Kapitel 1

Der Satz von Hahn-Banach

Der Satz von Hahn-Banach befasst sich mit der Fortsetzung von linearen Funktionalen, welche auf einem linearen Teilraum definiert sind, auf den ganzen Vektorraum - dies unter Beibehaltung der Norm bzw. verwandter Größen. Im nicht-separablen Fall wird zum Beweis des Lemma von Zorn verwendet.
Der Satz von Hahn-Banach zählt zu den grundlegenden Sätzen der Funktionalanalysis. Wir geben eine analytische und später eine geometrische Formulierung dieses Satzes.

1.1 Analytische Formulierung des Satzes von Hahn-Banach

1.1 Satz (Hahn-Banach, analytische Form). *Sei V ein Vektorraum und $p : V \to \mathbb{R}$ positiv homogen und subadditiv, d.h. $p(\lambda x) = \lambda p(x)$ für alle $x \in V$ und $\lambda \in \mathbb{R},\ \lambda > 0$, und $p(x + y) \leq p(x) + p(y)$ für alle $x, y \in V$. Ferner sei W ein linearer Teilraum von V und $\varphi : W \to \mathbb{R}$ ein lineares Funktional mit*

$$\varphi(x) \leq p(x) \qquad \text{für alle } x \in W\,.$$

Dann gibt es ein lineares Funktional $f : V \to \mathbb{R}$, das ein Fortsetzung von φ ist, d.h. $f|_W = \varphi$, so dass

$$f(x) \leq p(x) \qquad \text{für alle } x \in V\,.$$

• Besonders wichtig ist der Fall, dass $p(x) = \|x\|$, wenn V ein normierter Raum mit Norm $\|.\|$ ist.
Zum Beweis von Satz 1.1 zeigen wir zunächst, dass das Funktional φ in Satz 1.1 auf einen größeren Raum $W \oplus \langle x_0 \rangle$ unter Erhaltung der Ungleichung $\varphi(x) \leq p(x)$ fortgesetzt werden kann.

1.2 Lemma (Elementarerweiterung). *Unter den Voraussetzungen von Satz 1.1 gibt es zu jedem $x_0 \in V \setminus W$ eine lineare Fortsetzung $f_1 : W \oplus \langle x_0 \rangle \to \mathbb{R}$ von $\varphi : W \to \mathbb{R}$, so dass*

$$f_1(y) \leq p(y) \qquad \text{für alle } y \in W \oplus \langle x_0 \rangle\,. \tag{1.3}$$

Beweis : Wir setzen

$$f_1(x + tx_0) = \varphi(x) + t\alpha\,,$$

wobei α noch definiert werden muss. Jedenfalls ist f_1 linear auf $W \oplus \langle x_0 \rangle$. Um ein geeignetes α zu finden, welches die Ungleichung (1.3) erfüllt, beachten wir, dass wegen der positiven Homogenität von p nur sichergestellt werden muss, dass für alle $x \in W$

$$\begin{aligned} \varphi(x) + \alpha &\leq p(x + x_0)\,, \\ \varphi(x) - \alpha &\leq p(x - x_0) \end{aligned} \tag{1.4}$$

gilt. Ersetzt man in (1.4) x durch x/t, $t > 0$, und multipliziert mit t, so ergibt sich (1.3) für $y = x \pm tx_0$. Um (1.4) zu gewährleisten, muss man offensichtlich

$$\sup_{x\in W}\{\varphi(x) - p(x - x_0)\} \leq \alpha \leq \inf_{x\in W}\{p(x + x_0) - \varphi(x)\}$$

sicherstellen. Es gilt aber

$$\varphi(x) + \varphi(\xi) = \varphi(x + \xi) \leq p(x + \xi) \leq p(x + x_0) + p(\xi - x_0)\,,$$

woraus

$$\varphi(\xi) - p(\xi - x_0) \leq p(x + x_0) - \varphi(x)\,, \qquad \forall \xi, x \in W \tag{1.5}$$

folgt. ■

Bevor wir Satz 1.1 beweisen, bemerken wir, dass der wichtige Fall von separablen normierten Räumen sowie dem Funktional $p(x) = K\|x\|$ einfach zu beweisen ist.

Beweis (Satz 1.1, separabler, normierter Raum, $p(x) = K\|x\|$): Es sei $\{x_j |\ j \in \mathbb{N}\}$ eine Menge von linear unabhängigen Vektoren aus V, deren lineare Hülle in V dicht ist. Sei $\{x_j |\ j \in \Lambda \subset \mathbb{N}\}$ die Teilmenge von Vektoren x_j, die nicht in $\overline{W}$ liegen. Durch sukzessive Anwendung der Elementarerweiterung mit $x_0 = x_j$, $j \in \Lambda$, erhalten wir eine Fortsetzung f_1 von φ auf $W \oplus \langle x_j | j \in \Lambda \rangle$ unter Beibehaltung der Ungleichung

$$f_1(x) \leq K\|x\|.$$

Durch Abschluss erhalten wir eine Fortsetzung auf den ganzen Raum V, d.h. wir definieren

$$f(y) := \lim_{y_k \to y} f_1(y_k)\,, \qquad y_k \in W \oplus \langle x_j \,\big|\, j \in \Lambda \rangle\,.$$

Dies zusammen mit obiger Ungleichung liefert für alle $y \in X$

$$f(y) \leq K\|y\|.$$

■

Der Beweis des allgemeinen Satzes von Hahn-Banach geschieht mit Hilfe des *Lemmas von Zorn*. Zur Formulierung dieses Lemmas benötigen wir einige Begriffe: Es sei M eine Menge mit einer **Halbordnung** $\prec$, d.h. für gewisse Paare $(a, b) \in M \times M$ ist

$$a \prec b \qquad \text{oder} \qquad b \prec a\,.$$

Wenn einer der beiden Fälle zutrifft, sagt man *„a und b sind vergleichbar"*. Es gelten die Regeln

(i) $a \prec a$,

(ii) $\{a \prec b \text{ und } b \prec a\} \Rightarrow a = b$,

(iii) $\{a \prec b \text{ und } b \prec c\} \Rightarrow a \prec c$.

Beispiel: (i) $M = \mathbb{R}^n$ mit der Halbordnung $\leq$, die definiert ist durch

$$a \leq b \Leftrightarrow a_i \leq b_i \qquad \text{für die Komponenten } a_i \text{ und } b_i\,,\ i = 1, \ldots, n\,.$$

(ii) Die Menge der linearen Teilräume eines Vektorraumes bildet eine Halbordnung bezüglich der mengentheoretischen Inklusion.

Eine **Kette** Q ist eine Teilmenge einer Menge M mit Halbordnung, welche **total geordnet** ist, d.h. für je zwei Elemente $a, b \in Q$ gilt $a \prec b$ oder $b \prec a$.

Beispiel: Jede eindimensionale Strecke im $\mathbb{R}^n$ mit einem Richtungsvektor der nur nichtnegative Komponenten hat ist bzgl. der Halbordnung $\leq$ eine Kette.

Eine **obere Schranke** für eine Teilmenge S einer Menge M mit Halbordnung $\prec$ ist ein Element a mit

$$s \prec a \qquad \text{für alle } s \in S\,.$$

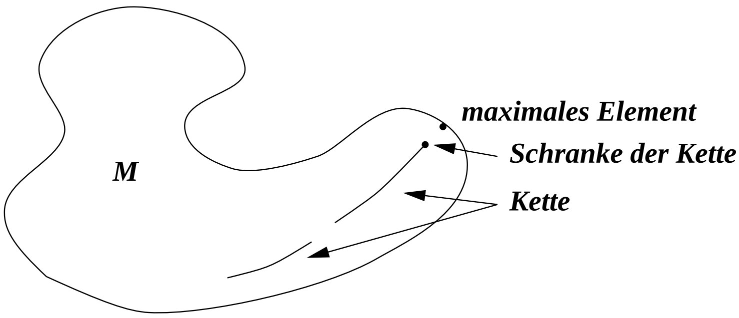
Ein Element $m \in M$ heißt **maximal** (auf M), wenn es von keinem anderen Element $x \in M$ bezüglich $\prec$ übertroffen werden kann, d.h. aus

$$m \prec x \qquad \text{folgt } x = m\,.$$

(Es ist aber durchaus möglich, dass m maximal ist und mit dem Element x nicht vergleichbar ist.)

1.6 Lemma (Zorn). *Sei M eine nichtleere Menge mit Halbordnung. Jede Kette aus M besitze eine obere Schranke. Dann besitzt M ein maximales Element.*

Beispiel: Sei $M \subset \mathbb{R}^2$ beschränkt und abgeschlossen und mit der Halbordnung $\leq$ aus Beispiel (i) versehen. Man überlegt sich - etwa mit Hilfe komponentenweiser Supremumsbildung - dass jede Kette eine obere Schranke besitzt. Es muss daher ein maximales Element in M geben. M braucht durchaus nicht „konvex“ sein.



Das Lemma von Zorn wird aus dem sogenannten *Auswahlaxiom* der Mengenlehre hergeleitet: *Ist $\mathcal{F}$ eine Familie von nichtleeren Mengen, so gibt es eine Funktion, die jeder Menge M aus $\mathcal{F}$ ein Element $m \in M$ zuordnet.*
Dies erscheint evident, aber die Aussage ist äquivalent zum sogenannten *Wohlordnungssatz*, der zumindest einem Analytiker Unwohlsein erzeugt, wie sogleich erläutert wird.
Der Wohlordnungssatz besagt, dass jede Menge M wohlgeordnet werden kann, d.h. es gibt eine Ordnungsrelation $\prec$ in M, die die Axiome der *Halbordnung* erfüllt, so dass je zwei Elemente $a, b \in M$ *vergleichbar* sind (d.h. es gilt $a \prec b$ oder $b \prec a$) und zusätzlich die Eigenschaft besitzt, dass jede nichtleere Teilmenge von M ein kleinstes Element bezüglich der Ordnung $\prec$ besitzt. Die letztere Eigenschaft bedeutet, dass die übliche Ordnung in $\mathbb{R}$, welche durch das $\leq$-Symbol gegeben ist, *keine* Wohlordnung ist, denn offene Intervalle in $\mathbb{R}$ besitzen *kein* kleinstes Element. In der Tat hat bisher noch niemand eine Wohlordnung der reellen Zahlen konkret konstruiert - man weiss nur, dass sie existiert und aus dem Auswahlaxiom folgt.

Nimmt man den Wohlordnungssatz als gegeben an, ist der Beweis des Zornschen Lemmas „einfach“. Wir geben ihn an, weil in ihm das Prinzip der *transfiniten Induktion* verwendet wird, welches man aus „erkenntnistheoretischen“ Gründen einmal gesehen haben sollte.

Beweis (Lemmas von Zorn mit Hilfe des Wohlordnungssatzes): Es sei M die im Lemma von Zorn genannte halbgeordnete Menge und $x_\alpha,\ \alpha \in I$, eine Wohlordnung der Elemente $x_\alpha \in M$, d.h. I ist eine wohlgeordnete Indexmenge bezüglich einer Ordnungsrelation, die wir mit dem Symbol $\leq$ bezeichnen und zu jedem $\alpha \in I$ gehört genau ein $x_\alpha \in M$. Wir bestimmen durch transfinite Induktion eine *Kette* G in M.

(i) Sei α_0 der kleinste Index aus I. Dann gehöre x_{α_0} zu G.

(ii) Ist für $\beta < \gamma,\ \beta, \gamma \in I$, entschieden, welche x_β zu G gehören, so gehöre x_γ genau dann zu G, wenn $x_\beta \prec x_\gamma$ für alle $x_\beta \in G,\ \beta < \gamma$.

Nach Konstruktion ist G eine Kette. Nach der Voraussetzung im Lemma von Zorn hat G eine obere Schranke z. Da z eines der x_α ist, muss z in G liegen, ist also das größte Element von G und maximal in M. (Wenn es nicht maximal wäre, gäbe es ein $x'_\alpha \succ z$, welches aber dann aufgrund der Konstruktion von G in G liegen müsste, z könnte dann nur obere Schranke sein, wenn $x'_\alpha = z$.) ∎

Der Kürze halber haben wir hier die Indexmenge I nicht konkretisiert und verweisen auf Bücher aus der Mengenlehre (Kamke).

Nach diesen Vorbereitungen beweisen wir den Satz von Hahn-Banach für reelle Vektorräume in voller Allgemeinheit.

Beweis (Satz 1.1): Wir betrachten die Menge

$$M = \big\{ h : D(h) \longrightarrow \mathbb{R} \,\big|\, D(h) \text{ linearer Teilraum von } V,\ W \subset D(h),\\ h \text{ linear},\ h|_W = \varphi,\ h(x) \leq p(x)\ \forall x \in D(h) \big\}\,.$$

In M besteht die Halbordnung $\prec$, welche erklärt ist durch

$$h_1 \prec h_2 \Leftrightarrow D(h_1) \subset D(h_2) \text{ und } h_2 \text{ ist Fortsetzung von } h_1\,.$$

Da $\varphi \in M$, ist $M \neq \emptyset$. Ferner erfüllt $(M, \prec)$ die Voraussetzung des Lemmas von Zorn. Ist nämlich K eine Kette in $M,\ K = \{h_i \in M \,\big|\, i \in I\}$, so definiert man

$$D(h) = \bigcup_{i \in I} D(h_i) \text{ und } h(x) = h_i(x)\,, \quad x \in D(h_i)\,.$$

Das so konstruierte Element $h \in M$ ist dann eine obere Schranke von K. Nach dem Lemma von Zorn gibt es daher ein maximales Element $f \in M$ bezüglich $\prec$. Wir behaupten, dass $D(f) = V$, was den Beweis dann vollenden würde. Angenommen, es wäre $D(f) \neq V$. Dann gibt es ein $x_0 \in V$, $x_0 \notin D(f)$ und wir führen eine Elementarerweiterung von f nach Lemma 1.2 durch. f wäre dann nicht maximal. ■

Wir geben sofort einige Folgerungen aus dem Satz von Hahn–Banach an. Dazu erinnern wir an die Definition $\|f\|_{V^*} = \sup\limits_{\substack{\|u\|_V \leq 1 \\ u \in V}} |\langle f, u\rangle|$

1.7 Lemma. *Sei V ein normierter Raum und $W \subset V$ ein linearer Teilraum. Jedes lineare, stetige Funktional $\varphi \in W^*$ lässt sich zu einem linearen, stetigen Funktional $f \in V^*$ fortsetzen, so dass*

$$\|f\|_{V^*} \leq \|\varphi\|_{W^*} .$$

Beweis : Man wendet Satz 1.1 mit $p(x) = \|\varphi\|_{W^*} \|x\|$ an. ■

1.8 Lemma. *Sei V ein normierter Vektorraum. Zu jedem $u \in V$ existiert ein $f \in V^*$ mit $\|f\|_{V^*} = \|u\|_V$ und $\langle f, u\rangle = \|u\|_V^2$.*

Beweis : Man wendet Lemma 2.12 mit $W = \langle u\rangle$ und $\varphi(tu) = t\|u\|^2$ an. Es gilt dann

$$\|\varphi\|_{W^*} = \sup\left\{ |t|\ \|u\|^2 \ \Big|\ |t|\|u\| \leq 1\right\} = \|u\| .$$

Nach dem Satz von Hahn-Banach gibt es eine Fortsetzung $f \in V^*$ von φ mit $\|f\| \leq \|u\|$. Da $f|_W = \varphi$ gilt

$$\langle f, u\rangle = \|u\|^2$$

und somit

$$\sup_{\|v\|\leq 1} |\langle f, v\rangle| = \|u\| ,$$

denn

$$\langle f, v\rangle \leq \|f\|\,\|v\| \leq \|u\|\|v\| .$$

und für $v = \frac{u}{\|u\|}$ wird die Gleichheit angenommen. ■

1.9 Lemma. *Sei V normierter Vektorraum und $u \in V$. Aus $\langle f, u\rangle = 0$ für alle $f \in V^*$ folgt $u = 0$.*

Beweis : Man wähle das in Lemma 1.8 konstruierte f und erhält $0 = \langle f, u\rangle = \|u\|^2$. ∎

• Das Element f in Lemma 1.8 muss nicht eindeutig sein. Dies gilt nur, wenn die Norm in V^* **strikt konvex** ist, d.h. für jedes Paar $f_1 \neq f_2 \in V^*$, $\|f_1\| = \|f_2\| = 1$ gilt

$$\|tf_1 + (1-t)\, f_2\| < 1\,, \qquad 0 < t < 1\,.$$

Die Abbildung, die jedem $u \in V$ die Menge der Elemente f zuordnet, so dass die Aussagen von Lemma 1.8 gelten, nennt man die **Dualitätsabbildung**.
Nach Definition von $\|f\|_{V^*},\ f \in V^*$, gilt

$$\|f\|_{V^*} := \sup\Big\{\, |\langle f, u\rangle| \;\Big|\; \|u\| \leq 1, u \in V \Big\}\,, \tag{1.10}$$

nach Lemma 1.8 hingegen gilt

$$\|u\| = \max\{|\langle f, u\rangle| \;\big|\; \|f\|_{V^*} \leq 1\,, \quad f \in V^*\}\,, \tag{1.11}$$

denn es gibt ein $f_0 \in V^*$ mit $\|f_0\|_{V^*} = \|u\|\,,\ \langle f_0, u\rangle = \|u\|^2$. Das Element $\|u\|^{-1} f_0$ realisiert das Maximum in (1.11).
Man kann beweisen, dass auch das „sup" in (1.10) genau dann angenommen wird, wenn der Raum reflexiv ist.

1.2 Geometrische Formulierung des Satzes von Hahn-Banach

Man kann den Satz von Hahn-Banach auch als Satz über die Trennung konvexer Mengen durch Hyperebenen formulieren. Dazu benötigen wir einige Begriffe. Wie im endlich-dimensionalen Fall definiert man Hyperebenen.

2.1 Definition. *Eine* ***Hyperebene*** *in einem Vektorraum V ist eine Menge der Gestalt*

$$H = \{x \in V |\; \varphi(x) = \alpha\} =: \{\varphi = \alpha\}$$

mit einer linearen Abbildung $\varphi : V \to \mathbb{R}$, $\varphi \neq 0$, und einer Zahl $\alpha \in \mathbb{R}$.

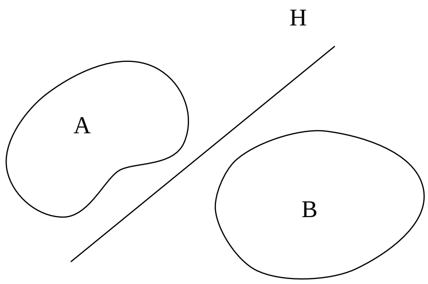
2.2 Lemma. *Sei V ein normierter Vektorraum. Eine Hyperebene $\{\varphi = \alpha\}$ ist genau dann abgeschlossen, wenn φ stetig ist.*

Beweis : Übungsaufgabe ∎

2.3 Definition. *Seien A, B Teilmengen des normierten Vektorraumes V. Man sagt, „die Hyperebene $\{\varphi = \alpha\}$* ***trenne*** *A und B“, wenn*

$$\varphi(x) \leq \alpha \leq \varphi(y) \quad \textit{für alle } x \in A \textit{ und } y \in B\,. \tag{2.4}$$

• Auf die Reihenfolge von A und B kommt es nicht an, d.h. (2.4) bedeutet auch, dass $\{\varphi = \alpha\}$ die Mengen B und A trennt.



Man spricht von **strikter Trennung**, wenn ein $\varepsilon > 0$ existiert mit

$$\varphi(x) + \varepsilon \leq \alpha \leq \varphi(y) - \varepsilon\,, \qquad x \in A,\ y \in B\,.$$

2.5 Satz (Hahn-Banach, geometrische Form). *Sei V ein normierter Raum und A, B zwei nichtleere, disjunkte, konvexe Teilmengen von V. Die Menge A sei offen. Dann gibt es eine abgeschlossene, A und B trennende Hyperebene.*

Zum Beweis benötigen wir das **Minkowski-Funktional** $p_C\colon V \to \mathbb{R} \cup \{\infty\}$ einer konvexen Menge $C \subset V,\ 0 \in C$

$$p_C(x) := \inf\{\alpha > 0 \,\big|\, \alpha^{-1}x \in C\}\,,$$

wobei $\inf \emptyset := \infty$.

2.6 Lemma. *Sei C eine offene, konvexe Teilmenge des normierten Vektorraumes V. Es sei $0 \in C$. Dann gibt es einer Konstante $M > 0$, so dass für alle $x \in V$*

$$0 \leq p_C(x) \leq M\,\|x\|\,. \tag{2.7}$$

Ferner gilt

$$C = \{x \in V \,\big|\, p_C(x) < 1\} \tag{2.8}$$

und p_C ist positiv homogen und subadditiv.

Beweis : Sei $B_r \subset C$, $r > 0$. Dann gilt $(r-\varepsilon)\frac{x}{\|x\|} \in C$, $x \in V$, und $p_C(x) \le \alpha_0 = \frac{\|x\|}{r}$ nach Definition von p_C. Daraus folgt (2.7). Die positive Homogenität von p_C folgt aus

$$\begin{aligned} p_C(\lambda x) &= \inf\{\alpha > 0 \,\big|\, \alpha^{-1}\lambda x \in C\} \\ &= \inf\{\beta\lambda \,\big|\, \beta^{-1} x \in C\} \\ &= \lambda \inf\{\beta \,\big|\, \beta^{-1} x \in C\} = \lambda\, p_C(x)\,, \qquad \lambda > 0\,. \end{aligned}$$

Wir beweisen nun (2.8). Ist $x \in C$, so gilt auch $(1+\varepsilon)\,x \in C$ für genügend kleines ε wegen der Offenheit von C. Daher $p_C(x) \le \frac{1}{1+\varepsilon} < 1$, d.h. $C \subset \{x \in V \,|\, p_C(x) < 1\}$. Ist umgekehrt $p_C(x) < 1$, dann folgt $\alpha^{-1}x \in C$ mit einem $\alpha > 0$ und somit $x = \alpha(\alpha^{-1}x) + (1-\alpha)\,0 \in C$. Es verbleibt der Nachweis der Subadditivität von p_C. Seien $x, y \in V$, $\varepsilon > 0$. Da

$$p_C\left(\frac{x}{p_C(x)+\varepsilon}\right) = \frac{1}{p_C(x)+\varepsilon}\, p_C(x) < 1\,,$$

folgt

$$\frac{x}{p_C(x)+\varepsilon} \in C \quad \text{und entsprechend } \frac{y}{p_C(y)+\varepsilon} \in C\,.$$

Aus Konvexitätsgründen ist

$$\frac{tx}{p_C(x)+\varepsilon} + \frac{(1-t)\,y}{p_C(y)+\varepsilon} \in C, \qquad 0 < t < 1\,.$$

Setzt man

$$t = \frac{p_C(x)+\varepsilon}{p_C(x)+p_C(y)+2\varepsilon}\,,$$

folgt

$$\frac{x+y}{p_C(x)+p_C(y)+2\varepsilon} \in C\,.$$

Dies zusammen mit (2.8) und der positiven Homogenität liefert

$$p_C(x+y) < p_C(x) + p_C(y) + 2\varepsilon\,.$$

Grenzübergang $\varepsilon \to 0$ ergibt die Behauptung. ■

2.9 Lemma (Trennung einer konvexen Menge und eines Punktes). *Sei V ein normierter Vektorraum und $C \subset V$ offen, konvex und nichtleer. Sei $x_0 \in V$, $x_0 \notin C$. Dann gibt es ein beschränktes, lineares Funktional $f \in V^*$, $f \ne 0$ mit*

$$f(x) < f(x_0) \qquad \textit{für alle } x \in C\,.$$

Beweis : O.B.d.A. sei $0 \in C$. Auf dem eindimensionalen Raum $\langle x_0 \rangle$ definieren wir das lineare Funktional φ durch

$$\varphi(tx_0) = t\,.$$

Es gilt $\varphi(x) \leq p_C(x)$, für alle $x \in \langle x_0 \rangle$ da andernfalls für ein $y \in \langle x_0 \rangle$ gelten würde:

$$p_C(y) < \varphi(y)\,, \qquad \text{d.h. } p_C(tx_0) < t \text{ und somit } p_C(x_0) < 1\,.$$

Dies würde nach Lemma 2.6 aber $x_0 \in C$ bedeuten. Nach dem Satz von Hahn-Banach in der analytischen Form (Satz 1.1) lässt sich φ zu einem linearen Funktional $f : V \to \mathbb{R}$ fortsetzen mit

$$f(x) \leq p_C(x)\,.$$

Insbesondere ist $f(x_0) = \varphi(x_0) = 1$ und andererseits $f(x) \leq p_C(x) < 1$ für $x \in C$ wegen (2.8). ∎

Beweis (Satz 2.5): Die Menge $C = A \ominus B \equiv \left\{a - b \,\middle|\, a \in A,\ b \in B\right\}$ ist konvex. Wegen der Darstellung $C = \bigcup\limits_{y \in B} (A - y)$ ist C offen. Da $A \cap B = \emptyset$, gilt $0 \notin C$. Wegen Lemma 2.9 gibt es $f \in V^*$ (d.h. ist f stetig und linear) mit

$$\begin{aligned} f(z) < f(0) = 0\,, &\qquad z \in C\,, \\ f(x) < f(y)\,, &\qquad x \in A,\ y \in B\,. \end{aligned} \tag{2.10}$$

Wir wählen α so, dass

$$\sup_{x \in A} f(x) \leq \alpha \leq \inf_{y \in B} f(y)\,.$$

Die Hyperebene $\{f = \alpha\}$ ist wegen Lemma 2.2 abgeschlossen und trennt A und B. ∎

Im Unendlichdimensionalen lassen sich beliebige konvexe Mengen nicht notwendig trennen.
Beispiel: $V = L^2(\Omega)$, Ω ein Gebiet des $\mathbb{R}^n$.

$$A = \left\{z \in C(\Omega) \cap L^2 \,\middle|\, \|z\|_2 < 1\right\}, \qquad B = \{y_0\}$$

mit fester L^2-Funktion $y_0 \notin C(\Omega)$, $\|y_0\|_{L^2} = \frac{1}{2}$. Man beachte, dass das Innere von A bezüglich der L^2-Norm *leer* ist. Ferner gilt $A \cap B = \emptyset$. Wir behaupten, dass A und B *nicht* durch eine abgeschlossene Hyperebene getrennt werden können. Gäbe es ein $f \in (L^2(\Omega))^*$ für das

$$f(z) < f(y_0)\,, \qquad z \in C(\Omega) \cap L^2\,, \quad \|z\|_{L^2} < 1\,,$$

gilt, so folgt, da $C(\Omega) \cap B_1$ dicht in $B_1 = \big\{w \in L^2 \big| \, \|w\|_{L^2} \le 1\big\}$ ist,

$$\begin{aligned} \|f\|_{(L^2(\Omega))^*} &= \sup_{\|z\|_{L^2(\Omega)} \le 1} |f(z)| = \sup_{z \in C(\Omega) \cap B_1} |f(z)| \\ &< |f(y_0)| \le \tfrac{1}{2} \|f\|_{(L^2(\Omega))^*}\,. \end{aligned}$$

Dies wäre ein Widerspruch.

Wir notieren eine Variante der geometrischen Form des Satzes von Hahn-Banach.

2.11 Satz. *Sei V ein normierter Vektorraum und $A, C \subset V$ konvexe, disjunkte nichtleere Teilmengen von V. Ferner sei A abgeschlossen und C kompakt. Dann gibt es eine abgeschlossene Hyperebene, die A und C strikt trennt.*

Beweis : Für $\varepsilon > 0$ setzen wir $A_\varepsilon = A + B_\varepsilon(0), C_\varepsilon = C + B_\varepsilon(0)$. Offensichtlich sind $A_\varepsilon, C_\varepsilon$ offen, konvex und nichtleer. Falls $\varepsilon > 0$ klein genug gewählt wurde gilt auch $A_\varepsilon \cap C_\varepsilon = \emptyset$. Falls dies nicht gelten würde, gäbe es $\varepsilon_n \searrow 0, x_n \in A, y_n \in C$, $\|x_n - y_n\| \le 2\varepsilon_n$. Da C kompakt ist gibt es eine Teilfolge $\Lambda \subseteq \mathbb{N}, y_n \to y \in C$, $n \to \infty, n \in \Lambda$. Dies ist aber ein Widerspruch. Satz 2.5 liefert die Existenz einer A_ε und C_ε trennenden, abgeschlossenen Hyperebene, d.h. es gibt ein $f \in V^*$ und $\alpha \in \mathbb{R}$ mit

$$f(x + \varepsilon z_1) \le \alpha \le f(y + \varepsilon z_2), \qquad \forall x \in A, y \in C, z_1, z_2 \in B_1(0)\,.$$

Somit gilt, da sowohl z als auch $-z$ gewählt werden können

$$f(x) + \varepsilon \|f\| \le \alpha \le f(y) - \varepsilon \|f\|\,, \qquad \forall x \in A, y \in C\,,$$

d.h. A und C sind strikt getrennt. ■

2.12 Folgerung. *Sei V ein normierter Vektorraum und W ein linearer Teilraum mit $\overline{W} \ne V$. Dann existiert ein $f \in V^*$, $f \ne 0$, so dass für alle $x \in W$ gilt:*

$$\langle f, x\rangle = 0$$

Beweis : Sei $x_0 \in V \setminus \overline{W}$. Dann folgt mit Satz 2.11 für $A = \overline{W}$ und $C = \{x_0\}$, dass es ein $f \in V^*$, $f \ne 0$, und $\alpha \in \mathbb{R}$ gibt, mit

$$\langle f, x\rangle < \alpha < \langle f, x_0\rangle \qquad \forall\, x \in \overline{W}\,.$$

Für $x \in W$ ist $\lambda x \in W$ für alle $\lambda \in \mathbb{R}$, und somit

$$\left.\begin{array}{r}\lambda \to \infty \Rightarrow \langle f, x\rangle \le 0 \\ \lambda \to -\infty \Rightarrow \langle f, x\rangle \ge 0\end{array}\right\} \Rightarrow \langle f, x\rangle = 0 \;\forall\; x \in \overline{W}\,.$$

■

• Die analytische Form des Satzes von Hahn-Banach lässt sich mit Hilfe der geometrischen Form beweisen - siehe Köthe, Topologische lineare Räume, Kap. 17.2. Der Satz ist dort noch erheblich allgemeiner dargestellt.

1.3 Konjugiert-konvexe Funktionen

Im Folgenden sei V ein normierter Raum. Wie üblich, verabreden wir: Die Funktion

$$g : V \to \mathbb{R} \cup \{\infty\}$$

heißt **konvex**, wenn

$$g\big(tx + (1-t)\,y\big) \le tg(x) + (1-t)\,g(y) \qquad \text{für alle } x, y \in V, 0 \le t \le 1.$$

Hierbei ist $\alpha + \infty = \infty$ gesetzt. Die Menge aller $x \in V$ mit $g(x) < \infty$ bezeichnen wir mit $D(g)$. Es ist klar, dass $D(g)$ konvex ist.

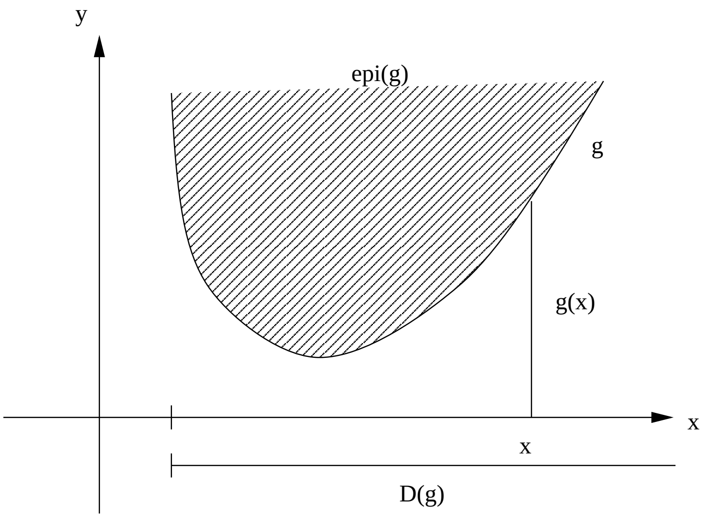
3.1 Definition. *Die Funktion $g : V \to \mathbb{R} \cup \{\pm\infty\}$ heißt **unterhalbstetig**, wenn für alle konvergenten Folgen (x_n) aus V mit $x_n \to x$ die Ungleichung*

$$g(x) \le \liminf_{n\to\infty} g(x_n)$$

gilt.

3.2 Definition. *Der **Epigraph** einer Funktion $g : V \to \mathbb{R} \cup \{\infty\}$ ist die Menge*

$$\operatorname{epi}(g) = \big\{(x, \lambda) \in V \times \mathbb{R} \,\big|\, g(x) \le \lambda\big\}\,.$$



3.3 Lemma. *g ist genau dann unterhalbstetig, wenn der Epigraph von g abgeschlossen ist.*

Beweis : Übungsaufgabe. ∎

• Wenn g konvex ist, ist auch epi(g) konvex; die Umkehrung ist ebenfalls richtig. (Übungsaufgabe)

3.4 Definition. *Sei $g : V \to \mathbb{R} \cup \{\infty\}$ und $g \not\equiv \infty$, d.h. $D(g) \neq \emptyset$. Die* ***duale Funktion*** *$g^* : V^* \to \mathbb{R} \cup \{\infty\}$ ist definiert durch*

$$g^*(f) = \sup_{u \in V} \big\{ \langle f, u \rangle - g(u) \big\} , \qquad f \in V^* .$$

• Es ist klar, dass g^* *konvex* ist, selbst wenn g nicht konvex sein sollte.
• Man überlegt sich leicht, dass g^* unterhalbstetig ist. (Übungsaufgabe)

Beispiel: $V = \mathbb{R}$, $\quad g(x) = \frac{|x|^p}{p}$

$$g^*(\xi) = \sup_{x \in \mathbb{R}} \left\{ \xi x - \frac{|x|^p}{p} \right\} .$$

Das Supremum $x^* = g^*(\xi)$ wird angenommen; es gilt dann $\xi - |x^*|^{p-1} \operatorname{sign} x^* = 0$ und somit

$$g^*(\xi) = |x^*|^{p-1} (\operatorname{sign} x^*)\, x^* - \frac{|x^*|^p}{p} = \left(1 - \frac{1}{p} \right) |x^*|^p .$$

Wegen $|x^*| = |\xi|^{\frac{1}{p-1}}$ folgt

$$g^*(\xi) = \frac{|\xi|^q}{q} \qquad \text{mit } q = \frac{p}{p-1} .$$

3.5 Lemma. *Sei V ein normierter Raum und $g : V \to \mathbb{R} \cup \{\infty\}$ konvex, unterhalbstetig und $g \not\equiv \infty$. Dann ist $g^* \not\equiv \infty$.*

Beweis : Sei $x_0 \in D(g)$ und $\lambda_0 < g(x_0)$. Wir wenden den Satz von Hahn-Banach in der Formulierung von Satz 2.11 an. Als Grundraum wählen wir $V \times \mathbb{R}$, als abgeschlossene, konvexe Menge $A = \operatorname{epi}(g)$ und $C = (x_0, \lambda_0)$. Nach der Definition von $\operatorname{epi}(g)$ ist $A \cap C = \emptyset$. A ist konvex und abgeschlossen und C kompakt, da letztere nur aus einem Punkt besteht. Es gibt daher eine trennende abgeschlossene Hyperebene H, die A und C strikt trennt. Die Hyperebene hat die Gestalt

$$H = \big\{(x,\lambda) \in V \times \mathbb{R} \,\big|\, \langle f, x\rangle + k\lambda = \alpha\big\}$$

mit einem $f \in V^*$, $k \in \mathbb{R}$, $\alpha \in \mathbb{R}$. Es gilt hierbei

$$\begin{aligned} \langle f, x\rangle + k\lambda &> \alpha \qquad\qquad \forall (x,\lambda) \in \operatorname{epi}(g)\,, \\ \langle f, x_0\rangle + k\lambda_0 &< \alpha\,. \end{aligned} \tag{3.6}$$

Da $\big(x_0, g(x_0)\big) \in \operatorname{epi}(g)$ folgt

$$\langle f, x_0\rangle + kg(x_0) > \alpha > \langle f, x_0\rangle + k\lambda_0\,.$$

Daraus folgt $k > 0$. Aus (3.6) folgt mit $\lambda = g(x)$

$$\big\langle -\tfrac{1}{k} f, x\big\rangle - g(x) < -\tfrac{\alpha}{k}\,, \qquad x \in D(g)\,,$$

und damit

$$g^*\big(-\tfrac{1}{k} f\big) \le -\tfrac{\alpha}{k} < \infty\,.$$

■

Die **biduale konvexe Funktion** zu g hat die Gestalt

$$g^{**}(x) = \sup_{f \in V^*} \big\{\langle f, x\rangle - g^*(f)\big\}\,. \tag{3.7}$$

Hierbei ist $\alpha - \infty = -\infty$ gesetzt.

• Offensichtlich ist g^{**} immer konvex.

3.8 Satz. *Sei V ein normierter Raum und $g : V \to \mathbb{R} \cup \{\infty\}$ konvex, unterhalbstetig und $g \not\equiv \infty$. Dann gilt $g^{**} = g$.*

Beweis : (i) Sei zunächst $g \ge 0$ und $g^*(f) < \infty$ (nur solche f sind für g^{**} relevant). Da $g(x) + g^*(f) \ge \langle f, x\rangle$ folgt $g(x) \ge \langle f, x\rangle - g^*(f)$ und durch Übergang zum Supremum[1]

$$g(x) \ge g^{**}(x)\,.$$

[1]Dies gilt auch ohne die Voraussetzung $g \ge 0$.

Angenommen, es würde $g^{**} \neq g$ gelten. Dann gäbe es ein x_0 mit

$$g^{**}(x_0) < g(x_0)\,. \tag{3.9}$$

($g(x_0) = \infty$ ist zugelassen, aber dann beinhaltet die Widerspruchsannahme $g^{**}(x_0) < \infty$.) Wir wenden den Satz von Hahn-Banach an und trennen die Menge $\operatorname{epi}(g)$ und den Punkt $(x_0, g^{**}(x_0))$, denn wegen (3.9) liegt $(x_0, g^{**}(x_0))$ nicht in $\operatorname{epi}(g)$. Ähnlich wie beim Beweis des letzten Lemmas erhält man dann ein $f \in V^*$ und $k \in \mathbb{R}$ sowie $\alpha \in \mathbb{R}$ mit

$$\langle f, x\rangle + k\lambda > \alpha \qquad \forall (x,\lambda) \in \operatorname{epi}(g)\,, \tag{3.10}$$

$$\langle f, x_0\rangle + k g^{**}(x_0) < \alpha\,. \tag{3.11}$$

Lässt man in (3.10) λ gegen ∞ gehen, folgt $k \geq 0$. Aus (3.10) schließt man mit $\lambda = g(x) \geq 0$, für $\varepsilon > 0$

$$\langle f, x\rangle + (k+\varepsilon)\, g(x) > \alpha\,, \qquad x \in D(g)\,,$$

und damit

$$\Big\langle -f\tfrac{1}{k+\varepsilon},\ x\Big\rangle - g(x) < -\tfrac{\alpha}{k+\varepsilon}\,,$$

und durch Übergang zum Supremum in x

$$g^*\big(-f\tfrac{1}{k+\varepsilon}\big) \leq -\tfrac{\alpha}{k+\varepsilon}\,.$$

Weiterhin folgt nach Definition von $g^{**}(x_0)$

$$\begin{aligned} g^{**}(x_0) &\geq \Big\langle -f\tfrac{1}{k+\varepsilon},\ x_0\Big\rangle - g^*\big(-f\tfrac{1}{k+\varepsilon}\big) \\ &\geq \Big\langle -f\tfrac{1}{k+\varepsilon},\ x_0\Big\rangle + \tfrac{\alpha}{k+\varepsilon}\,, \end{aligned}$$

und somit

$$\langle f, x_0\rangle + (k+\varepsilon)\, g^{**}(x_0) \geq \alpha\,.$$

Der Grenzübergang $\varepsilon \to 0$ führt zu einem Widerspruch mit (3.11). Dies war der Fall $g \geq 0$.
(*ii*) Den allgemeinen Fall erhält man durch Abändern von g mit einem $f_0 \in D(g^*)$

$$\overline{g}(x) := g(x) - \langle f_0, x\rangle + g^*(f_0) \geq 0\,.$$

Für $\overline{g}$ weiß man, dass $\overline{g}^{**} = \overline{g}$. Es gilt

$$\begin{aligned}
\overline{g}^*(f) &= \sup_{x\in V} \big\{ \langle f, x\rangle - \overline{g}(x) \big\} \\
&= \sup_{x\in V} \big\{ \langle f + f_0, x\rangle - g(x) - g^*(f_0) \big\} \\
&= g^*(f + f_0) - g^*(f_0) \\
\overline{g}^{**}(x) &= \sup_{f\in V^*} \{ \langle f + f_0, x\rangle - g^*(f + f_0) \} + g^*(f_0) - \langle f_0, x\rangle \\
&= \sup_{\tilde{f}\in V^*} \{ \langle \tilde{f}, x\rangle - g^*(\tilde{f}) \} - \langle f_0, x\rangle + g^*(f_0) \\
&= g^{**}(x) - \langle f_0, x\rangle + g^*(f_0)
\end{aligned}$$

Daraus folgt $g(x) = g^{**}(x)$. ■

Orlicz-Räume

3.12 Definition. *Es sei $g : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ eine nichtnegative, konvexe Funktion und $\Omega \subset \mathbb{R}^n$ messbar. Mit $\tilde{L}_g(\Omega)$ (**Orliczklasse**) bezeichnen wir die Menge aller messbaren Funktionen $f : \Omega \to \mathbb{R}$ mit*

$$G(f) := \int_\Omega g(|f(x)|)\, dx < \infty \,.$$

• Die Lebesgue-Räume $L^p(\Omega), 1 < p < \infty$, sind Spezialfälle der Orliczklassen $\tilde{L}_g$. Man setzt $g(t) = c\, t^p$, mit einer beliebigen positiven Konstanten c. Weitere Beispiele von Orliczklassen sind durch die Funktionen

$$\begin{aligned}
g_1(t) &= t \ln^+ t \,, \\
g_2(t) &= (1 + t) \ln(1 + t) - t \,, \\
g_3(t) &= e^t - t - 1 \,, \\
g_4(t) &= e^{t^2} - 1 \\
g_5(t) &= e^t \,.
\end{aligned}$$

gegeben. Die ersten beiden Funktionen g_1, g_2 werden oft benutzt um den Raum L^1 zu ersetzen und die Funktionen g_3, g_4, g_5 um den Raum L^∞ zu ersetzen. Beide Räume L^1, L^∞ sind Ausnahmefälle innerhalb der Lebesgue-Räume L^p.

• Die Orliczklasse $\tilde{L}_g(\Omega)$ ist im Allgemeinen kein Vektorraum, denn:
$\Omega = (0, 1) \subseteq \mathbb{R} \quad g(t) = e^t$. Wir betrachten $f(x) = -\frac{1}{2} \ln x$. $f(x)$ ist in $\tilde{L}_g(\Omega)$,

denn
$$\int_0^1 e^{-\frac{1}{2}\ln x} dx = \int_0^1 x^{-\frac{1}{2}} dx < \infty,$$
aber $2f(x) = -\ln x \notin \tilde{L}_g$, denn
$$\int_0^1 e^{-\ln x} dx = \int_0^1 x^{-1} dx = \infty.$$

3.13 Definition. *Die Funktion $\gamma : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ habe die Eigenschaften*

(i) $\gamma(0) = 0$,

(ii) $\gamma(s) > 0$, *für* $s > 0$,

(iii) γ *ist rechtsseitig stetig für alle* $s \geq 0$,

(iv) γ *ist nichtfallend auf* $(0, \infty)$,

(v) $\lim_{s\to\infty} \gamma(s) = \infty$.

Wir sagen, dass g eine ***Young-Funktion*** *ist, falls*
$$g(t) = \int_0^t \gamma(s)\, ds, \quad t \geq 0\,.$$

3.14 Lemma. *Eine Young-Funktion g ist stetig, nichtnegativ, strikt wachsend und konvex auf* $[0,\infty)$. *Darüber hinaus gilt:*

$$\begin{aligned} g(0) &= 0\,, & \lim_{t\to\infty} g(t) &= \infty\,,\\ \lim_{t\to 0^+} \tfrac{g(t)}{t} &= 0\,, & \lim_{t\to\infty} \tfrac{g(t)}{t} &= \infty\,,\\ g(\alpha t) &\le \alpha g(t)\,, & \alpha &\in [0,1], t\ge 0\,,\\ g(\beta t) &\ge \beta g(t)\,, & \beta &> 1, t \ge 0\,. \end{aligned}$$

Beweis : Offensichtlich ist g stetig, nichtnegativ und strikt wachsend. Weiterhin sieht man sofort, dass $g(0)=0$ und $g(t)\nearrow\infty$ für $t\to\infty$. Eine stetige Funktion f ist genau dann konvex, wenn

$$f\Big(\frac{1}{2}(s+t)\Big) \le \frac{1}{2}f(s) + \frac{1}{2}f(t)$$

gilt (Übungsaufgabe). Für $0\le s<t$ gilt

$$\begin{aligned} g\Big(\frac{s+t}{2}\Big) &= \int_0^{\frac{s+t}{2}} \gamma(\tau)d\tau\\ &= \int_0^s \gamma(\tau)d\tau + \frac{1}{2}\int_s^{\frac{s+t}{2}}\gamma(\tau)d\tau + \frac{1}{2}\int_s^{\frac{s+t}{2}}\gamma(\tau)d\tau\\ &\le \int_0^s \gamma(\tau)d\tau + \frac{1}{2}\int_s^{\frac{s+t}{2}}\gamma(\tau)d\tau + \frac{1}{2}\int_{\frac{s+t}{2}}^{t}\gamma(\tau)d\tau\\ &= \frac{1}{2}\int_0^s \gamma(\tau)d\tau + \frac{1}{2}\int_0^s \gamma(\tau)d\tau + \frac{1}{2}\int_s^{\frac{s+t}{2}}\gamma(\tau)d\tau + \frac{1}{2}\int_{\frac{s+t}{2}}^{t}\gamma(\tau)d\tau\\ &= \frac{1}{2}g(s) + \frac{1}{2}g(t) \end{aligned}$$

und somit ist g konvex. Daraus folgt weiterhin für $\alpha<1$

$$g(\alpha t) = g(\alpha t + (1-\alpha)0) \le \alpha g(t) + (1-\alpha)g(0) = \alpha g(t)\,.$$

Die Substitution $\beta = \frac{1}{\alpha}$ liefert sofort $g(\beta s)\ge \beta g(s)$ für $\beta>1$. Mitholfe des Hauptsatzes der Integralrechnung erhalten wir

$$\lim_{t\to 0^+}\frac{g(t)}{t} = \lim_{t\to 0^+}\frac{1}{t}\int_0^t \gamma(s)ds = \gamma(0) = 0\,.$$

Da γ nicht fallend ist gilt:

$$\lim_{t\to\infty}\frac{g(t)}{t}=\lim_{t\to\infty}\frac{1}{t}\int\limits_0^t\gamma(s)ds\geq\lim_{t\to\infty}\frac{1}{t}\int\limits_{\frac{t}{2}}^t\gamma(s)ds\geq\lim_{t\to\infty}\frac{1}{t}\gamma\Big(\frac{t}{2}\Big)\frac{t}{2}\to\infty$$

■

• In den obigen Beispielen ist g_1 keine Young-Funktion.
• Mit Hilfe von Lemma 3.14 kann man zeigen, dass $\tilde{L}_g(\Omega)$ eine konvexe Menge ist und

$$\tilde{L}_g(\Omega)\subseteq L^1(\Omega)\,,$$

falls g eine Young-Funktion ist und Ω ein beschränktes Gebiet.
• Sei $g\colon\mathbb{R}_0^+\to\mathbb{R}_0^+$ eine Young-Funktion. Wir setzen diese durch gerade Spiegelung zu einer Funktion $g\colon\mathbb{R}\to\mathbb{R}_0^+$ fort. Sei g^* die zur fortgesetzten Funktion g duale, konvexe Funktion (vgl. Definition 3.4 mit $V=\mathbb{R}$). Man kann zeigen, dass g^* auch eine Young-Funktion ist. Deshalb nennt man g^* auch **komplementäre** Young-Funktion. In der Tat, habe g die Darstellung

$$g(t)=\int_0^t\gamma(s)\,ds\,,$$

dann hat g^* die Darstellung

$$g^*(t)=\int_0^t\gamma^*(s)\,ds\,,$$

wobei

$$\gamma^*(s)=\sup_{\gamma(t)\leq s}t\,.$$

γ^* ist die Rechtsinverse der Funktion γ.

3.15 Definition. *Sei g eine Young-Funktion. Der **Orlicz-Raum** $L_g(\Omega)$ ist die Menge aller Äquivalenzklassen $[f]$ von Funktionen f, für die*

$$\|f\|_g = \sup\left\{ \int\limits_\Omega |\varphi(x)f(x)|\, dx \;\Big|\; \int_\Omega g^*(|\varphi(x)|)\, dx \le 1 \right\} < \infty$$

*gilt. Hierbei ist $[f]$ die Klasse aller Funktionen, die sich von f nur auf einer Menge vom Mass Null unterscheiden, und g^* die zu g komplementäre Young-Funktion. Die Größe $\|f\|_g$ nennt man die **Orlicz-Norm** von f.*

• Offensichtlich ist $L_g(\Omega)$ ein linearer Raum und $\|\cdot\|_g$ eine Norm. Somit ist $L_g(\Omega)$ ein normierter Vektorraum.

Auf Grund der Definition der dualen Funktion gilt folgende Erweiterung der Young-Ungleichung:

3.16 Lemma. *Für alle $u, v \in [0, \infty)$ gilt:*

$$uv \le g(u) + g^*(v)\,,$$

wobei g eine Young-Funktion ist, und g^* die zu g komplementäre Young-Funktion.

Mit Hilfe dieses Lemmas kann man die Hölder-Ungleichung auf Orlicz-Räume erweitern.

3.17 Lemma. *Es sei g eine Young-Funktion und g^* die zu g duale Funktion. Sei $u \in L_g(\Omega)$ und $v \in L_{g^*}(\Omega)$. Dann ist $uv \in L^1(\Omega)$ und es gilt*

$$\int\limits_\Omega |uv|\, dx \le \|u\|_g\, \|v\|_{g^*}\;. \tag{3.18}$$