xi_1\rangle \leq 0, \quad \langle x-\xi_2, \xi_1-\xi_2\rangle \leq 0.$$

Da die zweite Ungleichung äquivalent zu $\langle \xi_2 - x, \xi_2 - \xi_1\rangle \leq 0$ *ist, folgt durch Addition der Ungleichungen*

$$\langle x_2-\xi_1, \xi_2-\xi_1\rangle \leq 0$$

und damit die Eindeutigkeit. □

Eine wichtige theoretische Anwendung von Satz 4.9 ist:

Satz 4.10 (Orthogonalzerlegung, Projektionssatz)
Sei H *ein Hilbert–Raum und* $U \subset H$ *ein abgeschlossener Unterraum. Dann ist*

$$H = U \oplus U^\perp,$$

d. h. jeder Vektor $x \in H$ *hat eine eindeutige Zerlegung der Form* $x = u + u^\perp$ *mit* $u \in U$ *und* $u^\perp \in U^\perp$*. Die orthogonale Projektion* $\pi_U : H \to U$, $x \mapsto u$ *auf* U *ist linear, stetig und idempotent, d. h. es ist* $\pi_U^2 := \pi_U \circ \pi_U = \pi_U$*. Ferner ist* $\|\pi_U\| = 1$*, falls* $U \neq \{0\}$ *ist.*

Beweis.
Wir wählen in Satz 4.9 $K := U$*. Dann existiert zu jedem* $x \in H$ *ein abstandsminimaler Vektor* $u \in U$*, d. h. es ist* $\|x-u\| = \operatorname{dist}(x,U)$ *und* u *erfüllt*

$$\langle x-u, v-u\rangle \leq 0 \quad \textit{für alle} \quad v \in U.$$

Da U *ein Unterraum von* H *ist, durchläuft mit* v *auch* $v-u$ *ganz* U*, also ist* $\langle x-u, w\rangle \leq 0$ *sowie* $\langle x-u, -w\rangle \leq 0$ *für alle* $w \in U$*, und daher* $\langle x-u, w\rangle = 0$ *für alle* $w \in U$ *oder also* $x-u \in U^\perp$*. Die Aussagen über die orthogonale Projektion folgen unmittelbar, und seien dem Leser zur Übung überlassen.*

□

Korollar 4.11

Sei H ein Hilbert–Raum und $U \subset H$ ein (nicht notwendig abgeschlossener) Unterraum. Dann ist

$$U^{\perp\perp} := \left(U^{\perp}\right)^{\perp} = \overline{U}.$$

Beweis.
Wegen Lemma 4.8 ist $U^{\perp} = \overline{U}^{\perp}$ und nach Satz 4.10 ist

$$H = \overline{U} \oplus \overline{U}^{\perp} \quad \textit{sowie} \quad H = U^{\perp} \oplus U^{\perp\perp} = \overline{U}^{\perp} \oplus U^{\perp\perp},$$

und die Behauptung folgt sofort. □

Wir kommen nun zu dem entscheidenden Satz dieses §, dem Darstellungsatz von Riesz. *Dieser charakterisiert die stetigen linearen Abbildungen von einem Hilbert–Raum H nach $\mathbb{R}$ (und damit den Dualraum H^*) vollständig.*
Wir geben zunächst eine aus der Theorie der partiellen Differentialgleichungen stammende Motivation für diesen Satz (sog. Hilbert–Raum–Methode *zur Lösung partieller Differentialgleichungen): Sei $\Omega \subset \mathbb{R}^d$ offen, beschränkt und mit genügend glattem Rand. Weiter sei $f \in C^0(\overline{\Omega})$ eine vorgegebene Funktion. Wir betrachten das Dirichlet–Problem zur Poisson–Gleichung: Gesucht ist eine Lösung $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$ des Randwertproblems:*

$$\left.\begin{aligned} \Delta u &= f \quad \textit{auf} \quad \Omega \\ u &= 0 \quad \textit{auf} \quad \partial\Omega. \end{aligned}\right\} \tag{4.2}$$

Für eine solche Funktion u erhält man mit dem Divergenzsatz von Gauß die Beziehung

$$\int_\Omega \nabla u \cdot \nabla \eta \, dx = -\int_\Omega f\eta \, dx \quad \textit{für alle} \quad \eta \in C^\infty_\circ(\Omega). \tag{4.3}$$

Sei nun $\mathcal{W}(\Omega)$ ein Hilbert–Raum von Funktionen $u : \Omega \to \mathbb{R}$ mit folgenden Eigenschaften:

$$\left.\begin{array}{l} C^\infty_\circ(\Omega) \underset{\neq}{\subsetneq} \mathcal{W}(\Omega) \underset{\neq}{\subsetneq} L^2(\Omega); \\[1ex] „u = 0 \textit{ auf } \partial\Omega\text{“ } \textit{für alle } u \in \mathcal{W}(\Omega); \\[1ex] \textit{jede Funktion } u \in \mathcal{W}(\Omega) \textit{ besitzt eine} \\ \text{verallgemeinerte } \textit{Ableitung } \nabla u \in L^2(\Omega, \mathbb{R}^n). \end{array}\right\} \tag{4.4}$$

Ein Skalarprodukt auf $\mathcal{W}(\Omega)$ ist gegeben durch $\langle\cdot,\cdot\rangle_{\mathcal{W}} : \mathcal{W}(\Omega) \times \mathcal{W}(\Omega) \to \mathbb{R}$,

$$\langle u, w\rangle_{\mathcal{W}} := \int_\Omega \nabla u \cdot \nabla w \, dx, \tag{4.5}$$

wobei die rechte Seite in (4.5) *wegen der geforderten Eigenschaften von* $\mathcal{W}(\Omega)$ *wohldefiniert ist.*[a] *Die Beziehung* (4.3) *macht unter diesen Bedingungen Sinn und liest sich als*

$$\langle u, \eta \rangle_{\mathcal{W}} = T\eta := -\int_{\Omega} f\eta \, dx \quad \text{für alle} \quad \eta \in \mathcal{W}(\Omega), \tag{4.6}$$

wobei T*, aufgefasst als Abbildung auf* $\mathcal{W}(\Omega) \supset C_{\circ}^{\infty}(\Omega)$*, offenbar linear und stetig ist, d. h. es ist* $T \in \mathcal{W}(\Omega)^*$*. Das Problem eine Lösung* $u \in \mathcal{W}(\Omega)$ *für* (4.6) *zu finden, ist demnach gleichbedeutend mit dem Problem zu dem Funktional* $T \in \mathcal{W}(\Omega)^*$ *eine Funktion* $u \in \mathcal{W}(\Omega)$ *zu finden, so dass* $\langle u, \cdot \rangle_{\mathcal{W}} = T$ *auf* $\mathcal{W}(\Omega)$ *ist; diese Funktion* u *erfüllt dann insbesondere* (4.6) *(und damit* (4.3)*). Man nennt dann* u *eine* schwache *Lösung des Ausgangsproblems* (4.2)*, weil* u *a priori noch nicht einmal stetig ist. Genau die Existenz einer solchen Funktion garantiert der folgende*

Satz 4.12 (Darstellungssatz von Riesz)
Sei $\big(H, \langle \cdot, \cdot \rangle\big)$ *ein Hilbert–Raum und sei* $T \in H^*$*. Dann gibt es genau einen Vektor* $u \in H$ *mit*

$$\langle u, \cdot \rangle = T \text{ auf } H \quad \text{und} \quad \|u\| = \|T\|_{\infty}.$$

Die Abbildung $H \to H^*$*,* $x \mapsto \langle x, \cdot \rangle$ *ist also ein isometrischer Isomorphismus.*

Beweis.
Sei $x \in H$ *beliebig und sei* $S : H \to \mathbb{R}$, $y \mapsto \langle x, y \rangle$*. Mit der Ungleichung von Cauchy–Schwarz (Lemma 4.3 i)) wird* $\|S\|_{\infty} \le \|x\|$ *für alle* $x \in H$*. Für* $x \ne 0$ *ist andererseits* $S\big(\frac{x}{\|x\|}\big) = \|x\|$*, insgesamt also* $\|S\|_{\infty} = \|x\|$ *für alle* $x \in H$*. Dies zeigt, dass die Abbildung* $H \to H^*$, $x \mapsto \langle x, \cdot \rangle$ *eine isometrische Injektion ist. Bleibt die Surjektivität zu zeigen: Sei dazu* $T \in H^* \setminus \{0\}$ *und* $U := \ker T = T^{-1}\big(\{0\}\big)$*. Da* T *stetig ist, ist* U *abgeschlossen und nach Satz 4.10 ist* $H = U \oplus U^{\perp}$*.*
Wir zeigen nun, dass $\dim U^{\perp} = 1$ *ist. Wegen* $T \ne 0$ *ist offenbar* $\dim U^{\perp} \ge 1$*. Seien* $u_0 \in U^{\perp} \setminus \{0\}$ *und* $w \in U^{\perp}$ *beliebig. Dann ist*

$$T\left(w - \frac{Tw}{Tu_0}\, u_0\right) = 0 \quad \Longleftrightarrow \quad w - \frac{Tw}{Tu_0}\, u_0 \in U \cap U^{\perp},$$

und daher $w = \frac{Tw}{Tu_0} u_0$*, da* $U \cap U^{\perp} = \{0\}$ *ist. Mithin ist* $U^{\perp} = \operatorname{span}\{u_0\}$*, also* $\dim U^{\perp} = 1$.

[a] In der Tat existiert ein solcher Hilbert–Raum, wie wir im Kapitel Sobolev-Räume sehen werden. Dieser Raum wird üblicherweise mit $\mathring{W}^{1,2}(\Omega)$ bezeichnet wird. Die Bedingung „$u = 0$ auf $\partial\Omega$“ oben, ist in einem verallgemeinerten Sinn zu verstehen.

Sei nun $x \in H$*. Wegen* $H = U \oplus \operatorname{span}\{u_0\}$ *existieren ein* $v \in U$ *und ein* $\lambda \in \mathbb{R}$ *mit* $x = v + \lambda u_0$ *und es ist* $Tx = \lambda T u_0$*. Wegen*

$$\langle u_0, x\rangle = \langle u_0, v\rangle + \lambda \|u_0\|^2 = \lambda \|u_0\|^2$$

folgt $\lambda = \|u_0\|^{-2}\langle u_0, x\rangle$*. Damit haben wir* $Tx = \big\langle u, x\big\rangle$ *für alle* $x \in H$ *gezeigt, wobei* $u := \frac{Tu_0}{\|u_0\|^2} u_0$ *ist.* □

Bemerkung 4.13

Wegen Satz 4.12 heißen Hilbert–Räume auch selbstdual*, weil* $H \cong H^*$ *gilt. Speziell für* $H := L^2(X;\mu)$ *bedeutet Satz 4.12, dass jedes stetige, lineare Funktional* $T : L^2(X;\mu) \to \mathbb{R}$ *die Form* $Tw = \int_X uw\, d\mu = T_u w$ *hat, mit einer eindeutig bestimmten Funktion* $u \in L^2(X;\mu)$ *(vgl. § 3). Damit ist ein anderer Beweis für die Aussage* $L^2 \cong L^{2^*}$ *erbracht, der ganz ohne Maßtheorie auskommt.*

Zum Abschluss dieses § wollen wir noch eine Verallgemeinerung des Darstellungssatzes von Riesz zeigen, der als Darstellungssatz von Lax–Milgram *bekannt ist. Dieser sagt im wesentlichen aus, dass sich die Elemente* $T \in H^*$ *auch durch Bilinearformen* $\beta : H \times H \to \mathbb{R}$ *auf dem Hilbert–Raum* H *beschreiben lassen; genauer:*

$$\beta(u,\cdot) = T \quad \textit{für ein} \quad u \in H.$$

Zur Motivation betrachten wir eine Verallgemeinerung des Anfangswertproblems (4.2)*: Sei* $\Omega \subset \mathbb{R}^d$ *offen, beschränkt und mit genügend glattem Rand. Weiter sei* $f \in C^0(\overline{\Omega})$ *eine vorgegebene Funktion. Gesucht sei eine Funktion* $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$*, so dass gilt:*

$$\left.\begin{aligned} \operatorname{div}(A\nabla u) &= f \quad \textit{auf} \quad \Omega; \\ u &= 0 \quad \textit{auf} \quad \partial\Omega \end{aligned}\right\} \tag{4.7}$$

(partielle Differentialgleichung in Divergenzform*), wobei die Koeffizienten* $A := (a_{rs}) \in \mathbb{R}^{d\times d}$ *der* Elliptizitätsbedingung

$$A(P,P) := AP\cdot P = \sum_{r,s=1}^{d} a_{rs}\, P_r P_s > 0 \quad \textit{für alle} \quad P := (P_r) \in \mathbb{R}^d \setminus \{0\}$$

genügen mögen. (Man beachte, dass A *keine symmetrische Matrix sein muss.) Sei* $\mathcal{W}(\Omega)$ *der Banach–Raum mit den Eigenschaften aus* (4.4)*, versehen mit dem Skalarprodukt* $\langle\cdot,\cdot\rangle_{\mathcal{W}}$ *gem.* (4.5)*. Für eine Lösung* u *von* (4.7) *erhält man mit dem Divergenzsatz von Gauß die Beziehung (man beachte die augenfällige*

Analogie zu (4.3)):

$$\int_\Omega A(\nabla u, \nabla \eta)\, dx = -\int_\Omega f\eta\, dx =: T\eta \quad \text{für alle} \quad \eta \in C_\circ^\infty(\Omega). \tag{4.8}$$

Wie zuvor betrachten wir nun T als stetige, lineare Abbildung auf $\mathcal{W}(\Omega)$. Aufgrund der Eigenschaften (4.4) des Hilbert–Raumes $\mathcal{W}(\Omega)$ überlegt man sich sofort, dass durch die linke Seite von (4.8) eine (nicht symmetrische) Bilinearform $\beta_\mathcal{W}$ auf $\mathcal{W}(\Omega)$ induziert wird. Damit liest sich dann (4.8) als:

$$\beta_\mathcal{W}(u,\eta) := \int_\Omega A(\nabla u, \nabla \eta)\, dx = T\eta \quad \text{für alle} \quad \eta \in \mathcal{W}(\Omega). \tag{4.9}$$

Der nachfolgende Satz von Lax–Milgram sichert die Existenz einer Funktion $u \in \mathcal{W}(\Omega)$, so dass $\beta_\mathcal{W}(u,\cdot) = T$ auf $\mathcal{W}(\Omega)$ ist; u ist demnach eine schwache Lösung von (4.7). Dafür muss man allerdings noch wissen, dass die Bilinearform $\beta_\mathcal{W}$ stetig *und* koerziv *ist. Was das bedeuten soll, wollen wir nun allgemein fassen. (Wir überlassen es dem Leser als Übung, zu entscheiden ob bzw. unter welchen Voraussetzungen die obige Bilinearform $\beta_\mathcal{W}$ tatsächlich stetig und koerziv ist.)*

Definition 4.14 (Stetige/koerzive Bilinearform)
Sei $\big(X, \|\cdot\|\big)$ ein normierter Raum, und sei $\beta : X \times X \to \mathbb{R}$ eine Bilinearform auf X. Dann nennt man β

i) stetig (oder auch beschränkt), falls die Größe

$$\|\beta\|_\infty := \sup_{(x,y)\in\partial\mathbb{B}\times\partial\mathbb{B}} \big|\beta(x,y)\big|$$

endlich ist;

ii) koerziv (oder elliptisch), falls $\inf_{x\in\partial\mathbb{B}} \beta(x,x) > 0$ ist.

Wir bezeichnen mit $\mathscr{B}(X)$ den Raum aller stetigen und koerziven Bilinearformen auf X.

Bemerkung 4.15

i) Ist eine Bilinearform β stetig gem. obiger Definition, so ist β tatsächlich im üblichen Sinn stetig auf $X \times X$, sofern man $X \times X$ mit einer passenden Norm versieht. Beispielsweise werden durch $\big\|(x,y)\big\| := \|x\| + \|y\|$ oder $\big\|(x,y)\big\| := \sqrt{\|x\|^2 + \|y\|^2}$ für $(x,y) \in X \times X$ solche Normen auf $X \times X$ erklärt. (Natürlich gilt dies auch für jede zu einer dieser Normen äquivalenten Norm.)

ii) *Eine Bilinearform β über einem Raum X ist genau dann stetig und koerziv, wenn gilt:*

$$\bigl|\beta(x,y)\bigr| \le M\|x\|\|y\| \quad \textit{und} \quad \beta(x,x) \ge m\|x\|^2 \quad \textit{für alle} \quad x,y \in X$$

mit Konstanten $0 < m < M < \infty$.

Satz 4.16 (Darstellungssatz von Lax–Milgram)
Sei $\bigl(H, \langle\cdot,\cdot\rangle\bigr)$ ein Hilbert–Raum, und sei $\beta \in \mathscr{B}(H)$. Dann gibt es zu jedem $T \in H^*$ ein eindeutig bestimmtes $u \in H$, so dass

$$\beta(u,\cdot) = T \ \textit{auf } H \quad (\textit{bzw. } \beta(\cdot,u) = T \ \textit{auf } H)$$

ist. Die Abbildung $H \to H^*$, $x \mapsto \beta(x,\cdot)$ (bzw. $x \mapsto \beta(\cdot,x)$) ist ein Homöomorphismus.

Beweis.
Für jedes $x \in H$ ist $\beta(x,\cdot) \in H^*$ und wegen Bemerkung 4.15 i) folgt die Stetigkeit von $H \ni x \mapsto \beta(x,\cdot)$. Nach dem Satz von Riesz 4.12 existiert zu jedem $x \in H$ genau ein $u_x \in H$ mit $\langle u_x,\cdot\rangle = \beta(x,\cdot)$. Für $S : H \to H$, $x \mapsto u_x$ ist also

$$\langle Sx, y\rangle = \beta(x,y) \quad \textit{für alle} \quad y \in H$$

sowie

$$\|Sx\| = \bigl\|\beta(x,\cdot)\bigr\| = \sup_{y\in\mathbb{B}} \bigl|\beta(x,y)\bigr| \le M\|x\|,$$

wobei M gem. Bem. 4.15 definiert ist. Da S ersichtlich linear ist, ist also $S \in \mathscr{L}(H,H)$ mit $\|S\|_\infty \le M$. Wir zeigen nun, dass S bijektiv ist:

i) *Wir zeigen zunächst (mit m wie in Bem. 4.15)*

$$\|Sx\| \ge m\|x\| \quad \textit{für alle} \quad x \in H, \tag{4.10}$$

also folgt aus $Sx = 0$, dass $x = 0$ ist, was $\ker S = \{0\}$ bedeutet. Mithin ist S injektiv. Die Umkehrabbildung $S^{-1} : S(H) \to H$ existiert somit und ist wegen (4.10) *stetig mit $\|S^{-1}\|_\infty \le \frac{1}{m}$. Dabei erhält man (4.10) durch*

$$\begin{aligned}\|Sx\|\,\|x\| &\ge \langle Sx, x\rangle = \langle u_x, x\rangle \\ &= \beta(x,x) \ge m\,\|x\|^2 .\end{aligned}$$

ii) *$S(H)$ ist abgeschlossener Unterraum von H: Sei dazu $(Sx_n) \subset S(H) \subset H$ eine Folge mit $Sx_n \xrightarrow{n} y \in H$. Also ist (Sx_n) eine Cauchy–Folge in H*

und mit (4.10) *gilt:*

$$\|x_k - x_l\| \leq \frac{1}{m}\,\|Sx_k - Sx_l\| \xrightarrow{k,l} 0,$$

weshalb auch (x_n) *eine Cauchy–Folge in* H *sein muss, d. h. es existiert ein* $x \in H$ *mit* $x_n \xrightarrow{n} x$*, und wegen der Stetigkeit von* S *folgt* $Sx = y$*.*

iii) $S(H) = H$*: Nach ii) ist* $H = S(H) \oplus S(H)^\perp$*. Angenommen, es ist* $S(H) \subsetneqq H$*. Dann existiert ein* $0 \neq y \in H$ *mit* $y \in S(H)^\perp$ *und es ist* $\langle Sx, y\rangle = 0$ *für alle* $x \in H$*. Speziell ist also wegen* (4.10)

$$0 = \langle Sy, y\rangle = \beta(y,y) \geq m\|y\|^2 > 0,$$

da $y \neq 0$ *ist; ein Widerspruch.*

Nach dem Satz von Riesz 4.12 gibt es für jedes $T \in H^*$ *genau ein* $w \in H$ *mit* $T = \langle w, \cdot\rangle$*. Da* S *gem. obigen Überlegungen bijektiv ist, gibt es andererseits genau ein* $u \in H$ *mit* $Su = w$*. Daher ist also* $T = \langle w, \cdot\rangle = \langle Su, \cdot\rangle = \beta(u, \cdot)$*.* $\square$

§ 5. Schwache Konvergenz

Ist X ein endlich–dimensionaler, normierter Raum, so besitzt jede beschränkte Folge in X eine in X konvergente Teilfolge. Diese Aussage ist als Auswahlprinzip von Bolzano–Weierstraß *bekannt; man sagt auch der Raum X sei* (folgen–) kompakt [a]*. Wie das nachfolgende (einfache) Beispiel zeigt, ist diese Aussage in unendlich–dimensionalen Räumen i. a. nicht mehr richtig.*
Zur Motivation der Begriffsbildung „schwache Konvergenz“ betrachten wir eine abstrakte Variationsaufgabe: Sei $\big(X, \|\cdot\|\big)$ ein normierter Funktionenraum und sei $\emptyset \neq \mathscr{C} \subset X$ eine gegebene Teilmenge (in der konkreten Problemstellung wird diese Teilklasse durch Vorgabe von Randwerten oder durch andere Nebenbedingungen festgelegt). Wir betrachten ein Funktional [b] $T : X \to \mathbb{R}$*, welches in der Klasse $\mathscr{C}$ minimiert werden soll. Gesucht ist also eine Funktion $u \in \mathscr{C}$ mit*

$$T[u] = \inf_{w \in \mathscr{C}} T[w] =: \tau.$$

Dabei verlangt man, dass T koerziv *ist, d. h. es ist*

$$T[w] \geq \alpha\|w\| + \beta \quad \text{für alle} \quad w \in X \tag{5.1}$$

mit Konstanten $\alpha, \beta \in \mathbb{R}$, $\alpha > 0$. Wegen (5.1) *ist $T[w] \geq \beta$ für alle $w \in \mathscr{C}$, also insbesondere $\tau \in \mathbb{R}$. Nach Definition des Infimums existiert daher eine* Minimalfolge $(u_m) \subset \mathscr{C}$ *für T, d. h es gilt:*

$$\lim_m T[u_m] = \tau$$

und aus (5.1) *folgt unmittelbar die* Normbeschränktheit *dieser Folge:*

$$\sup_m \|u_m\| < \infty.$$

[a] In metrischen (und damit auch in normierten) Räumen ist „folgenkompakt“ gleichbedeutend mit „überdeckungskompakt“. In allgemeinen topologischen Räumen ist dies i. a. falsch! Die Topologie des Raums muss dafür wenigstens eine *abzählbare Basis* besitzen. Dies ist insbesondere in sog. *separablen* Räumen der Fall (s. u.).

[b] Damit sei hier nicht ein lineares Funktional (linearer Operator) gemeint, sondern eben irgendeine „Funktion“, deren Argumente Funktionen aus X sind; allgemein hat ein Funktional Werte in $\overline{\mathbb{R}}$.

Existiert nun eine in X konvergente Teilfolge $\left(u_{m_k}\right)_k \subset (u_m)$ mit $u_{m_k} \xrightarrow{k}: u \in X$, so prüfe man, ob die Grenzfunktion u wieder zur Klasse $\mathscr{C}$ gehört und ob ferner

$$T[u] = \lim_k T[u_{m_k}] \tag{5.2}$$

gilt. In diesem Fall ist dann offenbar u eine Lösung des Variationsproblems. Das Problem besteht nun darin, dass in einem Funktionenraum (der unendlich-dimensional ist) i. a. keine konvergente Teilfolge der Minimalfolge (u_m) existieren muss. Unsere Forderungen sind also offenbar zu stark. Daher benötigen wir eine Abschwächung des Konvergenzbegriffs: Wir werden noch sehen, dass in gewissen Räumen diese schwache Konvergenz *unter gewissen Voraussetzungen die Auswahl einer konvergenten Teilfolge erlaubt. Hat man eine solche Teilfolge $\left(u_{m_k}\right)_k \subset (u_m)$ gefunden, deren (schwache) Grenzfunktion u zur Klasse $\mathscr{C}$ gehört, kann man aber in der obigen Situation nicht mehr erwarten, dass* (5.2) *gilt. Um zu erreichen, dass u tatsächlich eine (schwache) Lösung des Variationsproblems ist, genügt es aber, dass*

$$T[u] \leq \liminf_k T[u_{m_k}] \tag{5.3}$$

ist. Letzere Eigenschaft nennt man (schwache) Unterhalbstetigkeit *(s.* [c]

Definition 5.2 (Schwache Konvergenz)
Sei $\left(X, \|\cdot\|\right)$ ein normierter Raum. Eine Folge $(x_n) \subset X$ heißt schwach konvergent gegen ein $x \in X$, falls gilt

$$\varphi(x_n) \xrightarrow{n} \varphi(x) \quad \textit{für alle} \quad \varphi \in X^*.$$

In diesem Fall schreiben wir $x_n \overset{n}{\rightharpoonup} x$.

Das schwache *Analogon zur Bolzano–Weierstraß–Eigenschaft in unendlich–dimensionalen normierten Räumen $\left(X, \|\cdot\|\right)$ lautet dann:*

$$(x_n) \subset X,\ \sup_n \|x_n\| < \infty \quad \Longrightarrow \begin{cases} x_{n_k} \overset{k}{\rightharpoonup}: x \in X \\ \textit{für eine Teilfolge } \left(x_{n_k}\right)_k. \end{cases} \tag{5.4}$$

Einen Raum X, in dem (5.4) *gilt nennt man auch* schwach (folgen–) kompakt. *Das* schwache Auswahlprinzip (5.4) *gilt jedoch keineswegs in jedem unendlich–dimensionalen, normierten Raum, wie wir weiter unten noch sehen werden. Die Räume, in denen dieses Prinzip Gültigkeit hat, sind genau die sog.* reflexiven *Räume, die wir gleich definieren werden. Da die schwache Konvergenz*

[c] Natürlicher wäre es hier schwache Stetigkeit zu verlangen, jedoch sind bereits sehr einfache Funktionale nicht schwach stetig.

durch die Konvergenz stetiger linearer Operatoren in X^* (vgl. § 3) erklärt ist, schicken wir noch einige allgemeine Tatsachen über Folgen linearer, stetiger Operatoren vorweg. Der folgende Satz, der auf Banach und Steinhaus zurückgeht, liefert ein Prinzip für die gleichmäßige Beschränktheit von Folgen stetiger linearer Operatoren auf Banach–Räumen, welches besagt: Aus der punktweisen Beschränktheit einer Folge stetiger linearer Operatoren, folgt bereits die gleichmäßige Beschränktheit dieser Folge.

Satz 5.3 (Banach–Steinhaus)
Seien X ein Banach–Raum, Y ein normierter Raum und $(T_n) \subset \mathscr{L}(X,Y)$ eine Folge stetiger linearer Operatoren derart, zu jedem $x \in X$ eine positive Konstante $c = c(x)$ existiert mit $\sup_n \|T_n x\| \leq c$. Dann ist die Folge (T_n) beschränkt (bzgl. der Operator–Norm), d. h. es ist

$$\sup_n \|T_n\|_\infty = \sup_n \sup_{x\in\overline{\mathbb{B}}} \|T_n x\| < \infty.$$

Beweis.
Wir nehmen an, dass für jede offene Kugel $B \subset X$

$$\sup_n \sup_{x\in\partial B} \|T_n x\| = \sup_n \sup_{x\in\partial B} \|T_n x\| = \infty \tag{5.5}$$

ist. Demnach gibt es ein $n_1 \in \mathbb{N}$ und ein $x_1 \in \mathbb{B}$ mit $\|T_{n_1} x_1\| > 1$. Da T_{n_1} stetig in x_1 ist, existiert eine Kugel $B_{r_1}(x_1) \Subset \mathbb{B}$ um x_1, so dass $\|T_{n_1} x\| \geq 1$ für alle $x \in \overline{B}_{r_1}(x_1)$ ist. Nach evtl. Vekleinern von $B_{r_1}(x_1)$ können wir ferner annehmen, dass

$$r_1 = \operatorname{diam} B_{r_1}(x_1) \leq \frac{1}{2} \operatorname{diam} \mathbb{B}$$

ist. Wieder mit (5.5) finden wir ein $n_2 > n_1$ und ein $x_2 \in B_{r_1}(x_1)$ mit $\|T_{n_2} x_2\| > 2$, und da T_{n_2} stetig ist in x_2, gibt es eine Kugel $B_{r_2}(x_2) \Subset B_{r_1}(x_1)$ um x_2 mit

$$\begin{aligned} &\|T_{n_2} x\| \geq 2 \quad \textit{für alle} \quad x \in \overline{B}_{r_2}(x_2); \\ &r_2 \leq \frac{1}{2} \operatorname{diam} r_1 \leq \frac{1}{4} \operatorname{diam} \mathbb{B}. \end{aligned}$$

Rekursiv so fortfahrend finden wir eine Folge (B_k) von Kugeln mit $B_{k+1} \Subset B_k \Subset \mathbb{B}$ für alle $k \in \mathbb{N}$ sowie eine Folge (n_k) von Indizes, so dass für jedes $k \in \mathbb{N}$ gilt

$$\|T_{n_k} x\| \geq k \quad \textit{für alle} \quad x \in \overline{B}_k;$$

$$\operatorname{diam} B_k \leq \frac{1}{2^k} \operatorname{diam} \mathbb{B}.$$

Da X ein Banach–Raum ist, gibt es nach dem Durchschnittsatz von Cantor (Schachtelungsprinzip) genau ein $z \in X$ mit

$$\bigcap_{k \in \mathbb{N}} \overline{B}_k = \{z\}.$$

Folglich ist $\|T_{n_k} z\| \geq k$ für alle $k \in \mathbb{N}$, im Widerspruch zur Voraussetzung des Satzes. Somit muss es eine Kugel um z geben, für die (5.5) nicht gilt. Daraus ergibt sich nun leicht die Behauptung. □

Aus dem Satz von Banach–Steinhaus ergibt sich unmittelbar das folgende Korollar, das die Sonderstellung stetiger linearer Operatoren zeigt: Die Stetigkeit bleibt bereits unter punktweiser Konvergenz erhalten.

Korollar 5.4 (Abgeschlossenheit von $\mathscr{L}(X,Y)$)
Seien X ein Banach–Raum, Y ein normierter Raum und $(T_n) \subset \mathscr{L}(X,Y)$ eine Folge mit $T_n(x) \xrightarrow{n} : T(x)$ für alle $x \in X$. Dann ist $T \in \mathscr{L}(X,Y)$.

Wir ziehen nun die für unsere Zwecke wichtigen Schlüsse, die schwache Konvergenz betreffend. Die Voraussetzung der folgenden Aussage ist insbesondere erfüllt, wenn die gegebene Folge $(x_n) \subset X$ schwach konvergent in X ist.

Korollar 5.5
Seien X ein normierter Raum und $(x_n) \subset X$ eine Folge mit

$$\big|\varphi(x_n)\big| \leq c \quad \textit{für alle} \quad n \in \mathbb{N},\ \varphi \in X^*$$

mit positiven Konstanten $c = c(\varphi)$. Dann ist $\sup_n \|x_n\| < \infty$.

Beweis.
Wir betrachten die kanonische *Abbildung*

$$\imath : X \to X^{**},\ x \mapsto \imath_x$$

in den Bidualraum *$X^{**} := \big(X^*\big)^*$ von X, d. h. jedem $x \in X$ wird eine stetige lineare Abbildung $\imath_x : X^* \to \mathbb{R}$ zugeordnet, wobei $\imath_x \varphi := \varphi(x)$ für alle $\varphi \in X^*$ ist. Wegen $|\imath_x \varphi| \leq \|\varphi\|_\infty \|x\|$ ist*

$$\|\imath_x\|_\infty = \sup_{\varphi \in X^* \setminus \{0\}} \frac{|\imath_x \varphi|}{\|\varphi\|_\infty} \leq \|x\| \quad \textit{für alle} \quad x \in X,$$

weshalb $\imath$ *eine* Einbettung *ist*[d]. *Man nennt* $\imath$ *die* kanonische Einbettung *von* X *in* X^{**} *und schreibt* $X \hookrightarrow X^{**}$.
Nach dem Satz von Hahn–Banach 3.19 können wir zu jedem $x \neq 0$ *ein* $\varphi \in X^*$ *so wählen, dass* $\|\varphi\|_\infty = 1$ *und* $\varphi(x) = \|x\|$, *also* $\|\imath_x\varphi\| = \|x\|$ *ist. Mithin ist* $\imath$ *eine lineare Isometrie . Sei nun* $T_n := \imath_{x_n} : X^* \to \mathbb{R}$ *für eine Folge* $(x_n) \subset X$. *Dann ist*

$$|T_n\varphi| = \big|\varphi(x_n)\big| \leq c \quad \text{für alle} \quad n \in \mathbb{N},\ \varphi \in X^*,$$

nach dem Satz von Banach–Steinhaus 5.3 also $\sup_n \|T_n\|_\infty = \sup_n \|x_n\| < \infty$.

□

Lemma 5.6
Seien X *ein normierter Raum,* $(x_n) \subset X$ *eine Folge und* $x \in X$. *Dann gilt:*

i) *Der schwache Limes von* (x_n) *ist, falls er existiert, eindeutig bestimmt.*

ii) $\|x_n - x\| \xrightarrow{n} 0 \quad \Longrightarrow \quad x_n \overset{n}{\rightharpoonup} x.$

iii) $x_n \overset{n}{\rightharpoonup} x \quad \Longrightarrow \quad \|x\| \leq \liminf_n \|x_n\|$ *(Unterhalbstetigkeit).*

iv) $x_n \overset{n}{\rightharpoonup} x \quad \Longrightarrow \quad \sup_n \|x_n\| < \infty$ *(Normbeschränktheit).*

Ist X *ein Hilbert–Raum, so gilt die partielle Umkehrung von i):*

$$x_n \overset{n}{\rightharpoonup} x \in X \text{ und } \|x_n\| \xrightarrow{n} \|x\| \quad \Longleftrightarrow \quad \|x_n - x\| \xrightarrow{n} 0. \tag{5.6}$$

Bemerkung.
Die Aussagen aus Lemma 5.6 gelten natürlich erstrecht bei starker Konvergenz. In diesem Fall hat man statt iii) allerdings die stärkere Aussage

$$x_n \xrightarrow{n} x \text{ (stark) in } X \quad \Longrightarrow \quad \|x\| = \lim_n \|x_n\|.$$

Dies ergibt sich aus der Tatsache, dass in jedem normierten Raum auch die Dreiecksungleichung nach unten, also

$$\big|\|x\| - \|y\|\big| \leq \|x - y\|$$

für alle $x, y \in X$ *gilt.*

□

[d] Sind X und Y normierte Räume, so heißt eine Abbildung $\imath : X \to Y$ eine *Einbettung*, falls X Untervektorraum von Y und $\imath$ eine lineare Abbildung mit

$$\|\imath x\| \leq M\|x\| \quad \text{für alle} \quad x \in X$$

mit einer positiven Konstanten M ist. Die lineare Abbildung $\imath$ ist dann insbesondere stetig, weshalb man auch von einer *stetigen Einbettung* spricht. Man schreibt dann $X \hookrightarrow Y$.

Bemerkung 5.7

Die Umkehrung der Aussage ii) in Lemma 5.6 gilt nur, falls $\dim X < \infty$ *ist. Dagegen behält die partielle Umkehrung* (5.6) *derselben allgemeiner in sog.* gleichmäßig konvexen *Räumen ihre Gültigkeit. Dazu zählen insbesondere die für uns wichtigen L^p–Räume mit $1 < p < \infty$.*
Es sei noch bemerkt, dass aus der schwachen Konvergenz in L^p i. a. keine fast überall punktweise Konvergenz (für eine Teilfolge) folgt, wie dies für die starke Konvergenz in L^p der Fall ist (vgl. Kor. 3.7).

Lemma 5.8

Seien H ein Hilbert–Raum und $(x_n) \subset H$ eine Folge, so dass $\lim_n \varphi(x_n)$ für alle $\varphi \in H^*$ existiert. Dann ist (x_n) schwach konvergent in H.

Beweis.

Aus Korollar 5.5 folgt, dass (x_n) beschränkt ist. Für ein $y \in H$ gehört $\langle \cdot, y\rangle$ zu H^*, d. h. die Abbildung φ, die gegeben ist durch $\varphi(y) := \lim_n \langle x_n, y\rangle$ ist wohldefiniert und linear. Mittels der Ungleichung von Cauchy–Schwarz (Lemma 4.3 i)) wird $\left|\varphi(y)\right| \leq \limsup_n \|x_n\|\|y\|$ und es ist $\sup_n \|x_n\| < \infty$, und damit $\varphi \in H^*$. Nach dem Satz von Riesz 4.12 existiert daher ein $x \in H$ mit $\varphi = \langle \cdot, x\rangle$, also

$$\varphi(x_n) = \langle y, x_n\rangle \xrightarrow{n} \langle y, x\rangle = \varphi(x) \quad \textit{für alle} \quad y \in H.$$

□

Wir kommen jetzt zu einem der entscheidenden Sätze dieses §, der schwachen Version des Satzes von Bolzano–Weierstraß in Hilbert–Räumen, welcher besagt, dass in Hilbert–Räumen das schwache Auswahlprinzip (5.4) *gilt.*

Satz 5.9 (Schwaches Auswahlprinzip in Hilbert–Räumen)

Seien H ein Hilbert–Raum und $(x_n) \subset H$ eine beschränkte *Folge, d. h. es gelte $\sup_n \|x_n\| < \infty$. Dann ist (x_n) schwach (folgen–) kompakt in H, d. h. es existiert eine Teilfolge $\left(x_{n_k}\right)$ von (x_n) und ein $x \in H$, so dass $x_{n_k} \xrightarrow{k} x$.* [e]

Bevor wir zum Beweis kommen, führen wir noch einen nützlichen Begriff ein: Wir sagen, ein normierter Raum X sei separabel, *falls X eine abzählbare Teilmenge Ξ besitzt, so dass $\overline{\Xi} = X$ ist. Mit anderen Worten soll es eine in X dicht liegende Menge Ξ geben, die gleichzeitig abzählbar ist. Ein einfaches endlichdimensionales Beispiel ist $\mathbb{R}$, da $\mathbb{Q}$ ja bekanntlich dicht in $\mathbb{R}$ liegt.*

[e] Man beachte, dass in Hilbert–Räumen *folgenkompakt* und *überdeckungskompakt* gleichbedeutend sind, dies jedoch für *schwach folgenkompakt* i. a. falsch ist; es sei denn, der Raum ist separabel.

Lemma 5.9
Ist H ein Hilbert–Raum, und $U \subset H$ ein separabler Unterraum.

a) Dann ist U selbst ein Hilbert–Raum.

b) Gilt das schwache Auswahlprinzip für alle seperablen Unterräume von H, so gilt es auch in H.

Beweis.
a) Sei (x_n) Cauchy-Folge in U, dann ist (x_n) auch Cauchy-Folge in H und damit konvergent gegen $x \in H$. Zu zeigen bleibt $x \in U$. Sei $\epsilon > 0$ gegeben, dann gibt es $x_n \in U$ mit $n = n(\epsilon)$ so dass

$$\|x - x_n\| < \frac{\epsilon}{2}.$$

Zu $x_n \in U$ gibt es wegen der Seperabilität von U ein $\xi_n \in \Xi$ (dabei gelte $\overline{\Xi} = U$) mit

$$\|x_n - \xi_n\| < \frac{\epsilon}{2} \quad \Rightarrow \quad \|x - \xi_n\| < \epsilon,$$

d.h. x lässt sich durch Elemente von Ξ approximieren also $x \in U = \overline{\Xi}$.
b) Sei $(x_n) \subset H$ eine beschränkte Folge, d. h. es gelte $\sup_n \|x_n\| < \infty$. Wir betrachten den seperablen Unterraum $U = \overline{\operatorname{span}\{x_1, x_2, ...\}}$. Dann ist $(x_n) \subset U$ beschränkt und daher existiert eine Teilfolge (x'_n) die schwach konvergent ist gegen ein $u \in U$ es folgt für $y \in H$ mit $y = w + v \in U \oplus U^\perp$

$$\langle x'_n, y\rangle = \langle x'_n, w\rangle \to \langle u, y\rangle,$$

d.h. $x'_n \rightharpoonup u$. □

Daher genügt es Satz 5.9 für sperable Räume zu beweisen.

Beweis von Satz 5.9.
Sei H separabel. Dann gibt es eine abzählbare Menge $\Xi := \{\xi_1, \xi_2, \xi_3, \ldots\} \subset H$, so dass $\overline{\Xi} = H$ ist. Wegen $\sup_n \|x_n\| < \infty$ ist $\big(\langle x_n, \xi_1\rangle\big)$ eine beschränkte Folge in $\mathbb{R}$. Nach dem Satz von Bolzano–Weierstraß existiert daher der Limes

$$\alpha^1 := \lim_n \langle x_n^1, \xi_1\rangle$$

für eine Teilfolge (x_n^1) von (x_n). Desweitern können wir aus (x_n^1) eine Teilfolge (x_n^2) wählen, so dass auch

$$\alpha^2 := \lim_n \langle x_n^2, \xi_2\rangle$$

existiert. Sukzessive so fortfahrend erhalten wir für jedes $k \in \mathbb{N}$ eine Folge (x_n^k), so dass

$$\alpha^k := \lim_n \langle x_n^k, \xi_k \rangle$$

existiert, wobei (x_n^k) eine Teilfolge von (x_n^{k-1}) ist (setze $(x_n^0) := (x_n)$). Wir betrachten nun die Diagonalfolge $(\zeta_n) := (x_n^n)$. Dieselbe ist ab einem gewissen Index $n_0 = n_0(k) \in \mathbb{N}$ eine Teilfolge von (x_n^k), und daher

$$\alpha^k = \lim_n \langle \zeta_n, \xi_k \rangle \quad \textit{für alle} \quad k \in \mathbb{N}, \ . \tag{5.7}$$

Seien nun $M := \sup_j \|x_j\|$, $\xi \in H$ und $\varepsilon > 0$ beliebig. Da Ξ nach Voraussetzung dicht in H liegt, können wir ξ_k ($k \in \mathbb{N}$ fest) so wählen, dass

$$\|\xi - \xi_k\| < \frac{\varepsilon}{2M}$$

ausfällt. Für beliebige $m, n \in \mathbb{N}$ erhalten wir dann mittels der Ungleichung von Cauchy–Schwarz die Abschätzung

$$\begin{aligned}
\big|\langle \zeta_m, \xi \rangle - \langle \zeta_n, \xi \rangle\big| &\le \big|\langle \zeta_m, \xi_k \rangle - \alpha^k\big| + \big|\langle \zeta_n, \xi_k \rangle - \alpha^k\big| \\
&\qquad + \big|\langle \zeta_m, \xi - \xi_k \rangle\big| + \big|\langle \zeta_n, \xi - \xi_k \rangle\big| \\
&\le \big|\langle \zeta_m, \xi_k \rangle - \alpha^k\big| + \big|\langle \zeta_n, \xi_k \rangle - \alpha^k\big| + \varepsilon,
\end{aligned}$$

weil offenbar $\|\zeta_n\| \le M$ für jedes $n \in \mathbb{N}$ ist. Zusammen mit (5.7) *ergibt sich daraus, dass $\big(\langle \zeta_n, \xi \rangle\big) \subset \mathbb{R}$ für jedes $\xi \in H$ eine Cauchy–Folge ist, so dass also* $\lim_n \langle \zeta_n, \xi \rangle$ *für jedes $\xi \in H$ existiert. Daraus folgt die Behauptung mit Lemma 5.8 und Satz 4.12.* □

Wir wollen nun der Frage nachgehen, welche Bedingungen ein Raum X notwendig erfüllen muss, damit das schwache Auswahlprinzip (5.4) *gilt.*

Definition 5.10 (Reflexivität)
Seien X ein normierter Raum und $\imath : X \hookrightarrow X^{**}$ die kanonische Einbettung von X in X^{**} (vgl. Beweis von Kor. 5.5). Dann heißt X reflexiv, falls $\imath$ surjektiv, also $X^{**} = \imath(X)$ ist.

Wie bereits eingangs angekündigt, wollen wir uns überlegen, dass die reflexiven Räume genau diejenigen sind, in denen das schwache Auswahlprinzip gilt. Wir beginnen mit dem

Satz 5.11 (Reflexivität)

Seien $(H, \langle\cdot,\cdot\rangle)$ ein Hilbert–Raum, (X,μ) ein Maßraum und $1 < p < \infty$.

i) H ist reflexiv.

ii) $L^p(X;\mu)$ ist reflexiv.

Beweis.

i) Sei $T \in H^{**}$, also eine lineare, stetige Abbildung $T : H^* \to \mathbb{R}$. Es ist zu zeigen, dass es ein $x \in H$ gibt mit $T = \imath_x$. Dazu bezeichne $R : H \to H^*$, $z \mapsto \langle z,\cdot\rangle$ den isometrischen Isomorphismus aus dem Satz von Riesz 4.12. Dann ist die Abbildung $S := T \circ R : H \to \mathbb{R}$ linear und stetig, also $S \in H^*$. Nach dem Satz von Riesz 4.12 gibt es ein $x \in H$, so dass $S = Rx$ auf H. Dies bedeutet aber:

$$T\big(\langle z,\cdot\rangle\big) = Sz = \langle x, z\rangle \quad \text{für alle} \quad z \in H.$$

Sei nun $\varphi \in H^*$ beliebig. Wiederum nach dem Satz von Riesz ist $\varphi = \langle z,\cdot\rangle$ für ein $z \in H$, und damit

$$T\varphi = T\big(\langle z,\cdot\rangle\big) = \langle x, z\rangle = \varphi(x) = \imath_x(\varphi),$$

woraus die Behauptung folgt.

ii) Nach Satz 3.17 ist $L^p(X;\mu) \cong L^p(X;\mu)^{**}$. Wir überlegen uns, dass diese Isomorphie von der kanonischen Einbettung herrührt. Für $s \in (1,\infty)$ sei $J_s : L^s(X;\mu) \to (L^{s^*}(X;\mu))^*$ mit $s^* := s/(s-1)$ definiert durch $J_s(u) = u^*$ mit

$$u^*(w) = \int_X uw \, d\mu \ \text{ für } w \in L^{s^*}(X;\mu).$$

Sei $T \in (L^p(X;\mu))^{**}$ und $q := p/(p-1)$. Wir betrachten $\phi \in (L^q(X;\mu))^*$ mit

$$\varphi(w) := T \circ J_q(w),\, w \in L^q(X;\mu).$$

Es folgt für $\psi \in (L^p(X;\mu))^*$,

$$T(\psi) = T(J_q(w)) = \varphi(w) = \int_X uw \, d\mu = J_q(w)(u) = \psi(u),$$

wobei $w \in L^q(X;\mu)$ und $u \in L^p(X;\mu)$ passend gewählt werden. Es gilt also $T = i_u$. □

Wie der Beweis von Satz 5.11 ii) zeigt, kann weder $L^1(X;\mu)$ *noch* $L^\infty(X;\mu)$ *i. a. reflexiv sein, da (vgl. § 3)* $L^1(X;\mu) \subsetneqq L^\infty(X;\mu)^*$ *ist.*

Bemerkung.
Jeder endlich–dimensionale Raum ist reflexiv.

Bevor wir den Zusammenhang zwischen der Reflexivität und dem schwachen Auswahlprinzip herstellen, verschaffen wir uns ein genaueres Verständnis von reflexiven Räumen, und stellen einige Hilfsmittel bereit. Das folgende Lemma ist als Trennungsatz von Mazur *bekannt; es gilt allgemeiner in* lokal konvexen, *topologischen Räumen (vgl. etwa [Yo], § IV.6, Thm. 3) und ist eine weitere Version des Satzes von Hahn–Banach 3.19.*

Lemma 5.12 (Mazur, Trennungssatz)
Seien X *ein normierter Raum und* $K \subset X$ *abgeschlossen und konvex. Dann gibt es zu jedem* $x \in X \setminus K$ *ein* $\varphi \in X^*$ *und eine Konstante* $c > 0$*, so dass*

$$\varphi(x) > c \quad \textit{und} \quad \sup_{z\in K} \big|\varphi(z)\big| \le c$$

ist. Ist K *ein abgeschlossener Unterraum von* X*, so darf man* $c = 0$ *wählen.*

Sei X *ein normierter Raum und sei* $A \subset X$*. Dann nennen wir* A schwach (folgen–) abgeschlossen, *falls jede Folge* $(x_n) \subset A$ *in* X *schwach konvergiert mit schwachem Limes in* A*. Als Korollar des Trennungssatzes erhält man:*

Satz 5.13 (Schwache Abgeschlossenheit)
Seien X *ein normierter Raum und* $K \subset X$ *abgeschlossen und konvex. Dann ist* K *erstrecht* schwach *abgeschlossen.*

Beweis.
Sei $(x_n) \subset K$ *eine Folge mit* $x_n \overset{n}{\rightharpoonup} x$ *und angenommen, es ist* $x \in X \setminus K$*. Nach dem Trennungssatz 5.12 gibt es dann ein* $\varphi \in X^*$ *und eine Konstante* $c > 0$*, so dass*

$$\varphi(x) > c \quad \textit{und} \quad \sup_{z\in K} \big|\varphi(z)\big| \le c$$

ist. Es strebt aber $\varphi(x_n) \xrightarrow{n} \varphi(x)$*, also ist* $\big|\varphi(x)\big| \le c$*; Widerspruch.* □

Die Beweise der folgenden Aussagen überlassen wir dem Leser als Übung (vgl. A 17).

Satz 5.14
Seien X ein Banach–Raum und $U \subset X$ ein Unterraum.

i) Ist X reflexiv und U abgeschlossen, so ist auch U reflexiv.

ii) X reflexiv $\iff$ X^* reflexiv.

iii) Endliche Produkte reflexiver Räume sind reflexiv.

Ist X ein normierter Raum, so nennen wir eine Folge $(x_n) \subset X$ eine schwache Cauchy–Folge*, falls $\big(\varphi(x_n)\big) \subset \mathbb{R}$ für alle $\varphi \in X^*$ eine Cauchy–Folge ist. Entsprechend heißt X* schwach vollständig*, wenn jede schwache Cauchy–Folge in X schwach konvergiert.*

Satz 5.15 (Schwache Vollständigkeit)
Jeder reflexive Raum X ist schwach vollständig.

Beweis.
Sei $(x_n) \subset X$ eine schwache Cauchy–Folge. Betrachte die Folge $(T_n) \subset X^{**}$, die gegeben wird durch $T_n\varphi := \varphi(x_n)$ für jedes $\varphi \in X^*$. Nach Voraussetzung existiert $T\varphi := \lim_n T_n\varphi$ für alle $\varphi \in X^*$ und nach dem Satz von Banach–Steinhaus 5.3 ist $T \in X^{**}$. Da X reflexiv ist, ist daher $T = \imath_x$ für ein $x \in X$. Mithin ist

$$\lim_n \varphi(x_n) = \lim_n T_n\varphi = T\varphi = \imath_x(\varphi) = \varphi(x) \quad \text{für alle} \quad \varphi \in X^*,$$

d. h. (x_n) ist schwach konvergent. □

Lemma 5.16 (Separabilität)
Sei X ein normierter Raum und X^* separabel. Dann ist auch X separabel.

Beachte. *Die Umkehrung der Aussage in Lemma 5.16 ist i. a. falsch).*

Beweis von Lemma 5.16.
Seien $S^* := \big\{\varphi \in X^*;\ \|\varphi\|_\infty = 1\big\}$ und Ξ eine abzählbare Menge mit $X^* = \overline{\Xi}$. Dann ist auch $\Sigma := \big\{\varphi \in \Xi;\ \varphi \not\equiv 0\big\}$ abzählbar und es ist

$$S^* = \overline{\big\{\varphi/\|\varphi\|_\infty;\ \varphi \in \Sigma\big\}} =: \overline{\Sigma_0},$$

also S^* separabel. Schreiben wir $\Sigma_0 := \big\{\psi_1, \psi_2, \dots\big\}$, so gibt es zu jedem $\psi_k \in \Sigma_0$ $(k \in \mathbb{N})$ eine Folge $(x_n^k) \subset X$ mit den Eigenschaften:

$$\|x_n^k\| = 1 \quad \text{und} \quad \psi(x_n^k) \xrightarrow{n} \|\psi\|_\infty = 1. \tag{5.8}$$

Für $U := \overline{\operatorname{span}\left\{x_n^k;\ (k,n) \in \mathbb{N}^2\right\}}$ *muss nun* $U = X$ *sein. Denn angenommen, es ist* $U \subsetneqq X$. *Dann existiert nach dem Satz von Hahn–Banach 3.19 ein* $\psi \in S^*$ *mit* $\psi|_U = 0$. *Andererseits gibt es ein* $\psi_k \in S^*$ *mit* $\|\psi - \psi_k\|_\infty \leq \frac{1}{2}$, *so dass*

$$\left|\psi_k(x_n^k)\right| = \left|(\psi - \psi_k)(x_n^k)\right| \leq \frac{1}{2}$$

ist, im Widerspruch zu (5.8). □

Wir kommen zum Hauptergebnis dieses Kapitels:

Satz 5.17 (Schwache Kompaktheit)
Jeder reflexive Raum X *ist schwach (folgen–) kompakt, d. h. es gilt das schwache Auswahlprinzip* (5.4).

Beweis.
Sei $(x_n) \subset X$ *eine beschränkte Folge.*

i) Sei zunächst X^* *separabel, d. h.* $X^* = \overline{\Xi}$ *für eine abzählbare Menge* Ξ. *Wir gehen wie im Beweis von Satz 5.9 vor: Wie dort zeigt man die Existenz einer Teilfolge von* (x_n) *(o. E.* (x_n) *selbst), so dass* $\lim_n \varphi(x_n)$ *für alle* $\varphi \in \Xi$ *existiert. Wegen* $X^* = \overline{\Xi}$ *folgt daraus die Existenz von* $\lim_n \varphi(x_n)$ *für alle* $\varphi \in X^*$, *und daher die Behauptung mit Satz 5.15.*

ii) Sei nun X^* *beliebig, und für eine Folge* $(\varphi_n) \subset X^*$ *sei*

$$U := \overline{\operatorname{span}\left\{\varphi_n;\ n \in \mathbb{N}\right\}}.$$

Dann ist U *separabel und als abgeschlossener Unterraum des reflexiven Raumes* X^* *selbst reflexiv (Satz 5.14). Also ist* $U^{**} = \imath(U)$ *separabel, und damit nach Lemma 5.16 auch* U^*. *Wie in i) erhält man eine Teilfolge von* (x_n) *(o. E.* (x_n) *selbst) und ein* $x \in U$ *mit* $\varphi(x_n) \xrightarrow{n} \varphi(x)$ *für alle* $\varphi \in U^*$. *Sei nun* $\psi \in X^*$ *beliebig. Dann ist* $\varphi := \psi|_U \in U^*$ *und es gilt*

$$\psi(x_n) = \varphi(x_n) \xrightarrow{n} \varphi(x) = \psi(x),$$

d. h. $x_n \overset{n}{\rightharpoonup} x$ *in* X. □

Wir kommen nun zu den für uns wichtigen L^p*–Räumen. Zunächst bemerken wir, dass aufgrund von Satz 3.17 folgendes gilt: Seien* $(u_n) \subset L^p(X;\mu)$ *eine*

Folge und $u \in L^p(X;\mu)$*, wobei* $1 \leq p < \infty$ *ist. Dann ist*

$$u_n \overset{n}{\rightharpoonup} u \ \text{in } L^p(X;\mu) \quad \Longleftrightarrow \quad \begin{cases} \displaystyle\int_X u_n \varphi \, d\mu \xrightarrow{n} \int_X u\varphi \, d\mu \\ \text{für alle} \quad \varphi \in L^{p'}(X;\mu). \end{cases} \tag{5.9}$$

Damit bekommt man aus Satz 5.17 (in Verbindung mit Satz 5.11):

Korollar 5.18 (Schwache Kompaktheit von L^p)

Seien (X,μ) *ein Maßraum,* $1 < p < \infty$ *und* $(u_n) \subset L^p(X;\mu)$ *eine* beschränkte *Folge, d. h. es ist* $\sup_n \|u_n\|_p < \infty$*. Dann gibt es eine Teilfolge* $\big(u_{n_k}\big) \subset (u_n)$ *und eine Funktion* $u \in L^p(X;\mu)$*, so dass (mit* p' *wie in (3.7))*

$$\int_X u\varphi \, d\mu = \lim_k \int_X u_{n_k} \varphi \, d\mu \quad \text{für alle} \quad \varphi \in L^{p'}(X;\mu), \tag{5.10}$$

d. h. es gilt: $u_{n_k} \overset{k}{\rightharpoonup} u$ *in* $L^p(X;\mu)$*.*

In Banach–Räumen (wie $\big(L^p(X;\mu), \|\cdot\|_p\big)$*) gilt auch die Umkehrung von Satz 5.17 (vgl. dazu [Yo], App. Chap. V, § 4), womit also die Reflexivität genau das richtige Konzept ist, um die schwache Bolzano–Weierstraß–Eigenschaft zu garantieren:*

Satz 5.19 (Eberlein–Shmulyan)

Sei X ein Banach–Raum. Dann ist X reflexiv genau dann, wenn in X die schwache Bolzano–Weierstraß–Eigenschaft (5.4) *gilt.*

Demnach ist die Aussage von Korollar 5.18 äquivalent zur Reflexivität von L^p *(*$p \in (1,\infty)$*). Da weder* L^1 *noch* L^∞ *reflexiv sind, zeigt dies umgekehrt, dass diese Aussage nicht für* $p = 1$ *bzw.* $p = \infty$ *gelten kann.*

Schlußbemerkung.

Alle Aussagen über die Räume $L^p(\Omega)$ *in diesem § gelten natürlich auch für die vektoriellen Räume* $L^p(\Omega)^D$*.*

§ 6. Schwache Differenzierbarkeit und Distributionen

Sei $\Omega \subset \mathbb{R}^d$ offen, beschränkt und mit genügend glattem Rand (z. B. Ω ein beschränktes Lipschitz–Gebiet). Wir betrachten (wie schon in § 4) das Dirichlet–Problem zur Poisson–Gleichung: Gesucht ist eine Lösung $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$ des Randwertproblems

$$\left.\begin{aligned} \Delta u &= f \quad \textit{auf} \quad \Omega; \\ u &= 0 \quad \textit{auf} \quad \partial\Omega, \end{aligned}\right\} \tag{6.1}$$

mit einer vorgegebenen Funktion $f \in C^0(\overline{\Omega})$. Wie wir bereits in § 3 gesehen haben, genügt eine Lösung u von (6.1) *der Beziehung*

$$\int_\Omega \nabla u \cdot \nabla \eta \, dx = -\int_\Omega f\eta \, dx \quad \textit{für alle} \quad \eta \in C^1_\circ(\Omega). \tag{6.2}$$

Dabei wird durch die linke Seite von (6.2) *ein Skalarprodukt β auf $C^1_\circ(\Omega)$ induziert, während die rechte Seite als stetige lineare Abbildung von $C^1_\circ(\Omega) \to \mathbb{R}$, also als ein Element von $C^1_\circ(\Omega)^*$, aufgefasst werden kann. Es läge nun nahe, den Darstellungssatz von Riesz 4.12 anzuwenden, um die Existenz einer Lösung zu zeigen. Diesen kann man hier jedoch nicht benutzen, da der Raum $C^1_\circ(\Omega)$ bzgl. β kein Hilbert–Raum ist (vgl. Beispiel 4.6). Bereits in* (4.4) *haben wir aber die Existenz eines Hilbert–Raums angekündigt, welcher eine Lösung für dieses Problem bietet. Der Ausweg besteht dabei im wesentlichen in einer Abschwächung des Differenzierbarkeitsbegriffs: Da in dem Skalarprodukt β die Ableitungen nicht punktweise ausgewertet, sondern nur integriert werden, ergibt sich ein gewisser Spielraum. (Den genannten Hilbert–Raum werden wir erst im folgenden § erklären.)*

Wir wollen uns nun ein vernünftiges Konzept für die sog. schwache Ableitung *einer Funktion $u : \Omega \to \mathbb{R}$, $\Omega \subset \mathbb{R}^d$ offen, überlegen.*

i) Zunächst kann man versuchen, die klassische Ableitung abzuschwächen in der Art, dass man ihre Existenz nur f. ü. auf Ω verlangt: Sei also ∇u f. ü. auf Ω vorhanden und es sei $\nabla u \in L^2(\Omega)^d$. Hat auch $w : \Omega \to \mathbb{R}$

diese Eigenschaft, so ist $\beta(u, w)$, gegeben durch die linke Seite von (6.2), wohldefiniert (Warum?).
Jedoch gehen bei dieser Begriffsbildung entscheidende Zusammenhänge zwischen u und ∇u verloren: Sei etwa $u := \mathbb{1}_{(0,\infty)}$. Dann existiert $u'(x)$ für alle $x \neq 0$ und es ist $u' = 0$ $\mathcal{L}^1$–f. ü. auf $\mathbb{R}$. Bei dieser Begriffsbildung impliziert $\nabla u = 0$ also nicht unbedingt — wie man erwarten würde —, dass u konstant ist. Das ist wenig sinnvoll.

ii) *Seien $u \in C^1(\Omega)$, $\gamma \in \{1, \dots, d\}$ und $h \in \mathbb{R}$ mit $0 < |h| < \operatorname{dist}(\omega, \partial\Omega)$ für ein Teilgebiet $\omega \Subset \Omega$. Dann ist der* Differenzenquotient

$$\Delta_\gamma^h u(x) := \frac{u(x + he_\gamma) - u(x)}{h} \quad (x \in \Omega) \tag{6.3}$$

von u in Richtung e_γ wohldefiniert auf Ω und es gilt:

$$\Delta_\gamma^h u \xrightarrow{h \to 0} \partial_\gamma u \quad \textit{lokal gleichmäßig auf } \Omega.$$

Dies lässt sich nun leicht zu einem vernünftigen Konzept für die „Ableitung" einer Funktion $u \in L^1_{loc}(\Omega)$ ausbauen: Man verlangt die Existenz von Funktionen $w_1, \dots, w_d \in L^1_{loc}(\Omega)$ mit

$$\Delta_\gamma^h u \xrightarrow{h \to 0} w_\gamma \quad \textit{in } L^1_{loc}(\Omega),$$

und nennt w_γ die γ–te schwache (partielle) Ableitung *von u und $w := (w_1, \dots, w_d)$ die* schwache Ableitung (Gradient) *von u. Leider ist diese Definition etwas unhandlich, wir werden später aber sehen, dass sie das Richtige leistet (siehe Satz 6.16).*

iii) *Für $\gamma \in \{1, \dots, d\}$ seien $u, w_\gamma \in L^1_{loc}(\Omega)$. Es wäre sinnvoll w_γ als die γ–te schwache partielle Ableitung von u anzusprechen, falls eine Folge $(u_n) \subset C^\infty(\Omega)$ existiert mit*

$$u_n \xrightarrow{n} u \quad \textit{und} \quad \partial_\gamma u_n \xrightarrow{n} w_\gamma \ \textit{in } L^1_{loc}(\Omega).$$

Natürlich muss man sich dabei davon überzeugen, dass diese Begriffsbildung nicht von der speziellen Folge abhängt. Auch diese Definiton leistet das Gewünschte (werden wir in der Nachfolgeveranstaltung sehen), ist jedoch ebenfalls unhandlich.

iv) *Das folgende Konzept der „Distributionsableitung" ist begrifflich am einfachsten zu verstehen und gleichzeitig — wie der Name schon andeutet — auf viel allgemeinere Situationen anwendbar.*

Seien $u \in C^1(\Omega)$, $w_\gamma \in L^1_{loc}(\Omega)$ *und* $\gamma \in \{1, \ldots, d\}$, *so dass gilt:*

$$\int_\Omega u\,\partial_\gamma \eta\,dx = -\int_\Omega w_\gamma \eta\,dx \quad \text{für alle} \quad \eta \in C_\circ^\infty(\Omega). \tag{6.4}$$

Partielle Integration liefert dann:

$$\int_\Omega \partial_\gamma u\,\eta\,dx = \int_\Omega w_\gamma \eta\,dx \quad \text{für alle} \quad \eta \in C_\circ^\infty(\Omega).$$

Also ist nach dem Fundamentallemma 3.22 $w_\gamma = \partial_\gamma u$. *Die linke Seite von* (6.4) *macht auch für* $u \in L^1_{loc}(\Omega)$ *Sinn, und motiviert die folgende Definiton.*

Definition 6.1 (Schwache Differenzierbarkeit)
Seien $\Omega \subset \mathbb{R}^d$ *offen,* $\gamma \in \mathbb{N}_0^d$ *und* $k \in \mathbb{N}_0$, *und seien* $u, w^\gamma \in L^1_{loc}(\Omega)$. *Dann heißt* w^γ *die* γ*–te schwache (partielle) Ableitung von* u, *falls gilt*

$$\int_\Omega u\,\partial^\gamma \eta\,dx = (-1)^{|\gamma|} \int_\Omega w^\gamma \eta\,dx \quad \text{für alle} \quad \eta \in C_\circ^\infty(\Omega); \tag{6.5}$$

u *heißt* k*–mal* schwach differenzierbar *auf* Ω, *falls für jedes* $\gamma \in \mathbb{N}_0^d$ *mit* $|\gamma| \le k$ *eine Funktion* $w^\gamma \in L^1_{loc}(\Omega)$, *welche der Beziehung (6.5) genügt, existiert. Den Raum aller* k*–mal auf* Ω *schwach differenzierbaren Funktionen* $u \in L^1_{loc}(\Omega)$ *bezeichnen wir mit* $W^k(\Omega)$. *Insbesondere sei* $W^0(\Omega) := L^1_{loc}(\Omega)$.

Bemerkung 6.2

i) In Definition 6.1 genügt es zu verlangen, dass (6.5) *für alle Funktionen* $\eta \in C_\circ^{|\gamma|}(\Omega)$ *gilt. Ferner lassen sich die obigen Konzepte auch in* L^p_{loc} *statt* L^1_{loc} *sowie für vektorwertige Funktionen (komponentenweise) formulieren. Die entsprechenden Räume vektorieller, schwach differenzierbarer Funktionen bezeichnen wir wie üblich mit* $W^k(\Omega)^D = W^k(\Omega, \mathbb{R}^D)$. *Die Eigenschaften der skalaren Räume* $W^k(\Omega)$ *übertragen sich sinngemäß auf* $W^k(\Omega)^D$.

ii) Per Definition ist $W^k(\Omega) \subset L^1_{loc}(\Omega)$, *d. h. schwach differenzierbare Funktionen sind Äquivalenzklassen* $\mathcal{L}^d$*–messbarer Funktionen* $\Omega \to \mathbb{R}$. *Zu* $u \in W^k(\Omega)$ *gibt es höchstens einen stetigen Vertreter (die Gleichheit in* Ω *bis auf eine Nullmenge stetiger Funktionen impliziert Gleichheit überall in* Ω*). Allgemein haben schwach differenzierbare Funktionen keine stetigen Vertreter (vgl. Übung).*

iii) *Hat $u \in L^1_{loc}(\Omega)$ eine $|\gamma|$–te schwache Ableitung w^γ, so ist diese nach dem Fundamentallemma 3.22 eindeutig bestimmt, und man schreibt wie üblich wieder $\partial^\gamma u$ statt w^γ, sofern sich die genaue Bedeutung von $\partial^\gamma u$ aus dem Kontext erschließt. Ebenso verwendet man alle anderen Bezeichnungen aus der klassischen Differentialrechnung für schwach differenzierbare Funktionen.*

iv) *Es ist $C^k(\Omega) \subset W^k(\Omega)$, und die schwachen Ableitungen einer C^k–Funktion stimmen mit den klassischen Ableitungen überein (bzw. werden von jenen erzeugt).*

v) *Eine Funktion u gehört genau dann zur Klasse $W^1(\Omega)$, falls gilt:*

$$\int_\Omega u \operatorname{div} \eta \, dx = -\int_\Omega \nabla u \cdot \eta \, dx \quad \text{für alle} \quad \eta \in C^\infty_\circ(\Omega)^d.$$

Entsprechend ist $u \in W^1(\Omega)^D$, falls gilt:

$$\int_\Omega u \cdot \operatorname{div} \eta \, dx = -\int_\Omega \nabla u : \eta \, dx \quad \text{für alle} \quad \eta \in C^\infty_\circ(\Omega)^{dD}.$$

Dabei bezeichnet für $w := (w_\nu)_{\nu=1,\dots,D}$ jetzt ∇w die (schwache) Jacobi–Matrix $\big(\partial_\gamma w^\nu\big)_{\substack{\gamma=1,\dots,d\\ \nu=1,\dots,D}} \in \mathbb{R}^{d\times D} \cong \mathbb{R}^{dD}$ und $\operatorname{div} w$ steht für die (schwache) vektorielle Divergenz:

$$\operatorname{div} w := \left(\sum_{\gamma=1}^{d} \operatorname{div} w^\nu \right)_{\nu=1,\dots,D} \in \mathbb{R}^D.$$

vi) *Es gibt Funktionen $u \in L^1_{loc}(\Omega)$, die weder klassisch noch schwach differenzierbar sind. Einfache Beispiele liefern charakteristische Funktionen.*

Beispiel 6.3

Funktion mit $u \in C^1\big(\Omega \setminus \{\xi\}\big)$. Dann ist i. a. sowohl $u \notin L^1_{loc}(\Omega)$ als auch $\nabla u \notin L^1_{loc}(\Omega)^d$, wobei ∇u die auf $\Omega \setminus \{\xi\}$ existierende klassische Ableitung von u bezeichnet. Ferner braucht eine bis auf eine Stelle stetig differenzierbare Funktion auch nicht in $W^1(\Omega)$ zu sein (siehe Übung).

Lemma 6.4

Seien $\Omega \subset \mathbb{R}^d$ offen, $\xi \in \Omega$, $s \in \mathbb{R}$ und $u : \Omega \to \mathbb{R}$, gegeben durch

$$u(x) := \begin{cases} \dfrac{1}{|x-\xi|^s} & ; \quad x \neq \xi \\ 0 & ; \quad x = \xi. \end{cases}$$

Für $k \in \mathbb{N}_0$ gilt dann:

$$u \in W^k(\Omega) \quad \Longleftrightarrow \quad s < d - k.$$

Wir haben gesehen, dass es Funktionen $u \in L^1_{loc}(\Omega)$ gibt, die nicht schwach differenzierbar auf Ω sind. Das bedeutet, es gibt keine Funktion $w \in L^1_{loc}(\Omega)$, welche der Beziehung

$$\int_\Omega u\, \partial_\gamma \eta \, dx = - \int_\Omega w \eta \, dx \quad \textit{für alle} \quad \eta \in C^\infty_\circ(\Omega)$$

($\gamma \in \{1, \dots, n\}$) genügt. Die linke Seite dieser Gleichung ist für jede Funktion $u \in L^1_{loc}(\Omega)$ wohldefiniert und induziert eine stetige lineare Abbildung auf $C^\infty_\circ(\Omega)$. Es liegt daher nahe, diesen linearen Operator als Ersatz für die fehlende schwache (partielle) Ableitung zu nehmen; diese „Ableitung“ von u bezeichnet man dann als distributionelle (partielle) Ableitung *von u.*
Hinter diesem Begriff steckt das allgemeinere Konzept der Distribution, *was im Prinzip nur ein anderer Begriff für eine stetige lineare Abbildung auf $C^\infty_\circ(\Omega)$ ist. Für $\Omega \subset \mathbb{R}^d$ offen und $\omega \Subset \Omega$ sei*

$$C^\infty_\omega(\Omega) := \bigl\{\eta \in C^\infty_\circ(\Omega);\ \operatorname{spt} \eta \Subset \omega\bigr\}.$$

Definition 6.5 (Distribution)

Sei $\Omega \subset \mathbb{R}^d$ offen. Eine Distribution T auf Ω ist eine Abbildung $T : C^\infty_\circ(\Omega) \to \mathbb{R}$ mit folgenden Eigenschaften.

i) *T ist linear.*

ii) *T ist stetig im folgenden Sinn: Zu jedem $\omega \Subset \Omega$ existiert ein $k = k(\omega) \in \mathbb{N}_0$ und ein $c = c(\omega) \in \mathbb{R}$ derart, dass gilt:*

$$|T\eta| \leq c \sum_{|\gamma| \leq k} \|\partial^\gamma \eta\|_{\infty;\,\omega} \quad \textit{für alle} \quad \eta \in C^\infty_\omega(\Omega).$$

Dabei heißt die kleinste Zahl k mit dieser Eigenschaft die Ordnung der Distribution. Den Raum aller Distributionen auf Ω bezeichnen wir mit $\mathscr{D}(\Omega)$, und *den Unterraum der Distributionen der Ordnung $\leq k \in \mathbb{N}_0$ auf Ω mit $\mathscr{D}^k(\Omega)$.*

Bemerkung 6.6

Aus Bedingung i) in Definition 6.5 folgt i. a. nicht die Stetigkeit von T, da $C_\circ^\infty(\Omega)$ unendlichdimensional ist. Die Bedingung ii) aus der Definition besagt, dass $T\eta$ auf jedem fixierten Kompaktum ω gleichmäßig *durch eine gewisse Zahl von Ableitungen kontrolliert werden kann, vorausgesetzt, dass $\operatorname{spt}\eta \Subset \omega$ ist. Offenbar ist die in ii) geforderte Bedingung gleichbedeutend mit*

$$|T\eta| \leq c \sup_{|\gamma|\leq k} \|\partial^\gamma \eta\|_{\infty;\,\omega} \quad \textit{für alle} \quad \eta \in C_\omega^\infty(\Omega)$$

*(was man in der Literatur auch als Defintion findet). Tatsächlich kann man zeigen (vgl. etwa [*Alt*], § 3.10):*

$$\mathscr{D}(\Omega) = C_\circ^\infty(\Omega)^* \quad \textit{und} \quad \mathscr{D}^k(\Omega) = C_\circ^k(\Omega)^*,$$

wenn man $C_\circ^\infty(\Omega)$ bzw. $C_\circ^k(\Omega)$ mit einer geeigneten Topologie versieht. Man beachte, dass ii) a priori nicht die Stetigkeit von T im Sinne von Definition 3.11 liefert: Zwar wird durch die rechte Seite eine Norm auf $C_\circ^\infty(\Omega)$ erklärt, aber man verlangt die Gültigkeit der Bedingung nur auf den Teilklassen $C_\omega^\infty(\Omega)$ und die Konstante c hängt von dem gewählten ω ab.

Beispiel 6.7

i) Für eine Funktion $u \in L^1_{loc}(\Omega)$ sei $T_u : C_\circ^\infty(\Omega) \to \mathbb{R}$,

$$T_u\eta := \int_\Omega u\eta\, dx.$$

Dann ist T linear und für $\omega \Subset \Omega$ ist

$$|T_u\eta| \leq \int_\omega |u\eta|\, dx \leq \|u\|_{1;\,\omega}\|\eta\|_{\infty;\,\omega} \quad \textit{für alle} \quad \eta \in C_\omega^\infty(\Omega),$$

also $T_u \in \mathscr{D}^0(\Omega)$. Die Distribution T_u heißt die von $u \in L^1_{loc}(\Omega)$ erzeugte reguläre Distribution, und wird auch mit $\langle u, \cdot\rangle$ (nicht zu verwechseln mit einem Skalarprodukt!) bezeichnet. Diese ist nach dem Fundamentallemma 3.22 durch u eindeutig bestimmt.

(ii) Sei μ ein Radon–Maß über Ω. Dann ist

$$\left|\int_\Omega \eta\, d\mu\right| \leq \|\eta\|_\infty\, \mu(\operatorname{spt}\eta) < \infty,$$

so dass durch $T_\mu : C_\circ^\infty(\Omega) \to \mathbb{R}$,

$$T_\mu \eta := \int_\Omega \eta \, d\mu$$

eine Distribution der Ordnung 0 *auf* Ω *erklärt wird. Eine „Umkehrung“ davon stellt der noch folgende Satz 6.12 dar.*

Definition 6.8 (Reguläre/Singuläre Distribution)
Seien $\Omega \subset \mathbb{R}^d$ *offen und* $T : C_\circ^\infty(\Omega) \to \mathbb{R}$ *eine Distribution auf* Ω. *Dann heißt* T *regulär, falls es eine Funktion* $u \in L^1_{loc}(\Omega)$ *gibt, so dass*

$$T\eta = \langle u, \eta \rangle := \int_\Omega u\eta \, dx \quad \text{für alle} \quad \eta \in C_\circ^\infty(\Omega)$$

ist. Sonst heißt T *singulär.*

Beispiel 6.9
Seien $\Omega \subset \mathbb{R}^d$ *offen und* $\xi \in \Omega$. *Dann wird durch* $T_\xi : C_\circ^\infty(\Omega) \to \mathbb{R}$, $\eta \mapsto \eta(\xi)$ *eine Distribution der Ordnung 0 erklärt. Diese ist bekannt als* Dirac–Distribution, *und wird mit* δ_ξ *bezeichnet.*
Durch Betrachtung der Funktionen $\eta_\varepsilon : \Omega \to \mathbb{R}$,

$$\eta_\varepsilon(x) := \begin{cases} \exp\left(\dfrac{\varepsilon^2}{|x-\xi|^2 - \varepsilon^2}\right) & ; \quad x \in B_\varepsilon(\xi) \\ 0 & ; \quad \textit{sonst}, \end{cases}$$

wobei $\varepsilon \in \big(0, \operatorname{dist}(x, \partial\Omega)\big)$ *ist, kann man zeigen, dass* δ_ξ *singulär ist.*

Eine weitere wichtige Charakterisierung von Distributionen liefert der folgende Satz. Dieser sagt im wesentlichen aus, dass der Raum der Distributionen der Ordnung 0 *auf* Ω *mit dem Raum* $\mathscr{M}(\Omega)$ *der Radon–Maße über* Ω *übereinstimmt:*

$$\mathscr{D}^0(\Omega) = C_\circ^0(\Omega)^* = \mathscr{M}(\Omega)$$

ist. (Mehr dazu findet man etwa in [Al], § 2.7 oder [AFP], § 1.4.)

Satz 6.12
Sei $\Omega \subset \mathbb{R}^d$ *offen und sei* $T \in \mathscr{D}^0(\Omega)$. *Dann gibt es* $\mu, \nu \in \mathscr{M}(\Omega)$ *mit*

$$T\eta = \langle \mu, \eta \rangle - \langle \nu, \eta \rangle \quad \text{für alle} \quad \eta \in C_\circ^\infty(\Omega).$$

Wir wollen nun erklären, was wir unter der Ableitung einer Distribution, der sog. Distributionsableitung *verstehen wollen. Zur Motivation betrachten wir eine Funktion $u \in C^\infty(\Omega)$, $\Omega \subset \mathbb{R}^d$ offen. Für jedes $\gamma \in \mathbb{N}_0^d$ ist dann nach der Regel der partiellen Integration*

$$\int_\Omega u\,\partial^\gamma \eta\,dx = (-1)^{|\gamma|} \int_\Omega \partial^\gamma u\,\eta\,dx \quad \textit{für alle} \quad \eta \in C_\circ^\infty(\Omega),$$

wobei die linke Seite für jede Funktion $u \in L^1_{loc}(\Omega)$ wohldefiniert ist, und eine stetige, lineare Abbildung auf $C_\circ^\infty(\Omega)$ induziert. Dies gibt Anlass zu der

Definition 6.13 (Distributionelle Ableitung)
Seien $\Omega \subset \mathbb{R}^d$ offen, $T \in \mathscr{D}(\Omega)$ und $\gamma \in \mathbb{N}_0^d$. Dann ist die distributionelle (partielle) Ableitung $\partial^\gamma T$ von T definiert durch die Distribution

$$\big(\partial^\gamma T\big)\eta := (-1)^{|\gamma|}\,T\big(\partial^\gamma \eta\big) \quad (\eta \in C_\circ^\infty(\Omega)).$$

Demnach ist jede Funktion $u \in L^1_{loc}(\Omega)$ beliebig oft distributionell *(oder im Sinne von Distributionen) differenzierbar mit* distributionellen (partiellen) Ableitungen

$$\langle \partial^\gamma u, \cdot\rangle := \partial^\gamma \langle u, \cdot\rangle \quad (\gamma \in \mathbb{N}_0^d).$$

Ferner gilt offenbar:

$$u \in W^k(\Omega) \quad \Longleftrightarrow \quad u \in L^1_{loc}(\Omega),\ \langle \partial^\gamma u, \cdot\rangle \textit{ regulär für alle } \gamma \in \mathbb{N}_0^d \textit{ mi } |\gamma| \le k.$$

Beispiel 6.14
Seien $\Omega \subset \mathbb{R}^d$ offen und $u := (u_1, \dots, u_d) \in L^1_{loc}(\Omega)^d$. Dann definiert man die distributionelle Divergenz durch die Distribution

$$\langle \operatorname{div} u, \cdot\rangle := \sum_{k=1}^d \langle \partial_k u_k, \cdot\rangle.$$

Für jede Funktion $\eta \in C_\circ^\infty(\Omega)$ ist

$$\langle \operatorname{div} u, \cdot\rangle = -\sum_{k=1}^d \langle u_k, \partial_k \eta\rangle = -\sum_{k=1}^d \int_\Omega u_k\,\partial_k \eta\,dx = -\int_\Omega u \cdot \nabla \eta\,dx.$$

Auf analoge Weise kann man den distributionellen Laplace–Operator einer Funktion $w \in L^1_{loc}(\Omega)$ erklären, und kann zeigen, dass dieser im Falle $w \in W^1(\Omega)$ mit $\langle \operatorname{div} \nabla w, \cdot\rangle$ übereinstimmt — wie es sein sollte (vgl. Übung).

Sei $\Omega \subset \mathbb{R}^d$ offen. Eine Familie (ω_m) $(m \in \mathbb{N})$ offener Mengen in Ω heißt eine Ausschöpfung *von Ω, falls gilt:*

$$\omega_m \Subset \omega_{m+1} \quad \text{für alle} \quad m \in \mathbb{N} \quad \text{und} \quad \Omega = \bigcup_{m \in \mathbb{N}} \omega_m .$$

Beispielsweise wird durch

$$\omega_m := B_m(0) \cap \left\{ x \in \Omega;\ \operatorname{dist}(x, \partial\Omega) > \frac{1}{m} \right\}$$

eine Ausschöpfung (ω_m) von Ω definiert. [a]

Der folgende Satz zeigt insbesondere (Teil iv)), dass die schwache Differenzierbarkeit (ähnlich wie die Differenzierbarkeit im klassischen Sinn) eine lokale Eigenschaft ist. Natürlich darf man den Bogen nicht überspannen, und von „punktweiser" schwacher Differenzierbarkeit sprechen, die es ja nicht geben kann (wir haben es ja mit Äquivalenzklassen von Funktionen zu tun).

Satz 6.15 (Schwache Differenzierbarkeit)

Seien $\Omega \subset \mathbb{R}^d$ offen, $\gamma \in \mathbb{N}_0^d$ und $u \in L^1_{loc}(\Omega)$. Dann sind folgende Aussagen äquivalent:

i) *u besitzt auf Ω eine γ–te schwache (partielle) Ableitung $\partial^\gamma u \in L^1_{loc}(\Omega)$.*

ii) *Für jede offene Menge $\omega \subset \Omega$ hat $u|_\omega$ auf U eine γ–te schwache (partielle) Ableitung mit $\partial^\gamma \big(u|_\omega\big) = \big(\partial^\gamma u\big)|_\omega$ auf ω.*

iii) *Für jede Ausschöpfung (ω_m) von Ω hat $u|_{\omega_m}$ eine γ–te schwache (partielle) Ableiung auf ω_m.*

iv) *Zu jedem $x \in \Omega$ gibt es eine Umgebung $\omega \subset \Omega$, so dass $u|_\omega$ eine γ–te schwache (partielle) Ableitung auf ω besitzt, d. h. u hat lokal auf Ω eine γ–te schwache Ableitung.*

Beweis.

Die Implikationen i) $\Rightarrow$ ii), iii), iv) *sind trivial (man wertet die definierende Integralrelation nur mit Funktionen mit kompakten Träger in einer kleineren offenen Menge aus). Die Richtung* ii) $\Rightarrow$ iii) *ergibt sich durch Spezialisierung. Wir beweisen die noch fehlenden Implikationen:*

Ad iii) $\Rightarrow$ i). *Sei (ω_m) eine Ausschöpfung von Ω. Nach Voraussetzung existiert*

[a] Man beachte, dass man für beschränktes Ω in der Definiton der ω_m den Schnitt mit der Kugel $B_m(0)$ fortlassen kann.

zu jedem $m \in \mathbb{N}$ eine Funktion $w_m \in L^1_{loc}(\omega_m)$ mit

$$\int_{\omega_m} u\, \partial^\gamma \eta\, dx = (-1)^{|\gamma|} \int_{\omega_m} w_m \eta\, dx \quad \text{für alle} \quad \eta \in C^\infty_\circ(\omega_m). \tag{6.6}$$

Nach dem Fundamentallemma 3.22 ist dann aber $w_m = w_{m+1}$ f. ü. auf ω_m für alle $m \in \mathbb{N}$. Für $x \in \omega_m$ sei $w(x) := w_m(x) \in L^1_{loc}(\Omega)$ und sei $\eta \in C^\infty_\circ(\Omega)$ beliebig. Dann existiert ein $m \in \mathbb{N}$, so dass $\operatorname{spt}\eta \Subset \omega_m$ ist, woraus mit (6.6) *die Behauptung folgt.*

Ad iv) $\Rightarrow$ i). *Schreibe $\Omega = \bigcup_m \omega_m$ mit offenen Mengen ω_m, auf denen $\partial^\gamma u$ existiert. Zunachst erhält man zu $x \in \Omega$ eine Umgebung $V(x)$ auf der u schwach difffabr ist und $\Omega = \bigcup_x V(x)$ und wählt abzählbar viele Punkte x_m mit $\Omega = \bigcup_m V(x_m)$. Weiter sei $(\eta_k) \subset C^\infty_\circ(\Omega)$ eine* Zerlegung der Eins *(auch: $C^\infty_\circ$–*Zerlegung der Eins*) für Ω bzgl. (ω_m), i. e.:*

$$0 \le \eta_k \le 1 \text{ in } \Omega \text{ für alle } k \in \mathbb{N}. \tag{6.7}$$

$$\text{Für alle } k \in \mathbb{N} \text{ gibt es ein } m_k \in \mathbb{N} \text{ mit } \operatorname{spt}\eta_k \Subset \omega_{m_k}. \tag{6.8}$$

$$\sharp\big\{k \in \mathbb{N};\ \operatorname{spt}\eta_k \cap \omega \neq \emptyset\big\} < \infty \text{ für alle } \omega \Subset \Omega. \tag{6.9}$$

$$\sum_{k\in\mathbb{N}} \eta_k \equiv 1 \text{ in } \Omega. \tag{6.10}$$

Dabei sind in (6.10) *immer nur endlich viele Summanden $\neq 0$ sind. (Vgl. dazu etwa [Ad], Thm. 3.14 oder [Yo], § I.12.)*
Ist nun $\eta \in C^\infty_\circ(\Omega)$ beliebig, so ist nach (6.9) $\operatorname{spt}\eta \cap \operatorname{spt}\eta_k \neq \emptyset$ *nur für endlich viele $k \in \mathbb{N}$ und wegen* (6.10) *folgt*

$$\int_\Omega u\, \partial^\gamma \eta\, dx = \int_\Omega u\, \partial^\gamma \left(\sum_{k\in\mathbb{N}} \eta\eta_k \right) dx = \sum_{k\in\mathbb{N}} \int_\Omega u\, \partial^\gamma (\eta\eta_k)\, dx. \tag{6.11}$$

Nun ist $\operatorname{spt}(\eta\eta_k) \Subset \omega_{m_k}$ *für ein $m_k \in \mathbb{N}$ (gem.* (6.8)*) und es existiert eine Funktion $w_k \in L^1_{loc}(\omega_{m_k})$ derart, dass gilt:*

$$\int_{\omega_{m_k}} u\, \partial^\gamma \eta\, dx = (-1)^{|\gamma|} \int_{\omega_{m_k}} w_k \eta\, dx \quad \text{für alle} \quad \eta \in L^\infty_o(\omega_{m_k}).$$

Mit (6.11) *folgt, dass u auf Ω eine γ–te schwache (partielle) Ableitung besitzt, nämlich $w := \sum_k w_k \eta_k \in L^1_{loc}(\Omega)$.* □

Wir geben noch die folgenden Charakterisierungen der schwachen Differenzierbarkeit an, die den Zusammenhang zu den zu Beginn des Kapitels hergeleiteten

Zugängen herstellen. Insbesondere sei an die Differenzenquotienten $\Delta_\gamma^h u$ *aus* (6.3) *erinnert.*

Satz 6.16 (Schwache Ableitung)
Seien $\Omega \subset \mathbb{R}^d$ *offen,* $\gamma \in \{1, \dots, d\}$ *und* $u, w \in L^1_{loc}(\Omega)$. *Dann sind äquivalent:*

i) w *ist die* γ*–te schwache (partielle) Ableitung von* u *auf* Ω*, also* $w = \partial_\gamma u$.

ii) Es gibt eine Folge $(u_n) \subset C^\infty(\Omega)$ *mit*

$$u_n \xrightarrow{n} u \quad \textit{und} \quad \partial_\gamma u_n \xrightarrow{n} w \textit{ in } L^1_{loc}(\Omega).$$

iii) Für $0 < |h| \ll 1$ *ist* $\Delta_\gamma^h u$ $\mathcal{L}^d$*–f. ü. in* Ω *definiert und es strebt*

$$\Delta_\gamma^h u \xrightarrow{h \to 0} w \textit{ in } L^1_{loc}(\Omega).$$

Beweis.
Wir zeigen hier nur, dass ii) und iii) hinreichend für i) sind, die Umkehrung i) $\Rightarrow$ ii) *beispielsweise wird sich später aus viel allgemeineren Sätzen (Konstruktion glatter Approximationen; GdV II) ergeben.*
Ad ii) $\Rightarrow$ i). *Sei* $\eta \in C^\infty_\circ(\Omega)$ *beliebig. Dann wird mit partieller Integration (Satz von Gauß)*

$$\int_\Omega u_n \, \partial_\gamma \eta \, dx = - \int_\Omega \partial_\gamma u_n \, \eta \, dx,$$

mit $n \to \infty$*, unter Beachtung von* $\operatorname{spt} \eta \Subset \Omega$ *und* $\eta, \partial_\gamma \eta \in L^\infty(\Omega)$*, also:*

$$\int_\Omega u \, \partial_\gamma \eta \, dx = - \int_\Omega \partial_\gamma u \, \eta \, dx.$$

Ad iii) $\Rightarrow$ i). *Seien* $\eta \in C^\infty_\circ(\Omega)$ *und* $|h| \ll 1$ *(so klein, dass* $x + he_\gamma \in \Omega$ *liegt für jedes* $x \in \operatorname{spt} \eta$*). Dann hat* $\eta \, \Delta_\gamma^h u$ *kompakten Träger in* Ω *und es gilt (nachrechnen!):*

$$\int_\Omega \Delta_\gamma^h u \, \eta \, dx = - \int_\Omega u \, \Delta_\gamma^{-h} \eta \, dx.$$

Da η *glatt ist, strebt natürlich* $\Delta_\gamma^{-h} \eta \xrightarrow{h \to 0} \partial_\gamma \varphi$ *gleichmäßig in* Ω*. Dies zusammen mit der Voraussetzung liefert i).* □

Bemerkung 6.17

i) Nach Satz 6.16 ist eine Funktion $u \in L^1_{loc}(\Omega)$ *genau dann in* $W^1(\Omega)$*, wenn eine der gleichwertigen Bedingungen ii) oder iii) aus dem Satz erfüllt ist.*

ii) Analoge Aussagen gelten auch für höhere schwache Ableitungen.

Schlußbemerkung.
Natürlich kann man die Begriffsbildungen in diesem § — wie teilweise schon erwähnt — sinngemäß auf vektorwertige Funktionen $u : \Omega \to \mathbb{R}^D$ *übertragen. Insbesondere gelten die oben gemachten Aussagen über die schwache Differenzierbarkeit auch für Funktionen* $u \in W^k(\Omega)^D$.

§ 7. Absolutstetige Funktionen

Wenn nichts anderes gesagt wird, bezeichne im Folgenden I ein offenes Intervall in $\mathbb{R}$ und $u : I \to \mathbb{R}$ eine punktweise (also nicht durch Vertreter) erklärte Funktion. Ferner bezeichnen wir mit $\fint_J u\, d\mathcal{L}^1$ den Mittelwert *von u über $J \subset I$, d. h.*

$$\fint_J u(t)\, d\mathcal{L}^1(t) := \frac{1}{\mathcal{L}^1(J)} \int_J u(t)\, d\mathcal{L}^1(t).$$

Definition 7.1 (Absolutstetige Funktion)
Eine Funktion $u : I \to \mathbb{R}$ heißt absolutstetig, falls u das unbestimmte Lebesgue–Integral einer Funktion $v \in L^1_{loc}(I)$ ist, d. h. es gibt Konstanten $a \in I$ und $c \in \mathbb{R}$ mit

$$u(x) = \int_{[a,x]} v(t)\, d\mathcal{L}^1(t) + c$$

für alle $x \in I$. Die Klasse der absolutstetigen Funktionen über I wird mit $AC(I)$ bezeichnet.

Für die Beweise der folgenden Aussagen sei auf [HS], § 18 verwiesen.

Lemma 7.2
Sei $u \in AC(I)$. Dann gilt:

i) $u \in C^0(I)$.

ii) *u ist $\mathcal{L}^1$–f. ü. in I (im klassischen Sinne) differenzierbar mit $u' = v$ $\mathcal{L}^1$–f. ü. in I. Ferner ist*

$$\frac{u(x+h) - u(x)}{h} = \fint_{[x,x+h]} v(t)\, d\mathcal{L}^1(t)$$

für alle $x \in I$ und $h \in \mathbb{R} \setminus \{0\}$ genügend klein. [a]

[a] Allgemein sind für jede Funktion $v \in L^1_{loc}(I)$ $\mathcal{L}^1$–f. a. $x \in I$ *Lebesgue–Punkte* von v, d. h. aus dem Mittelwert von v läßt sich für diese x der Funktionswert rekonstruieren:

$$v(x) = \lim_{h \to 0} \fint_{[x,x+h]} v(t)\, d\mathcal{L}^1(t).$$

Diese Aussage ist als *Differentiationssatz von Lebesgue* bekannt (siehe [HS], Lem. 18.4). Die Menge aller Lebesgue–Punkte von u heißt die *Lebesgue–Menge* für u. Da $\mathcal{L}^1$–f. a. $x \in I$ Lebesgue–Punkte von u sind, hat die Lebesgue–Menge $J \subset I$ für u volles Maß, d. h. es ist $\mathcal{L}^1(I \setminus J) = 0$.

Bemerkung 7.3

Es gibt viele stetige Funktionen, die $\mathcal{L}^1$–f. ü. (im klassischen Sinne) differenzierbar sind (dazu zählen etwa alle monotonen Funktionen), die aber nicht absolutstetig sind. Die f. ü.–Existenz der Ableitung u' einer Funktion reicht nicht aus, um einen sinnvollen Zusammenhang zwischen u und u' herzustellen (vgl. Übung).

Lebesgue konstruierte eine Funktion $u : [0,1] \to \mathbb{R}$ mit den Eigenschaften

- *i)* u *ist stetig und monoton wachsend,*
- *ii)* $u' = 0$ $\mathcal{L}^1$–*f. ü. in* $[0,1]$,
- *iii)* $u\big([0,1]\big) = [0,1]$,

die aber nicht zur Klasse $AC[0,1]$ gehört.

Satz 7.4

Folgende Aussagen äquivalent:

- *i)* $u \in AC(I)$.
- *ii)* $u \in C^0(I)$ *hat die Eigenschaft, dass zu jedem* $\varepsilon > 0$ *ein* $\delta = \delta(\varepsilon) > 0$ *existiert derart, dass für jedes* $J \subset I$ *mit* $\mathcal{L}^1(J) < \delta$
$$\mathcal{L}^1\big(u(J)\big) < \varepsilon$$
ausfällt.
- *iii) Zu jedem* $\varepsilon > 0$ *gibt es ein* $\delta = \delta(\varepsilon) > 0$ *derart, dass*
$$\sum_{\nu=1}^{n} \big|u(s_\nu) - u(t_\nu)\big| < \varepsilon$$
ausfällt für alle $s_\nu, t_\nu \in I$ *mit* $s_\nu \le t_\nu \le s_{\nu+1}$ *für alle* $\nu \in \{1, \dots, n-1\}$, $s_n \le t_n$ *und* $\sum_{\nu=1}^{n} |s_\nu - t_\nu| < \delta$.
- *iv) Es gibt eine Folge* $(u_m) \subset C^\infty(I)$ *mit* $u_m \xrightarrow{m} u$ *lokal gleichmäßig auf* I *und* $u'_m \xrightarrow{m} v$ *in* $L^1_{loc}(I)$ *mit einer Funktion* $v \in L^1_{loc}(I)$ *wie in Definition 7.1.*

Aus Satz 7.4 folgt relativ leicht, dass jede absolutstetige Funktion lokal von beschränkter Variation ist, und daher als Differenz monotoner Funktionen geschrieben werden kann.

Ferner liefert Satz 7.4 iv) in Verbindung mit Satz 6.16 (ii) $\Rightarrow$ i)) offenbar

$$AC(I) \subset W^1(I),$$

d. h. jede auf I absolutstetige Funktion ist schwach differenzierbar, und die schwache Ableitung wird von der f. ü. existierenden punktweisen (klassischen) Ableitung erzeugt. Es gilt auch die Umkehrung:

Satz 7.5
Sei $I \subset \mathbb{R}$ offen. Dann ist $AC(I) = W^1(I)$, d. h. jede W^1–Funktion hat genau einen Vertreter der Klasse $AC(I)$.

Beweis.
Es ist noch $W^1(I) \subset AC(I)$ zu zeigen. Sei also $u \in W^1(I)$ und sei $(u_m) \subset C^\infty(I)$ eine Folge mit

$$u_m \xrightarrow{m} u \quad \text{und} \quad u'_m \xrightarrow{m} v \qquad \text{in } L^1_{loc}(I),$$

wobei $v \in L^1_{loc}(I)$ die schwache Ableitung von u bezeichne (eine solche Folge existiert gem. Satz 6.16). Nach Wahl eines Vertreters für u (den wir wieder mit u bezeichnen) und evtl. Übergang zu einer Teilfolge von (u_m) (welche wieder mit (u_m) bezeichnet werde) haben wir auch

$$u_m \xrightarrow{m} u \text{ punktweise } \mathcal{L}^1\text{–f. ü. in } I.$$

Offenbar ist

$$u_m(x) = u_m(a) + \int_{[a,x]} u'_m(t)\, d\mathcal{L}^1(t)$$

für alle $a, x \in I$. Mit $c := \lim_m u_m(a)$ folgt daraus

$$u(x) = c + \int_{[a,x]} u'(t)\, d\mathcal{L}^1(t) \quad \text{für } \mathcal{L}^1\text{–f. a. } x \in I.$$

Damit ist aber $\tilde{u} : I \to \mathbb{R}$ mit

$$\tilde{u}(x) := c + \int_{[a,x]} u'(t)\, d\mathcal{L}^1(t)$$

ein Vertreter der Klasse $AC(I)$ für u. □

Bemerkung 7.5

i) Im Fall $d \geq 2$ erhält man folgende (deutlich schwächere) Verallgemeinerung von 7.4:
Eine Funktione $u \in L^1_{loc}(\Omega)$, $\Omega \subset \mathbb{R}^d$, ist genau dann schwach differen-

zierbar in Richtung $\gamma \in \{1, ..., d\}$, *wenn eine* $v_\gamma \in L^1_{loc}(\Omega)$ *existiert mit*

$$\begin{aligned} u(x) &= u(x_1, .., x_\gamma, ..., x_d) \\ &= c_\gamma + \int_{[a_\gamma, x_\gamma]} v_\gamma(x_1, ..., x_{\gamma-1}, t, x_{\gamma+1}, ..., x_d)\, d\mathcal{L}^1(t) \end{aligned}$$

für $\mathcal{L}^{d-1}$*-f.a.* $(x_1, ..., x_{\gamma-1}, x_{\gamma+1}, ..., x_d)$ *mit passenden Konstanten* $a_\gamma, c_\gamma \in \mathbb{R}$.

ii) Für $d \geq 2$ *ist es allgemein nicht richtig, dass ein geeigneter Vertreter der schwach differenzierbaren Funktion* u *f.ü. im klassischen Sinne differenzierbar ist; eine solche Aussage trifft nur im Ausnahmefall* $d = 1$ *zu.*

§ 8. Sobolev–Räume

Die Räume $W^k(\Omega)$ der k–mal auf $\Omega \subset \mathbb{R}^d$ schwach differenzierbaren Funktionen, die wir im vorangegangenen Kapitel kennengelernt haben, sind ersichtlich lineare Räume. In diesem § wird es darum gehen, geeignete Unterräume auszuwählen, die wir mit einer vollständigen Norm versehen können. Dies geht so, dass man von Funktionen und ihren schwachen Ableitungen (bis zur Ordnung k) die Zugehörigkeit zu den Räumen $L^p(\Omega)$ mit einem $p \in [1, \infty]$ verlangt.
Unsere Überlegungen werden insbesondere den bereits in § 4 (vgl. (4.4)*) beschriebenen Hilbert–Raum (dort mit $\mathcal{W}$ bezeichnet) liefern, und damit das zu Beginn von § 6 beschriebene Problem lösen (vgl. auch § 4), welches uns dazu veranlasst hatte, den Differenzierbarkeitsbegriff abzuschwächen.*

Definition 8.1 (Sobolev–Raum)
Seien $\Omega \subset \mathbb{R}^d$ offen, $1 \le p \le \infty$ und $k \in \mathbb{N}_0$. Dann heißt der lineare Raum

$$W^{k,p}(\Omega) := \Big\{ u \in W^k(\Omega);\ \partial^\gamma u \in L^p(\Omega) \text{ für alle } \gamma \in \mathbb{N}_0^k \text{ mit } |\gamma| \le k \Big\}$$

der Sobolev–Raum mit Differenzierbarkeitsstufe k und Integrabilitätsexponent p. (Insbesondere ist $W^{0,p}(\Omega) = L^p(\Omega)$.) Auf $W^{k,p}(\Omega)$ betrachtet man die Norm:

$$\|u\|_{k,p} := \|u\|_{\Omega;k,p} := \begin{cases} \left(\sum\limits_{|\gamma| \le k} \left\|\partial^\gamma u\right\|_p^p \right)^{\frac{1}{p}} & ; \quad p < \infty \\ \max_{|\gamma| \le k} \left\|\partial^\gamma u\right\|_\infty & ; \quad p = \infty. \end{cases}$$

Die Elemente von $W^{k,p}(\Omega)$ heißen Sobolev–Funktionen. [a]

---

[a] Man erinnere sich daran, dass es sich bei den Elementen von $W^{k,p}(\Omega)$, die ja insbesondere Elemente von $L^1(\Omega)$ sind, eigentlich nicht um Funktionen, sondern um Äquivalenzklassen von Funktionen handelt (vgl. § 3). In der Literatur findet man häufig auch die Bezeichnungen $W_p^k(\Omega)$, $H^{k,p}(\Omega)$ oder $H_p^k(\Omega)$, wobei die „H–Notation“ einen tieferliegenden Grund hat, wie wir später noch sehen werden (Satz von Meyers und Serrin). Ferner werden die Räume $W^{k,2}(\Omega)$ auch häufig mit $H^k(\Omega)$ bezeichnet, weil diese dadurch ausgezeichnet sind, dass sie Hilbert–Raum sind (s. u.).

Bemerkung 8.2

i) Dass durch $\|\cdot\|_{k,p}$ tatsächlich eine Norm auf $W^{k,p}(\Omega)$ erklärt ist, zeigen einfache Rechnungen.

ii) Zu $\|\cdot\|_{k,p}$ äquivalente Normen auf $W^{k,p}(\Omega)$ werden gegeben durch:

$$\sum_{|\gamma|\le k} \left\|\partial^\gamma u\right\|_p \quad bzw. \quad \max_{|\gamma|\le k} \left\|\partial^\gamma u\right\|_p$$

für $u \in W^{k,p}(\Omega)$ (mit $1 \le p \le \infty$). Betrachet man die dadurch erklärten Normen, so ändern sich Konvergenzaussagen, Vollständigkeit oder andere topologische Begriffe natürlich nicht. Die von uns bevorzugte Norm auf $W^{k,p}(\Omega)$ hat den entscheidenden Vorteil, dass sie für $p = 2$ von dem durch

$$\langle u, v\rangle_k := \langle u, v\rangle_{W^{k,2}(\Omega)} := \sum_{|\gamma|\le k} \int_\Omega \partial^\gamma u\, \partial^\gamma v\, dx$$

gegebenen Skalarprodukt auf $W^{k,2}(\Omega)$ herrührt.

Wir führen noch die folgende Bezeichnung ein, die uns später zum „Randverhalten“ von Sobolev–Funktionen führen wird. Zunächst ist völlig unklar, ob und wie man Sobolev–Funktionen „Randwerte“ zuordnen soll.

Definition 8.3

Seien $\Omega \subset \mathbb{R}^d$ offen, $1 \le p \le \infty$ und $k \in \mathbb{N}_0$. Dann bezeichnen wir mit $\mathring{W}^{k,p}(\Omega)$ den Normabschluss des Raumes $C^\infty_\circ(\Omega)$ in $W^{k,p}(\Omega)$ (also den Abschluss bzgl. der Norm $\|\cdot\|_{k,p}$):

$$\mathring{W}^{k,p}(\Omega) := \overline{C^\infty_\circ(\Omega)}^{\|\cdot\|_{k,p}},$$

d. h. es ist $u \in \mathring{W}^{k,p}(\Omega)$ genau dann, wenn es eine Folge $(\eta_n) \subset C^\infty_\circ(\Omega)$ gibt mit

$$\partial^\gamma \eta_n \xrightarrow{n} \partial^\gamma u \ \text{in}\ L^p(\Omega)$$

für alle $\gamma \in \mathbb{N}_0^d$ mit $|\gamma| \le k$. [b]

Man kann zeigen, dass wie erwartet $\mathring{W}^{k,p}(\Omega) \subset W^{k,p}(\Omega)$ und $\mathring{W}^{0,p}(\Omega) = W^{0,p}(\Omega) = L^p(\Omega)$ ist; ferner ist $\mathring{W}^{k,p}(\mathbb{R}^d) = W^{k,p}(\mathbb{R}^d)$.

Entsprechend zu den Lebesgue–Räumen, erklären wir auch für die Sobolev–Räume lokale sowie vektorielle Versionen: Seien $\Omega \subset \mathbb{R}^d$ offen, $1 \le p \le \infty$

[b] In der Literatur findet man häufig auch die Bezeichnungen $\mathring{W}^k_p(\Omega)$, $W^{k,p}_\circ(\Omega)$, $W^k_{p\,\circ}(\Omega)$ sowie entsprechende „H–Notationen“. Ferner sind die Bezeichnungen $\mathring{H}^k(\Omega)$ bzw. $H^k_\circ(\Omega)$ für die Hilbert–Räume $\mathring{W}^{k,2}(\Omega)$ gebräuchlich.

und $k \in \mathbb{N}_0$. *Dann setzt man:*

$$W^{k,p}_{loc}(\Omega) := \Big\{ u \in W^k(\Omega);\ \partial^\gamma u \in L^p_{loc}(\Omega) \text{ für alle } \gamma \in \mathbb{N}_0^k \text{ mit } |\gamma| \le k \Big\}.$$ [c]

Insbesondere ist $W^k(\Omega) = W^{k,1}_{loc}(\Omega)$ *(man erinnere sich an Satz 6.15, wonach eine* L^1_{loc}*–Funktion genau dann schwach differenzierbar auf* Ω *ist, wenn dies lokal auf* Ω *der Fall ist).*
Für vektorwertige Funktionen $u := (u^1, \ldots, u^D) : \Omega \to \mathbb{R}^D$ $(D \in \mathbb{N})$ *definieren wir:*

$$W^{k,p}(\Omega)^D := W^{k,p}(\Omega, \mathbb{R}^D) := \left\{ u \in L^p(\Omega)^D;\ \begin{array}{c} u^\nu \in W^{k,p}(\Omega) \text{ für alle} \\ \nu = 1, \ldots, D \end{array} \right\}$$

und versehen diesen Raum mit der Norm:

$$\|u\|_{k,p} := \|u\|_{\Omega;k,p} := \begin{cases} \left(\sum\limits_{\nu=1}^{D} \left\| u^\nu \right\|^p_{k,p} \right)^{\frac{1}{p}} & ;\quad p < \infty \\ \max_{\nu=1}^{D} \left\| u^\nu \right\|_{k,\infty} & ;\quad p = \infty. \end{cases}$$

Schließlich erklärt man wie oben auch noch lokale Versionen $W^{k,p}_{loc}(\Omega)^D$ *der Räume* $W^{k,p}(\Omega)^D$ *sowie den Raum* $\mathring{W}^{k,p}(\Omega)^D$. *(Wie dies im Detail auszusehen hat, dürfte ersichtlich sein).*

Satz 8.4
Seien $\Omega \subset \mathbb{R}^d$ *offen,* $1 \le p \le q \le \infty$ *und* $k, l \in \mathbb{N}_0$ *mit* $k \ge l$. *Dann gilt:*

i) Die Einbettugen
$$W^{k,p}(\Omega) \hookrightarrow W^{l,p}(\Omega) \hookrightarrow L^p(\Omega)$$
sind linear und stetig.

ii) Ist $\mathcal{L}^d(\Omega) < \infty$, *so ist die Einbettung*
$$W^{k,q}(\Omega) \hookrightarrow W^{k,p}(\Omega)$$
stetig.

iii) Für $p < \infty$ *liegt* $W^{k,p}(\Omega)$ *(bzgl. der Norm* $\|\cdot\|_p$*) dicht in* $L^p(\Omega)$.

Beweis.
Die Aussage i) ist völlig trivial; bei der zweiten Einbettung „vergisst" man einfach, dass die Funktionen auch schwache Ableitungen in $L^p(\Omega)$ *besitzen. Die*

[c] Auch hier sind andere Bezeichnungen gebräulich: $W^k_{p,loc}(\Omega)$ bzw. entsprechende „H–Notationen".

Aussage in ii) folgt sofort aus der Hölder–Ungleichung (Lem. 3.6 i)). Bleibt iii) nachzuweisen: Nach Satz 3.21 liegt sogar $C_\circ^\infty(\Omega)$ *dicht in* $L^p(\Omega)$, *aus*

$$C_\circ^\infty(\Omega) \subset \mathring{W}^{k,p}(\Omega) \subset W^{k,p}(\Omega) \subset L^p(\Omega)$$

folgt dann natürlich die Dichtheit von $W^{k,p}(\Omega)$ *in* $L^p(\Omega)$. □

Bemerkung 8.5

Für $p = \infty$ *ist iii) in Satz 8.4 falsch: Wie später gezeigt wird, besteht* $W^{k,\infty}(\Omega)$ *für* $k \geq 1$ *aus stetigen Funktionen. Läge also* $W^{k,\infty}(\Omega)$ *dicht in* $L^\infty(\Omega)$, *so hieße das: Jede Funktion* $u \in L^\infty(\Omega)$ *läßt sich bzgl.* $\|\cdot\|_\infty$ *durch stetige Funktionen approximieren. Es gibt aber bekanntlich beschränkte Funktionen* (L^∞–*Funktionen), die man nicht gleichmäßig durch stetige Funktionen approximieren kann. (Etwa solche, die keinen stetigen Vertreter besitzen.)*

Mit der Einbettung $W^{k,p}(\Omega) \hookrightarrow L^p(\Omega)$ $(p < \infty)$ *aus dem obigen Satz läßt sich nichts anfangen, da* $W^{k,p}(\Omega)$ *gem. Teil iii) von Satz 8.4 kein abgeschlossener Teilraum von* $L^p(\Omega)$ *ist. (Andernfalls folgte ja* $W^{k,p}(\Omega) = L^p(\Omega)$. *Wie wir wissen, gibt es aber* L^p–*Funktionen, die nicht einmal schwach differenzierbar sind.) Andererseits übertragen sich die funktionalanalytischen Eigenschaften eines normierten Raumes (hier* L^p*) nur auf abgeschlossene Unterräume. Aus diesem Grund konstruieren wir eine andere Einbettung für* $W^{k,p}(\Omega)$, *welche den Ableitungseigenschaften Rechnung trägt: Seien* $k \in \mathbb{N}$ *und* $1 \leq p \leq \infty$ *fixiert. Für*

$$D := \sharp\big\{\gamma \in \mathbb{N}_0^d;\ |\gamma| \leq k\big\} \tag{8.1}$$

definieren wir eine Einbettung $\Phi : W^{k,p}(\Omega) \to L^p(\Omega)^D$ *durch*

$$\Phi u := \big(\partial^\gamma u\big)_{|\gamma|\leq k}, \tag{8.2}$$

wobei $\big(\partial^\gamma u\big)_{|\gamma|\leq k}$ *o. E. in Zeilenform angeordnet sei. Für* $k = 1$ *hat man dann beispielsweise*

$$\Phi u = \big(u, \partial_1 u, \ldots, \partial_d u\big) = \big(u, \nabla u\big).$$

Offenbar ist $\|u\|_{k,p} = \|\Phi u\|_p$ *für alle* $u \in W^{k,p}(\Omega)$, *und es gilt:*

Satz 8.6

Die durch (8.2) *definierte Einbettung* $\Phi : W^{k,p}(\Omega) \hookrightarrow L^p(\Omega)^D$ *ist eine lineare Isometrie und der Unterraum* $W^{k,p}(\Omega)$ *ist abgeschlossen in* $L^p(\Omega)^D$.

Beweis.

Bis auf die Abgeschlossenheit sind alle Aussagen trivial. Wir zeigen diese exemplarisch für $k = 1$ *und überlassen dem Leser den allgemeinen Fall. Sei also*

$(u_m) \subset W^{1,p}(\Omega)$ eine Folge, so dass $\Phi u_m = (u_m, \nabla u_m)$ in $L^p(\Omega)^{1+d}$ konvergiert, also

$$u_m \xrightarrow{m} : u \text{ in } L^p(\Omega) \quad \text{und} \quad \nabla u_m \xrightarrow{m} : v \text{ in } L^p(\Omega)^d$$

mit Funktionen $u \in L^p(\Omega)$ und $v \in L^p(\Omega)^d$. Wir müssen zeigen, dass wieder $u \in W^{1,p}(\Omega)$ ist. Offensichtlich muss dann gerade $\nabla u = v$ erfüllt sein. Sei dazu $\eta \in C_\circ^\infty(\Omega)^d$ beliebig. Dann ist

$$\int_\Omega u \operatorname{div} \eta \, dx = \lim_m \int_\Omega u_m \operatorname{div} \eta \, dx = -\lim_m \int_\Omega \nabla u_m \cdot \eta \, dx = -\int_\Omega v \cdot \eta \, dx,$$

also ist u schwach differenzierbar mit schwacher Ableitung $\nabla u = v$, womit die Behauptung folgt. □

Damit übertragen sich nun die funktionalanalytischen Eigenschaften des Raumes $L^p(\Omega)^D$ (mit D wie in (8.1)) auf den Raum $W^{k,p}(\Omega)$. Man bekommt:

Korollar 8.7

Für alle $k \in \mathbb{N}_0$ und $1 \le p \le \infty$ ist $W^{k,p}(\Omega)$ ein Banach–Raum (bzgl. der Norm $\|\cdot\|_{k,p}$). Ferner sind $W^{k,2}(\Omega)$ sowie $\mathring{W}^{k,2}(\Omega)$ Hilbert–Räume bzgl. dem Skalarprodukt $\langle\cdot,\cdot\rangle_k$ aus Bem. 8.2.

Wir wollen nun klären, was schwache Konvergenz in $W^{k,p}(\Omega)$ bedeutet: Sei dazu eine Folge $(u_m) \subset W^{k,p}(\Omega)$ vorgegeben. Per Definition bedeutet $u_m \overset{m}{\rightharpoonup}: u$ für eine Funktion $u \in W^{k,p}(\Omega)$, dass

$$\varphi(u_m) \xrightarrow{m} \varphi(u) \quad \text{für alle} \quad \varphi \in W^{k,p}(\Omega)^*$$

strebt, was für konkrete Anwendungen aber zu unhandlich ist. Vermöge der Einbettung Φ gem. (8.2) lässt sich jedoch zeigen:

$$u_m \overset{m}{\rightharpoonup} u \text{ in } W^{k,p}(\Omega) \quad \Longleftrightarrow \quad \partial^\gamma u_m \overset{m}{\rightharpoonup} \partial^\gamma u \text{ in } L^p(\Omega) \text{ für alle } |\gamma| \le k,$$

so dass also schwache Konvergenz in $W^{k,p}(\Omega)$ komponentenweise schwache Konvergenz in $L^p(\Omega)^D$ (mit D gem. (8.1)) bedeutet. Als Anwendung des Satzes von Riesz für Lebesgue–Räume (Satz 3.17) ergibt sich (vgl. auch (5.9)):

Satz 8.8

Seien $k \in \mathbb{N}_0$, $1 \le p < \infty$ und $u_m, u \in W^{k,p}(\Omega)$. Dann sind äquivalent:

i) $u_m \overset{m}{\rightharpoonup} u$ *in* $W^{k,p}(\Omega)$.

ii) Für alle $\varphi \in L^{p'}(\Omega)$ (mit p' gem. (3.7)) gilt:

$$\int_\Omega \partial^\gamma u_m \, \varphi \, dx \xrightarrow{m} \int_\Omega \partial^\gamma u \, \varphi \, dx.$$

Für $1 < p < \infty$ erhalten wir aus der Reflexivität von L^p (Satz 5.11) die Reflexivität von $W^{k,p}$, und damit wg. Satz 5.17 das folgende — für die Variationsrechnung wichtige — schwache Auswahlprinzip, das wegen der Irreflexivität von L^1 und L^∞ in den Räumen $W^{1,1}$ und $W^{1,\infty}$ nicht gilt. Diese sind umgekehrt nach dem Satz von Eberlein–Shmulyan (Satz 5.19) auch selbst irreflexiv.

Satz 8.9 (Schwaches Auswahlprinzip in $W^{k,p}$)
Seien $k \in \mathbb{N}_0$, $1 < p < \infty$ und $(u_m) \subset W^{k,p}(\Omega)$ eine beschränkte Folge, d. h. es sei $\sup_m \left\| u_m \right\|_{k,p} < \infty$. Dann gibt es eine Teilfolge $\left(u_{m_k}\right)$ von (u_m) und eine Funktion $u \in W^{k,p}(\Omega)$, so dass gilt:

$$u_{m_k} \overset{k}{\rightharpoonup} u \quad \text{in } W^{k,p}(\Omega).$$

Beweis.
Nach Voraussetzung ist die Folge (u_m) beschränkt in $L^p(\Omega)^D$ mit D gem. (8.1), so dass es nach dem schwachen Auswahlprinzip für L^p–Räume (Kor. 5.18) eine Teilfolge $\left(u_{m_k}\right)$ und eine Funktion $v \in L^p(\Omega)^D)$ gibt mit $u_{m_k} \overset{k}{\rightharpoonup} v$ in $L^p(\Omega)^D)$. Da aber $W^{k,p}(\Omega)$ normabgeschlossen in $L^p(\Omega)^D)$ ist (Satz 8.6), ist $W^{k,p}(\Omega)$ erstrecht schwach abgeschlossen in $L^p(\Omega)^D$ (Satz 5.13), so dass v zu $W^{k,p}(\Omega)$ gehört, und damit die Folge $\left(u_{m_k}\right)$ in $W^{k,p}(\Omega)$ schwach gegen v konvergiert. □

Bemerkung 8.10
Nach Definition ist $\mathring{W}^{k,p}(\Omega)$ normabgeschlossen, und daher auch schwach abgeschlossen (Satz 5.13) in $W^{k,p}(\Omega)$, d. h. ist $(u_m) \subset \mathring{W}^{k,p}(\Omega)$ und strebt

$$u_m \overset{m}{\rightharpoonup}: u \quad \text{in } W^{k,p}(\Omega)$$

mit einer Funktion $u \in W^{k,p}(\Omega)$, so ist auch $u \in \mathring{W}^{k,p}(\Omega)$.

Zum Abschluß dieses § betrachten wir für Funktionen $w : \Omega \to \mathbb{R}$ ($\Omega \subset \mathbb{R}^d$ ein beschränktes Gebiet mit genügend glattem Rand) und $p \in (1, \infty)$ das Funktional

$$J[w] := \int_\Omega |\nabla w|^p \, dx,$$

das offenbar für Funktionen $w \in W^{1,p}(\Omega)$ wohldefiniert ist. Mit den bisher erworbenen Kenntnissen sind wir in der Lage die Existenz eines eindeutugen Minimierers (bei gegebenen Randwerten) nachzuweisen. Dies stellt das finale

Ziel der Vorlesung dar.
Obiges Funktional soll in der Teilklasse

$$\mathscr{C} := \left\{w \in W^{1,p}(\Omega);\ w - u_0 \in \mathring{W}^{1,p}(\Omega)\right\} =: u_0 + \mathring{W}^{1,p}(\Omega)\big\},$$

mit einer fixierten Funktion $u_0 \in W^{1,p}(\Omega)$*, minimiert werden.*[d] *Wesentliche Hilfsmittel zum Beweis werden in den folgenden Lemmata erarbeitet:*

Lemma 8.11
Sei X *ein normierter Raum,* (x_k) *konvergiere schwach gegen ein* $x \in X$*. Dann gibt es eine Folge* (y_k)*, so dass* y_k *in der konvexen Hülle von* $\{x_l,\, l \geq k\}$ *liegt mit* $y_k \to x$*,* $k \to \infty$*.*

Beweis.
Sei M_i *definiert als die konvexe Hülle von* $\{x_i, x_{i+1}, ...\}$*,* $i \in \mathbb{N}$*.*
Nehmen wir an es gebe eine Zahl $\epsilon > 0$ *mit* $\|x - y\| \geq \epsilon$ *für alle* $y \in M_1$*. Wir definieren*

$$L := \left\{z \in X : \exists u \in M_1 \text{ mit } \|z - u\| \leq \frac{\epsilon}{2}\right\}.$$

Dann gilt

1) $L \supset M_1$ *und damit* $\overline{L} \supset M_1$*.*

2) $\overline{L}$ *ist konvex, da* L *konvex ist.*

3) $x \notin \overline{L}$

Die Konvexität von L *sieht man wie folgt: Seien* $z_1, z_2 \in L$ *und* $0 \leq t \leq 1$*. Man wähle* $u_1, u_2 \in M_1$ *mit*

$$\|z_i - u_i\| \leq \frac{\epsilon}{2}, \quad i \in \{1, 2\}.$$

Es folgt

$$\begin{aligned}&\|tz_1 + (1-t)z_2 - \{tu_1 + (1-t)u_2\}\| \\ &\leq t\,\|z_1 - u_1\| + (1-t)\,\|z_2 - u_2\| \leq \frac{\epsilon}{2}.\end{aligned}$$

Es gilt also $tz_1 + (1-t)z_2 \in L$*, weshalb* L *konvex ist (man beachte, dass aufgrund der Konvexistät von* M_1 *mit* u_1 *und* u_2 *auch* $tu_1 + (1-t)u_2$ *in* M_1 *liegt).*
Nach dem Trennungssatz (Satz 5.12) existiert ein $\Phi \in X^*$ *und ein* $\alpha \in \mathbb{R}$ *mit* $\Phi \leq \alpha$ *auf* $\overline{L}$ *und* $\Phi(x) > \alpha$*, was im Widerspruch zur schwachen Konvergenz*

[d] Später werden wir sehen, dass $w \in u_0 + \mathring{W}^{1,p}(\Omega)$ gerade bedeutet, dass w verallgemeinerte Randwerte u_0 hat, also als „$w = u_0$ auf $\partial\Omega$“ interpretiert werden kann. Demnach wird durch $\mathscr{C}$ eine „Randwertbedingung“ realisiert.

$x_k \to x$ steht.
Unsere Annahme war demnach falsch, d.h. zu jedem $\epsilon > 0$ gibt es ein $y \in M_1$ mit $\|x - y\| \leq \epsilon$. M.a.W.: es existiert eine Folge $(v_k^1) \subset M_1$ mit

$$\left\|v_k^1 - x\right\| \to 0,\, k \to \infty.$$

Wir definieren y_1 als das erste Folgenglied von (v_k^1) mit $\left\|v_k^1 - x\right\| \leq 1$. Weiter seien $y_1, ..., y_n$ konstruiert mit

$$y_i \in M_i,\ \|y_i - x\| \leq \frac{1}{i},\ i = 1, ..., n.$$

Offenbar konvergiert auch die verschobene Folge $(x_{k+n})_{k\in\mathbb{N}}$ schwach gegen x. Wiederholung des ersten Beweisteils angewendet auf $(x_{k+n})_{k\in\mathbb{N}}$ ergibt die Existenz einer Folge (w_l) Teilmenge der konvexen Hülle von $\{x_{1+n}, x_{2+n}, ...\} = M_{k+1}$ mit

$$\|w_l - x\| \to 0,\, l \to \infty.$$

Man wähle l_0 als kleinsten Index mit $\|w_l - x\| \leq \frac{1}{i+1}$ und setzt $y_{k+1} := w_{l_0}$. □

Lemma 8.12
Sei $\Omega \subset \mathbb{R}^d$ offen und beschränkt mit Lipschitz-Rand sowie $1 < p < \infty$.

a) Für alle $u \in W_0^{1,p}(\Omega)$ gilt

$$\|u\|_p \leq c(\Omega)\, \|\nabla u\|_p\,.$$

b) Für alle $u \in W^{1,p}(\Omega)$ gilt

$$\|u - (u)_\Omega\|_p \leq c(\Omega)\, \|\nabla u\|_p\,, \quad (u)_\Omega = \frac{1}{\mathcal{L}^d(\Omega)} \int_\Omega u\, dx.$$

Beweis.
a) Nehmen wir an die Ausage sei falsch. Dann finden wir eine Folge $u_k \in W_0^{1,p}(\Omega)$ mit

$$\|u_k\|_p \geq k\, \|\nabla u_k\|_p\,.$$

Definieren wir

$$v_k := \frac{u_k}{\|u_k\|_p} \in W_0^{1,p}(\Omega),$$

so folgt

$$\|v_k\|_p = 1, \quad \|\nabla v_k\|_p \leq \frac{1}{k}. \tag{8.3}$$

Demnach ist (v_k) *eine beschränkte Folge im reflexiven Raum* $W^{1,p}(\Omega)$ *(Satz 8.9), die eine schwach konvergente Teilfolge* $(\widetilde{v}_k) \subset (v_k)$ *besitzt, d.h.*

$$\widetilde{v}_k \rightharpoonup: v \in W^{1,p}(\Omega).$$

Insbesondere ist $W_0^{1,p}(\Omega)$ *schwach abgeschlossen (Bem. 8.10), also* $v \in W_0^{1,p}(\Omega)$. *Es gilt wegen (8.3)*

$$\nabla \widetilde{v}_k \to 0 \in L^p(\Omega)$$

und damit $\nabla \widetilde{v}_k \rightharpoonup 0 \in L^p(\Omega)$. *Die Endeutigkeit des schwachen Limes in* $L^p(\Omega)$ *impliziert* $\nabla v = 0$ *und damit* ($v = 0$ *auf* $\partial\Omega$) $v = 0$. *Wir benutzen an dieser Stelle die kompakte Einbettung* $W^{1,p}(\Omega) \hookrightarrow L^p(\Omega)$ *(beschränkte Folgen in* $W^{1,p}(\Omega)$ *haben Teilfolgen, die in* $L^p(\Omega)$ *stark konvergieren), die in GdV II bewiesen wird. Nach erneuter Teilfolgenwwahl folgt*

$$\widetilde{\widetilde{v}}_k \to v \in L^p(\Omega)$$

und wir erhalten aus (8.3) $\|v\|_p = 1$ *ein Widerspruch zu* $v = 0$.
b) Vorgehen analog, man betrachte zunächst Funktionen mit

$$(u)_\Omega = \frac{1}{\mathcal{L}^d(\Omega)} \int_\Omega u \, dx = 0$$

und erhält die gewünschte Ungleichung durch Subtrahieren des Mittelwerts. □

Satz 8.13
Sei $\Omega \subset \mathbb{R}^d$ *offen und beschränkt mit Lipschitz-Rand sowie* $1 < p < \infty$. *Dann hat das Minimierungsproblem*

$$J[w] := \int_\Omega |\nabla w|^p \, dx \longrightarrow \min$$

in der Klasse $\mathscr{C}$ *eine eindeutige Lösung.*

Beweis.
Wir betrachten eine Minimalfolge $(u_n) \subset \mathscr{C}$:

$$\int_\Omega |\nabla u_n|^p dx \to \inf_{\mathscr{C}} J.$$

Offenbar gilt $\sup_n \|\nabla u_n\|_p < \infty$ *und aus der Poincaré-Ungleichung Lemma 8.12 a) folgt*

$$\|u_n\|_p \leq \|u_n - u_0\|_p + \|\nabla u_0\|_p \leq c\, \|\nabla(u_n - u_0)\|_p + \|u_0\|_p$$

$$\leq c\left\|\nabla u_n\right\|_p + c\left\|u_0\right\|_{1,p}.$$

Wir erhalten $\sup_n \|u_n\|_p < \infty$ *und insgesamt*

$$\sup_n \|u_n\|_{1,p} < \infty.$$

Mit Satz 8.9 finden wir eine Teilfolge $(\widetilde{u}_n)$ *mit*

$$\widetilde{u}_n \rightharpoonup: u \in W^{1,p}(\Omega).$$

Aufgrund der schwachen Abgeschlossenheit von $\mathscr{C}$ *(vgl. Bem. 8.10) ist* $u \in \mathscr{C}$*. Zu zeigen bleibt also, dass* u *tatsächlich* J*-minimal ist. Dazu benutzen wir Lemma 8.11 und wählen eine Folge*

$$\sum_{i=k}^{N(k)} \alpha_i^k \widetilde{u}_i \in u_0 + W_0^{1,p}, \quad \sum_{i=k}^{N(k)} \alpha_i^k = 1, \alpha_i \geq 0$$

die in $W^{1,p}(\Omega)$ *stark gegen* u *konvergiert. Es folgt mit der Konvexität der Abbildung* $t \mapsto t^p$

$$\begin{aligned}\int_\Omega |\nabla u|^p dx &= \lim_k \int_\Omega \left|\sum_{i=k}^{N(k)} \alpha_i^k \nabla \widetilde{u}_i\right|^p dx \leq \lim_k \int_\Omega \left(\sum_{i=k}^{N(k)} \alpha_i^k |\nabla \widetilde{u}_i|\right)^p dx \\ &\leq \lim_k \int_\Omega \sum_{i=k}^{N(k)} \alpha_i^k |\nabla \widetilde{u}_i|^p dx = \inf_{\mathscr{C}} J.\end{aligned}$$

Dabei folgt die letzte Gleichheit aus

$$\begin{aligned}\left|\sum_{i=k}^{N(k)} \alpha_i^k \int_\Omega |\nabla \widetilde{u}_i|^p - \inf_{\mathscr{C}} J\right| &= \left|\sum_{i=k}^{N(k)} \alpha_i^k \left(\|\nabla \widetilde{u}_i\|_p^p - \inf_{\mathscr{C}} J\right)\right| \\ &\leq \sum_{i=k}^{N(k)} \alpha_i^k \left|\|\nabla \widetilde{u}_i\|_p^p - \inf_{\mathscr{C}} J\right| \\ &\leq \sum_{i=k}^{N(k)} \alpha_i^k \epsilon = \epsilon\end{aligned}$$

für $n(k) \geq n_0$*, da* $(\widetilde{u}_i)$ *Minimalfolge ist. Es gilt also*

$$\int_\Omega |\nabla u|^p dx \leq \inf_{\mathscr{C}} J$$

und wegen $u \in \mathscr{C}$

$$\int_\Omega |\nabla u|^p dx = \inf_{\mathscr{C}} J,$$

d.h. u *is ein* J*-Minimierer. Zu zeigen bleibt die Eindeutigkeit: Nehmen wir an es gebe zwei Minimierer* $u_1, u_2 \in \mathscr{C}$. *Da* $\mathscr{C}$ *konvex ist, gehört auch* $\frac{1}{2}u_1 + \frac{1}{2}u_2$ *zur Klasse* $\mathscr{C}$ *und es folgt mir der strengen Konvexität der Abbildung* $t \mapsto t^p$ *für* $p \in (1, \infty)$

$$\int_\Omega \left|\nabla\left(\frac{1}{2}u_1 + \frac{1}{2}u_2\right)\right|^p dx < \frac{1}{2}\int_\Omega |\nabla u_1|^p dx + \frac{1}{2}\int_\Omega |\nabla u_2|^p dx = \inf_{\mathscr{C}} J,$$

ein Widerspruch. □

Bemerkung 8.14

a) *Satz 8.13 lässt sich deutlich verallgemeiner: Sei* $F : \mathbb{R}^d \to [0, \infty)$ *stetig und konvex, dann hat das Minimierungsproblem*

$$\int_\Omega F(\nabla w)\, dx \longrightarrow \min$$

in $\mathscr{C}$ *eine Lösung* (Ω *wie in Satz 8.13), falls*

$$F(Z) \geq c_0|Z|^p - c_1 \text{ für alle } Z \in \mathbb{R}^d$$

gilt (mit Konstanten $c_0 > 0$ *und* $c_1 \geq 0$*). Falls* F *streng konvex ist, ist die Lösung eindeutig.*

b) *Die Forderung der Konvexität ist sogar notwendig für die Existenz eines Minimierers, wie Acerbi und Fusco 1984 zeigen konnten (vgl. [Da]). Im Fall von vektorwertigen Mimimierern tritt an diese Stelle ein abgeschwächter Konvexitätsbegriff.*

Schlußbemerkung.

Alle Aussagen über die Räume $W^{k,p}(\Omega)$ *(oder deren Teilräume) können sinngemäß auf die Räume* $W^{k,p}(\Omega)^D$ *(bzw. deren Teilräume) übertragen werden.*

Literaturverzeichnis

[Ad] *R. A. Adams.* Sobolev Spaces. *Pure Appl. Math.* ***65****, Academic Press, New York/ London (1975).*

[AFP] *L. Ambrosio, N. Fusco, D. Pallara.* Functions of Bounded Variation and Free Discontinuity Problems. *Oxford Mathematical Monographs, Clarendon Press, Oxford (2000).*

[Al] *M. A. Al-Gwaiz.* Theory of Distributions. *Pure Appl. Math.* ***159****, Marcel Dekker Inc., New York (1992).*

[Alt] *H. W. Alt.* Lineare Funktionalanalysis — Eine anwendungsorientierte Einführung. *Springer–Lehrbuch. Zweite, verbesserte Auflage, Springer Verlag, Berlin et. al. (1992).*

[Da] *B. Dacorogna.* Direct Methods in the Calculus of Variations. *Appl. Math. Sci.* ***78****, Springer Verlag, Berlin et. al. (1989).*

[HS] *E. Hewitt, K. Stromberg.* Real and Abstract Analysis — A Modern Treatment of the Theory of Functions of a Real Variable. *Zweite, korrigierte Auflage, Springer Verlag, Berlin (1969).*

[Yo] *K. Yosida.* Functional Analysis. *Grundl. math. Wiss.* ***123****, Springer Verlag, Berlin (1965).*

Index