Vorlesungsskript

Variationsrechnung und Sobolevräume

Patrizio Neff,
nach einer Vorlage von H.D. Alber (TU Darmstadt)

Vorbemerkung

Das Dirichletsche Prinzip erlaubt die Lösung von Randwertproblemen zu vielen linearen und nichtlinearen elliptischen partiellen Differentialgleichungen auf die Lösung von Variationsproblemen zurückzuführen. Ein wichtiges Verfahren zur Lösung dieser Randwertprobleme besteht daher in der Anwendung der "Direkten Methoden der Variationsrechnung", und es ist ein wesentliches Ziel der vorliegenden Einführung in die Theorie der Sobolevräume und in die Variationsrechnung, die Grundlagen für das Studium elliptischer Differentialgleichungen zu legen. Wegen dieser Zielsetzung werden die klassischen Methoden der Variationsrechnung nur kurz gestreift.

Inhaltsverzeichnis

tion. 5.7 Eigenschaften der Hamiltonfunktion. 5.8 Hamiltonsche Differentialgleichungen, kanonische Form der Eulergleichung. 5.9 Beispiele. 5.9.1 Das Hamiltonsche Prinzip. 5.9.2 Das Fermatsche Prinzip. 5.9.3 Der Fall $f(x,u,\xi) = f(\xi)$. 5.9.4 Der Fall $f(x,u,\xi) = f(x,\xi)$. 5.9.5 Der Fall $f(x,u,\xi) = f(u,\xi)$. 5.10 Hamilton-Jacobi Gleichung. 5.11 Eigenschaften der Lösung der Hamilton-Jacobi Gleichung. 5.12 Existenz von Lösungen der Hamilton-Jacobi Gleichung. 5.13 Beispiel (Geometrische Optik, Eikonalgleichung).

6. Schwache Topologie.
6.1 Konvexe Mengen. 6.4 Projektion auf konvexe Mengen. 6.7 Reell lineare Abbildungen. 6.8 Trennungssatz. 6.10 Schwache Topologie. 6.13 Abschluß konvexer Mengen in der schwachen und in der Normtopologie. 6.14 Beschränkte Mengen. 6.15 Äquivalenz von schwacher und Normbeschränktheit. 6.17 Schwache Folgenkompaktheit der Einheitskugel.

7. Konvexe Funktionale.
7.1 Unterhalbstetige Funktionen. 7.3 Äquivalenz von Unterhalbstetigkeit und schwacher Unterhalbstetigkeit bei konvexen Funktionalen. 7.4 Konvexität der Norm. 7.5 Koerzitivität. 7.7 Existenz eines Minimums bei konvexen Funktionalen. 7.8 Beschränktheit konvexer, unterhalbstetiger Funktionale. 7.10 Subdifferential. 7.11 Beispiele für Subdifferentiale. 7.12 Wertebereich des Subdifferentials. 7.14 Surjektivität und Injektivität des Subdifferentials.

8. Direkte Methoden.
8.1 Strikte Koerzitivität. 8.2 Existenz eines Minimums für Variationsfunktionale. 8.3 Konvexität und Unterhalbstetigkeit von Variationsfunktionalen auf $H_1(\Omega)$. 8.5 Konvexität und Unterhalbstetigkeit von Variationsfunktionalen auf $L^2(\Omega)$. 8.6 Eine Version der Poincaréschen Ungleichung. 8.7 Subdifferential und Variationsungleichung. 8.8 Variationsungleichung zum Variationsfunktional. Schwache Form der Eulergleichung. 8.9 Das Subdifferential. 8.10 Schwache Divergenz. 8.11 Schwache Definition von Differentialoperatoren mit Randbedingung. 8.12 Dirichletsches Prinzip. Äquivalenz zwischen Randwertproblem, Variationsungleichung und Variationsproblem.

Lehrbücher

1. Einführung, Beispiele für Variationsprobleme, Inhalt der Vorlesung

1.1 Das allgemeine Problem der Variationsrechnung. Die Variationsrechnung ist die mathematische Theorie, die sich mit der Lösung des folgenden Grundproblems befaßt:
Seien n und N natürliche Zahlen, sei $\Omega \subseteq \mathbb{R}^n$ eine offene Menge und sei

$$f : \Omega \times \mathbb{R}^N \times \mathbb{R}^{nN} \to \mathbb{R}$$

eine stetige Funktion. Einer Funktion

$$u = (u_1, \ldots, u_N) : \Omega \to \mathbb{R}^N$$

werde die reelle Zahl

$$I(u) = \int_\Omega f(x, u(x), \nabla u(x))\, dx$$

zugeordnet. Hierbei sei

$$\nabla u(x) = \begin{pmatrix} \frac{\partial u_1}{\partial x_1}(x) & \cdots & \frac{\partial u_1}{\partial x_n}(x) \\ \vdots & & \\ \frac{\partial u_N}{\partial x_1}(x) & \cdots & \frac{\partial u_N}{\partial x_n}(x) \end{pmatrix}$$

die $N \times n$–Matrix der ersten Ableitungen von u, die ich auch als Gradient bezeichnen werde. I ist also eine Funktion vom Raum der Funktionen $u : \Omega \to \mathbb{R}^N$ in die reellen Zahlen. Man bezeichnet I auch oft als Funktional. Gesucht ist nun eine Funktion u, für die das Funktional I seinen minimalen Wert annimmt, wobei man oft noch verlangt, daß u gewisse Nebenbedingungen erfüllt. Beispiele für solche Nebenbedingungen sind zum Beispiel, daß u zum Raum $C_1(\Omega)$ gehören soll, oder daß u am Rand $\partial\Omega$ von Ω mit einer vorgegebenen Funktion u_0 übereinstimmen soll.

Man kann das Grundproblem der Variationsrechnung daher auch folgendermaßen formulieren: Sei D eine Menge von "zulässigen" Funktionen $v : \Omega \to \mathbb{R}^N$. Gesucht ist eine Funktion $u \in D$ mit

$$I(u) = \min_{v \in D} I(v)\,.$$

u heißt Lösung des Variationsproblems.

Ich möchte hier gleich auf die entscheidende Schwierigkeit aufmerksam machen. Die Menge

$$\{I(v) : v \in D\}$$

ist eine Teilmenge der reellen Zahlen. Das Variationsproblem ist lösbar, wenn diese Menge ein Minimum besitzt. Dazu muß sie nach unten beschränkt sein, also müssen das Funktional I und damit insbesondere die Funktion f und auch die Menge D so vorgegeben sein, daß das Infinuum $\inf\limits_{v \in D} I(v)$ existiert. Natürlich folgt hieraus noch nicht, daß die Bildmenge $\{I(v) : v \in D\}$ auch ein Minimum besitzt, und die gesamte Variationsrechnung dreht sich im wesentlichen darum, zu zeigen, daß das Minimum existiert, und eine oder mehrere Funktionen $u : \Omega \to \mathbb{R}^N$ zu finden, für die dieses Minimum angenommen wird.

Zahlreiche Probleme der reinen und angewandten Mathematik und ihren Anwendungen in Physik, Technik, Biologie, Chemie lassen sich auf diese Weise formulieren, und natürlich ist die Variationsrechnung aus diesen Anwendungen heraus entstanden. Sie ist ein klassisches Teilgebiet der Mathematik.

Ihre Ursprünge kann man bis zu Zenodor zurückverfolgen, der 200 Jahre vor Chr. lebte, und der das Problem der Isoperimetrischen Ungleichungen studierte. Fermat studierte das Problem der Brechung von Lichtstrahlen und entdeckte 1662, daß der Weg des Lichtes zwischen zwei Punkten derjenige Weg ist, den es in der kürzesten Zeit zurücklegt. Seither haben viele der besten Mathematiker sich mit Problemen der Variationsrechnung beschäftigt. Der Name "Variationsrechnung" für dieses Gebiet stammt von Euler (1707–1783). Seit dieser Zeit ist die Variationsrechnung ein sehr lebendiges und sich schnell entwickelndes Teilgebiet der Mathematik geblieben. Der Ursprung der Funktionalanalysis ist mit der Variationsrechnung verknüpft, und heutzutage steht die Variationsrechnung im engsten Zusammenhang mit der Theorie der partiellen Differentialgleichungen und der Kontrolltheorie.

1.2 Beispiele für Variationsprobleme. Ich will nun eine Reihe von Beispielen für Variationsprobleme angeben.

1.2.1 Das Fermatsche Prinzip. Ein Lichtstrahl verlaufe in der Ebene vom Punkt (a, α) zum Punkt (b, β), wobei ich $a < b$ voraussetze. Der vom Ort abhängige Brechungsindex des Mediums, das der Lichtstrahl durchquert, sei $n(x, y)$. Der Lichtweg kann dann durch den Graphen einer Funktion

$u : [a,b] \to \mathbb{R}$ mit $u(a) = \alpha$, $u(b) = \beta$ beschrieben werden. Da $\frac{1}{n(x,y)}$ die Ausbreitungsgeschwindigkeit des Lichtes im Punkt (x,y) ist, ist die Zeit, die das Licht von (a,α) nach (b,β) braucht, gegeben durch

$$t = \int_a^b n(x,u(x)) \sqrt{1+u'(x)^2}\, dx\,.$$

Unter allen möglichen Wegen von (a,α) nach (b,β) wählt das Licht nun einen solchen, für den diese Zeit minimal ist. Also gilt für den Lichtweg mit

$$f(x,u,\xi) = n(x,u)\sqrt{1+\xi^2}\,,$$
$$I(u) = \int_a^b f(x,u(x),u'(x))\, dx$$

und

$$D = \{u \in C_1([a,b],\mathbb{R}) : v(a) = \alpha\,,\ v(b) = \beta\}\,,$$

daß $u \in D$ und

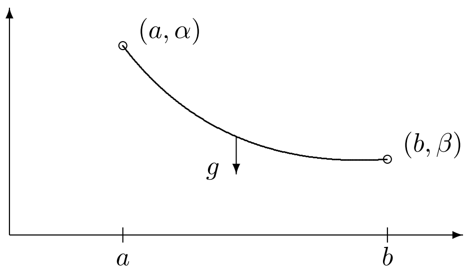
$$I(u) = \min_{v\in D} I(v)\,.$$

Falls die Lösung dieses Variationsproblems eindeutig ist, also falls es nur ein $u \in D$ gibt, für das das Funktional I das Minimum annimmt, kann man durch Lösung des Variationsproblems auch den Lichtweg bestimmen. Andererseits ist durchaus nicht sicher, ob das so formulierte Variationsproblem überhaupt eine Lösung hat. Denn wenn der Lichtstrahl auf dem Weg von (a,α) nach (b,β) von Luft in Wasser übergeht, dann hat der Lichtweg an der Wasseroberfläche einen Knick und kann also nicht durch den Graphen einer Funktion in $C_1([a,b])$ dargestellt werden. Man wird also in diesem Fall nicht erwarten, daß das oben formulierte Variationsproblem lösbar ist. Der Grund ist natürlich, daß der Brechungsindex sich in diesem Fall nicht stetig ändert, sondern an der Wasseroberfläche vom kleineren Wert für Luft auf den größeren Wert für Wasser springt. Damit dieses Variationsproblem lösbar ist, wird man also mindestens voraussetzen müssen, daß der Brechungsindex $n(x,y)$ eine stetige Funktion ist.

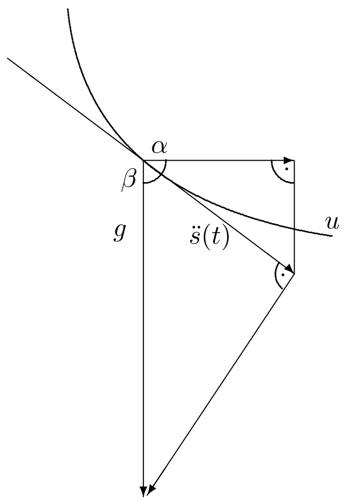
Andererseits gilt das Fermatsche Prinzip auch für den Übergang eines Lichtstrahles von Luft nach Wasser, und es stellt sich die Frage, ob man das Variationsproblem nicht auch in diesem Fall lösbar machen kann, indem man einen größeren Raum D zuläßt, der auch stückweise stetig differenzierbare

Funktionen enthält. Dies ist tatsächlich der Fall, wenn man für D den Sobolevraum $H_1([a,b])$ von "schwach differenzierbaren" Funktionen zuläßt, und wie sich zeigen wird, ist es vom mathematischen Problem her viel natürlicher, Lösungen in $H_1([a,b])$ zu suchen als in $C_1([a,b])$.

1.2.2 Das Brachistochronenproblem. Gesucht ist die Bahn, die ein sich in einer Ebene bewegender Massenpunkt unter dem Einfluß der Gravitation nehmen nuß, um in der kürzesten Zeit von einem gegebenen Punkt (a,α) zu einem gegebenen Punkt (b,β) zu gelangen.



Um das Variationsfunktional zu diesem Problem aufzustellen, sei die gesuchte Bahn durch den Graphen der Funktion $u : [a,b] \to \mathbb{R}$ gegeben. Auf den Massenpunkt wirkt die Erdbeschleunigung g in vertikaler Richtung.



Die Bahnbeschleunigung des Massenpunktes ist gleich der Tangentialkomponente $g\cos\beta$ von g an die Bahn. Wenn s die Bogenlänge der Bahn zwischen (a,α) und dem Ort des Massenpunktes ist, folgt also für den Massenpunkt zur Zeit t an der Stelle $s(t)$:

$$
\begin{aligned}
\ddot{s}(t) &= \frac{d^2s}{dt^2}(t) = g\cos\beta = g\cos\big(\frac{\pi}{2}-\alpha\big) = g\sin\alpha \\
&= g\,\frac{\tan\alpha}{\sqrt{1+\tan^2\alpha}} = -g\,\frac{u'(x)}{\sqrt{1+u'(x)^2}}\,,
\end{aligned}
$$

wobei $x = x(t)$ die x–Komponente des Ortes des Massenpunktes zur Zeit t ist. Man beachte nun, daß

$$
\begin{aligned}
\frac{d}{ds}\dot{s}(s) &= \frac{d}{ds}\Big[\frac{ds}{dt}(t(s))\Big] = \frac{d^2s}{dt^2}(t(s))\cdot\frac{dt}{ds}(s) \\
&= \ddot{s}(t(s))\,\frac{1}{\frac{ds}{dt}(t(s))} = \ddot{s}(s)\,\frac{1}{\dot{s}(s)}\,,
\end{aligned}
$$

also

$$
\Big[\frac{d}{ds}\,\dot{s}(s)\Big]\dot{s}(s) = \frac{d}{ds}\Big(\frac{1}{2}\,\dot{s}(s)^2\Big) = \ddot{s}(s)\,,
$$

und somit ergibt sich für die Bahngeschwindigkeit

$$
\begin{aligned}
\dot{s}(s)^2 &= 2\int_0^s \ddot{s}(\sigma)d\sigma + \dot{s}(0)^2 = 2\int_0^s \ddot{s}(\sigma)ds \\
&= 2\int_0^s \Big(\frac{-gu'(x(\sigma))}{\sqrt{1+u'(x(\sigma))^2}}\Big)d\sigma = 2\int_0^{x(s)} \frac{-gu'(x)}{\sqrt{1+u'(x)^2}}\,\frac{ds}{dx}\,dx \\
&= 2g\int_0^{x(s)} (-u'(x))dx = -2g[u(x(s))-\alpha] = -2g\,u(x(s))\,.
\end{aligned}
$$

Hierbei sei $x(s)$ die x–Komponente des Massenpunktes. Ich habe benützt, daß die Geschwindigkeit $\dot{s}(0)$ am Punkt (a,α) verschwindet. Die Vereinfachung im letzten Schritt ergibt sich, wenn man das Koordinatensystem so wählt, daß $\alpha = 0$ gilt.

Wenn ℓ die Bogenlänge zwischen (a,α) und (b,β) ist, ergibt sich also für die Laufzeit T zwischen diesen Punkten

$$
\begin{aligned}
T &= t(\ell)-t(0) = \int_0^\ell \frac{dt}{ds}(s)\,ds = \int_0^\ell \frac{1}{\frac{ds}{dt}(t(s))}\,ds = \int_0^\ell \frac{1}{\dot{s}(s)}\,ds \\
&= \int_0^\ell \frac{1}{\sqrt{-2gu(x(s))}}\,ds = \int_a^b \frac{1}{\sqrt{-2gu(x)}}\,\frac{ds}{dx}\,dx = \int_a^b \sqrt{\frac{1+u'(x)^2}{-2gu(x)}}\,dx\,.
\end{aligned}
$$

Zur Lösung des Brachistochronenproblems setze also $m = N = 1$,

$$f(x,u,\xi) = \sqrt{\frac{1+\xi^2}{-2gu}}$$

und

$$D = \{v \in C_1([a,b],\mathbb{R}) : v(a) = 0\,,\ v(b) = \beta\}\,.$$

Gesucht ist $u \in D$ mit

$$I(u) = \int_a^b f(x,u(x),u'(x))\,dx = \min_{v\in D} I(v)\,.$$

1.2.3 Systeme von Massenpunkten. Gegeben seien ℓ Massenpunkte mit den Massen $m_1,\dots,m_\ell$. Zur Zeit t sei

$$u_i(t) = (x_i(t), y_i(t), z_i(t)) \in \mathbb{R}^3$$

der Ort des i–ten Massenpunktes. In der gegebenen Zeit τ bewege sich dieser Punkt vom gegebenen Ort $u_i^{(0)}$ zum gegebenen Ort $u_i^{(1)}$. In der Zeit zwischen $t = 0$ und $t = \tau$ bewegen sich dann die Massenpunkte so, daß das Integral

$$\int_0^\tau T(u'(t)) - U(t,u(t))\,dt = \int_0^\tau \frac{1}{2}\sum_{i=1}^{\ell} m_i|u_i'(t)|^2 - U(t,u(t))\,dt$$

einen Minimalwert annimmt, wobei ich

$$u(t) = (u_1(t),\dots u_\ell(t)) \in \mathbb{R}^{3\ell}$$

gesetzt habe, $U(t,u(t))$ die potentielle Energie des Systems von Massenpunkten zur Zeit t sei und

$$T(u'(t)) = \frac{1}{2}\sum_{i=1}^{\ell} m_i|u_i'(t)|^2$$

die kinetische Energie bezeichne. In diesem Fall erhält man also ein Variationsproblem mit $n = 1$, $N = 3\ell$,

$$f(t,u,\xi) = \frac{1}{2}\sum_{i=1}^{\ell} m_i|\eta_i|^2 - U(t,u)\,,\ \xi = (\eta_1,\dots,\eta_\ell) \in \mathbb{R}^N\,,$$

und

$$D = \{u \in C_1([0,\tau], \mathbb{R}^N) : u(0) = u^{(0)}, u(\tau) = u^{(1)}\}\,.$$

Gesucht ist $u \in D$ mit

$$I(u) = \int_0^\tau f(t, u(t), u'(t))dt = \min_{v \in D} I(v)\,.$$

Die Aussage, daß sich die Massenpunkte so bewegen, daß $\int_0^\tau T - U\,dt$ einen Minimalwert annimmt, heißt Hamiltonsches Prinzip.

1.2.4 Dirichletsches Integral. Sei $n \geq 1$, sei $\Omega \subseteq \mathbb{R}^n$ ein Gebiet mit Rand $\partial\Omega$ und sei $u_0 : \partial\Omega \to \mathbb{R}$ eine gegebene Funktion. Gesucht ist eine Funktion $u : \overline{\Omega} \to \mathbb{R}$, die am Rand mit u_0 übereinstimmt, und für die das "Dirichletsche Integral"

$$I(u) = \int_\Omega \frac{1}{2}\, |\nabla u(x)|^2\, dx$$

einen Minimalwert annimmt. Mit $N = 1$,

$$D = \{v \in C_1(\overline{\Omega}, \mathbb{R}) : v\big|_{\partial\Omega} = u_0\}$$

und

$$f(x, u, \xi) = \frac{1}{2}\, |\xi|^2$$

ist also $u \in D$ gesucht, so daß

$$I(u) = \int_\Omega f(\nabla u(x))\, dx = \min_{v \in D} I(v)$$

gilt.

Dieses Problem hat in der Entwicklung der Variationsrechnung und der Theorie der partiellen Differentialgleichungen eine entscheidende Rolle gespielt. Es ist ein Modellproblem der Variationsrechnung, weil viele Eigenschaften komplizierterer Variationsprobleme sich an diesem Problem in einfacher Form studieren lassen. Wie jedem Variationsproblem kann man diesem Problem ein Randwertproblem zu einer gewöhnlichen oder partiellen Differentialgleichung zuordnen, so daß unter gewissen Zusatzbedingungen die

Lösung des Variationsproblems eine Lösung des Randwertproblems ist. In diesem Fall ist das Randwertproblem

$$\Delta u(x) = 0 \quad \text{in } \Omega$$

$$u|\partial\Omega = u_0\,,$$

wobei $\Delta = \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2}$ der Laplaceoperator ist. Die Gleichung $\Delta u(x) = 0$ heißt Potentialgleichung, weil zum Beispiel das Newtonsche Gravitationspotential und das elektrostatische Potential diese Gleichung erfüllen.

Der erwähnte Zusammenhang zwischen der Variationsrechnung und der Theorie der partiellen Differentialgleichungen beruht auf dieser Zuordnung von Randwertproblemen zu Variationsproblemen, weil man durch Lösung des Variationsproblems das Randwertproblem lösen kann.

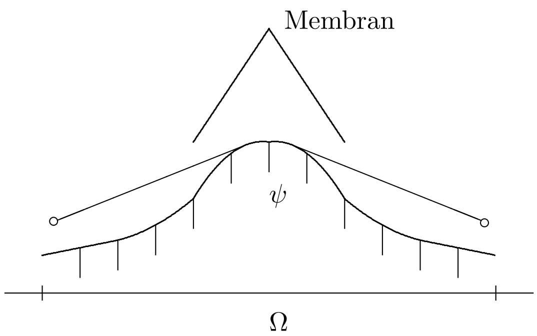
1.2.5 Hindernisprobleme. In den letzten Jahrzehnten haben Hindernisprobleme in der Variationsrechnung eine große Rolle gespielt. Ich will an folgendem physikalischen Beispiel erläutern, worum es sich dabei handelt:
Sei $\Omega \subseteq \mathbb{R}^2$ ein beschränktes Gebiet, und sei $u_0 : \partial\Omega \to \mathbb{R}$ eine gegebene Funktion. Der Graph dieser Funktion beschreibe einen geschlossenen Drahtbügel. Also sei

$$x \mapsto (x, u_0(x)) : \partial\Omega \to \mathbb{R}^3$$

ein "Parametrisierung des Drahtbügels". An diesem Drahtbügel sei eine elastische Membran befestigt. Außerdem sei eine Funktion $\psi : \overline{\Omega} \to \mathbb{R}$ gegeben mit

$$\psi(x) \leq u_0(x)\,, \quad x \in \partial\Omega\,.$$

Der Graph dieser Funktion beschreibe die Oberfläche eines Hindernisses, das unter Umständen in den Bereich der Membran hineinragt und die Membran nach oben drückt. In gewissen Bereichen liegt die Membran auf dem Hindernis auf. Gesucht ist die Form, die die Membran annimmt.



Das Problem läßt sich in einfacher Weise als Variationsproblem formulieren. An der Stelle $x \in \overline{\Omega}$ sei $u(x)$ die Höhe der Membran über der Ebene. Es gilt dann $u(x) = u_0(x)$ für $x \in \partial\Omega$. Die potentielle Energie $V(u)$ der elastischen Membran ist

$$I(u) = \int_{\Omega} \frac{1}{2} |\nabla u(x)|^2 \, dx \, .$$

Die Membran nimmt diejenige Form an, bei der diese potentielle Energie minimal ist. Natürlich ist die Menge der möglichen Formen, die die Membran annehmen kann, von der Hindernisoberfläche mitbestimmt, und man erhält folgendes Variationsproblem: Sei

$$D : \{v \in C_1(\overline{\Omega}, \mathbb{R}) : v|_{\partial\Omega} = u_0 \text{ und } v(x) \geq \psi(x) \text{ für alle } x \in \Omega\} \, .$$

Gesucht ist $u \in D$ mit

$$I(u) = \min_{v \in D} I(v) \, .$$

Vom vorangehenden Beispiel 1.2.4 unterscheidet sich dieses Problem nur durch die andere Wahl der Menge D von zulässigen Vergleichsfunktionen. Die jetzige Menge D ist kleiner als die Menge vom Beispiel 1.2.4.

1.2.6 Nichtparametrische Minimalflächen. Das Minimalflächenproblem ist berühmt, und die Theorie zu diesem Problem hat sich in den letzten Jahren schnell entwickelt und eine große Anziehungskraft ausgeübt. In seiner einfachsten Form geht es darum, zu einer gegebenen geschlossenen Kurve im $\mathbb{R}^3$ eine Fläche mit der gegebenen Kurve als Rand zu finden, die den kleinsten Flächeninhalt hat. Die genaue Formulierung des Problems hängt davon ab, was für Flächen man zuläßt. Ich will das Problem erst für Flächen formulieren, die als Graph einer Funktion dargestellt werden können.

Wie im vorangehenden Beispiel sei $\Omega \subseteq \mathbb{R}^2$ ein Gebiet, $u_0 : \partial\Omega \to \mathbb{R}$ eine Funktion, und

$$x \mapsto (u, u_0(x)) : \partial\Omega \to \mathbb{R}^3$$

die Parametrisierung einer geschlossenen Kurve. Man stelle sich unter dieser Kurve einen Drahtbügel vor. Taucht man diesen Drahtbügel in Seifenwasser, dann spannt sich eine Seifenhaut zwischen diesem Drahtbügel aus, die unter allen möglichen Flächen mit demselben Rand den kleinsten Flächeninhalt hat. Ich nehme an, daß diese Fläche als Graph einer Funktion $u : \overline{\Omega} \to \mathbb{R}$ beschrieben werden kann. Es gilt dann $u|_{\partial\Omega} = u_0$. Der Flächeninhalt dieser Fläche ist

$$I(u) = \int_\Omega \sqrt{1 + |\nabla u(x)|^2}\, dx\,.$$

Sei $D = \{v \in C_1(\overline{\Omega}, \mathbb{R}) : v|_{\partial\Omega} = u_0\}$. In diesem Variationsproblem ist also $u \in D$ gesucht mit

$$I(u) = \int_\Omega \sqrt{1 + |\nabla u(x)|^2}\, dx = \min_{v \in D} I(v)\,.$$

1.2.7 Parametrisierte Minimalflächen. Nicht jede Fläche kann als Graph einer Funktion dargestellt werden, und deswegen muß man im allgemeinen Fall mit Minimalflächen, die Mannigfaltigkeiten sind, arbeiten. Ich nehme der Einfachheit halber an, daß Σ eine Minimalfäche sei, die durch eine einzige Karte $u : \overline{\Omega} \to \Sigma \subseteq \mathbb{R}^3$ parametrisiert werden kann, wobei $\Omega \subseteq \mathbb{R}^2$ eine offene Menge sei. Sei also

$$\Sigma = \{u(x,y) \in \mathbb{R}^3 : (x,y) \in \overline{\Omega}\}\,.$$

Hierbei habe ich Σ mit der Bildmenge von u identifiziert. Der Flächeninhalt von Σ ist

$$I(u) = \int_\Omega |u_x(x,y) \times u_y(x,y)|\, d(x,y)\,,$$

wobei $u_x \times u_y$ das Vektorprodukt der Tangentialvektoren $u_x = \frac{\partial}{\partial x} u$ und $u_y = \frac{\partial}{\partial y} u$ sei. Sei $u_0 : \partial\Omega \to \mathbb{R}^3$ eine Parametrisierung der Randkurve von Σ, und sei

$$D = \{v \in C_1(\overline{\Omega}, \mathbb{R}^3) : v|_{\partial\Omega} = u_0\}\,.$$

In diesem Variationsproblem ist $u \in D$ gesucht, so daß

$$I(u) = \int_\Omega |u_x \times u_y|\, d(x,y) = \min_{v \in D} I(v)$$

1.2.8 Isoperimetrische Ungleichung. In der isoperimetrischen Ungleichung für ebene Gebiete wird die Länge der Randkurve des Gebietes mit dem Flächeninhalt des Gebietes verglichen. Genau lautet das Problem folgendermaßen: Sei $\Omega \subseteq \mathbb{R}^2$ eine offene beschränkte Menge, deren Rand $\partial\Omega$ eine einfach geschlossene, "hinreichend reguläre" Kurve ist. Wenn $L(\partial\Omega)$ die Länge der Randkurve und $\operatorname{meas}(\Omega)$ das Lebesguemaß von Ω ist, dann gilt nach der Isoperimetrischen Ungleichung

$$(L(\partial\Omega))^2 \geq 4\pi \operatorname{meas}(\Omega)\,,$$

und Gleichheit tritt in dieser Ungleichung genau dann ein, wenn Ω ein Kreis ist. Mit den Methoden der Variationsrechnung kann diese Ungleichung bewiesen werden. Dazu muß die Aussage im Rahmen, der in Abschnitt 1.1 angegeben wurde, formuliert werden. Hierzu sei $u : [a,b] \to \mathbb{R}^2$ eine Parametrisierung von $\partial\Omega$. Nach dem Gaußschen Satz gilt

$$\begin{aligned}
\operatorname{meas}\Omega &= \int_\Omega dy = \frac{1}{2}\int_\Omega \Big(\frac{\partial}{\partial y_1}\,y_1 + \frac{\partial}{\partial y_2}\,y_2\Big)\,d(y_1,y_2) \\
&= \frac{1}{2}\int_\Omega (\mathrm{div\ y})\,\mathrm{dy} = \frac{1}{2}\int_{\partial\Omega} \mathrm{n(y)}\cdot \mathrm{y\,ds_y} \\
&= \frac{1}{2}\int_a^b \frac{1}{\sqrt{u_1'(x)^2+u_2'(x)^2}} \begin{pmatrix} u_2'(x) \\ -u_1'(x)\end{pmatrix} \cdot \begin{pmatrix} u_1(x) \\ u_2(x)\end{pmatrix} \sqrt{u_1'(x)^2+u_2'(x)^2}\,dx \\
&= \frac{1}{2}\int_a^b (u_1(x)u_2'(x) - u_2(x)u_1'(x))\,dx\,.
\end{aligned}$$

Außerdem gilt

$$L(\partial\Omega) = \int_a^b \sqrt{u_1'(x)^2+u_2'(x)^2}\,dx\,.$$

Für $A > 0$ sei also

$$D(A) = \{v \in C_1([a,b],\mathbb{R}^2) : v(a) = v(b) \text{ und } \int_a^b (v_1v_2' - v_2v_1')dx = 2A\}$$

und sei

$$I(v) = \int_a^b \sqrt{v_1'(x)^2+v_2'(x)^2}\,dx\,.$$

Gesucht ist $u \in D(A)$ mit

$$I(u) = \min_{v \in D(A)} I(v)\,.$$

Weiterhin muß gezeigt werden, daß $I(u) = \sqrt{4\pi A}$, daß das Minimum u eindeutig ist, und daß u die Parametrisierung eines Kreises ist.

1.3 Klassische und direkte Methoden der Variationsrechnung. Das Problem der Variationsrechnung besteht darin, das Minimum eines Funktionals auf einer Teilmenge eines Funktionenraumes zu finden, der normalerweise unendlichdimensional ist. Aus der Infinitesimalrechung ist das Problem wohlbekannt, das Minimum einer reellwertigen Funktion auf einer Teilmenge des $\mathbb{R}^m$ zu finden. Gegenüber diesem Problem ergeben sich in den Minimumproblemen der Variationsrechnung neue Schwierigkeiten, die hauptsächlich in der Unendlichdimensionalität des Definitionsbereiches des Funktionals begründet sind.

Grundsätzlich jedoch kann man zur Lösung von Variationsproblemen genauso vorgehen wie im Problem, das Minimum einer Funktion auf einer Teilmenge des $\mathbb{R}^m$ zu finden. Deswegen will ich die Lösungswege zu diesem Problem kurz besprechen:

Eine Möglichkeit, das Minimum einer Funktion $f : U \subseteq \mathbb{R}^m \to \mathbb{R}$ zu finden, besteht darin, die stationären Punkte von f zu bestimmen, das heißt die Punkte $x \in U$ mit

$$\nabla f(x) = 0\,.$$

Durch Studium der höheren Ableitungen von f in den stationären Punkten können dann Kriterien gefunden werden, mit denen man entscheiden kann, ob ein stationärer Punkt von f auch ein Minimum von f ist oder nicht. Die zweite Methode ist, f direkt zu minimieren: Man wählt eine Minimalfolge aus, das heißt eine Folge $\{x_n\}_{n=1}^{\infty}$ mit $x_n \in U$ und

$$\lim_{n\to\infty} f(x_n) = \inf_{x \in U} f(x)\,.$$

Dann muß gezeigt werden, daß $\{x_n\}_{n=1}^{\infty}$ eine in U konvergente Teilfolge besitzt. Hierzu benötigt man Kompaktheitskriterien. Nach dem Satz von Bolzano-Weierstraß genügt es zum Beispiel, daß $\{x_n\}_{n=1}^{\infty}$ eine beschränkte Folge ist und U abgeschlossen ist, um eine konvergente Teilfolge $\{x_{n_\ell}\}_{\ell=1}^{\infty}$

auswählen zu können. Wenn f stetig ist, gilt dann für den Grenzwert x_0 dieser Folge

$$f(x_0) = f(\lim_{\ell \to \infty} x_{n_\ell}) = \lim_{\ell \to \infty} f(x_{n_\ell}) = \inf_{x \in U} f(x)\,,$$

also ist x_0 eine Minimalstelle von f.

Auch in der Variationsrechnung werden beide Methoden angewandt. Die erste Methode wurde in der Entstehung der Variationsrechnung zuerst entwickelt und heißt daher auch klassische Methode der Variationsrechnung. Die Gleichung $\nabla f(x) = 0$ wird dabei ersetzt durch die gewöhnliche oder partielle Differentialgleichung

$$\sum_{i=1}^{n} \frac{\partial}{\partial x_i} \left[f_{\xi_i}(x, u(x), \nabla u(x)) \right] = f_u(x, u(x), \nabla u(x))\,, \quad x \in \Omega\,,$$

die als notwendige Bedingung von jeder Lösung u des Variationsproblems erfüllt werden muß, die zu $C_2(\Omega)$ gehört, und die Eulersche Differentialgleichung genannt wird. Zu dieser Differentialgleichung kommen in der Regel noch weitere Bedingungen hinzu, die die Lösung des Variationsproblems erfüllen muß und die durch die Wahl des Raumes D der zulässigen Vergleichsfunktionen bestimmt ist. Häufig tritt als Bedingung zum Beispiel auf, daß die Lösung u am Rande $\partial\Omega$ mit einer vorgegebenen Funktion übereinstimmen muß:

$$u\big|_{\partial\Omega} = u_0\,.$$

Diese beiden Bedingungen zusammen bilden ein Randwertproblem für die Eulersche Differentialgleichung, und jede Lösung $u \in C_2(\Omega)$ des Variationsproblems, also jedes zweimal stetig differenzierbare Minimum des Variationsfunktionales muß dieses Randwertproblem erfüllen. Genau wie im Fall einer Funktion auf dem $\mathbb{R}^n$ ist es aber nicht sicher, ob eine Lösung dieses Randwertproblems auch ein Minimum des Variationsfunktionales ist. Analog zu den Kriterien für Funktionen im $\mathbb{R}^n$, die garantieren, daß ein stationärer Punkt auch ein Minimum der Funktion ist, und die man durch Untersuchung der höheren Ableitungen von f erhält, sind in der Variationsrechnung eine ganze Reihe verschiedener Kriterien entwickelt worden, die garantieren, daß eine Lösung des Randwertproblems zur Eulerschen Differentialgleichung auch ein Minimum des Variationsfunktionales ist. Solche Kriterien sind hauptsächlich für den Fall $n = 1$ aufgestellt worden. Dazu gehören die Theorie der Felder und die Kriterien von Jacobi, Weierstraß, Zermelo, Legendre und andere.

Abgesehen davon, daß diese Kriterien häufig schwer nachzuprüfen sind, bestehen die Probleme der klassischen Methode der Variationsrechnung darin, daß man von einem Minimum des Funktionals verlangt, daß es zweimal stetig differenzierbar sein soll, und daß man das Minimum dann als Lösung eines Randwertproblems für eine Differentialgleichung zweiter Ordnung bestimmt. Minima von Variationsfunktionalen sind aber häufig gar nicht zweimal stetig differenzierbar, und überdies ist die Lösung eines Randwertproblems für eine Differentialgleichung zweiter Ordnung oft eine schwierige Aufgabe. Wie man heute weiß, führt man auf diese Weise in vielen Fällen ein leichter zu lösendes Problem auf ein schwieriger zu lösendes "zurück".

Beginnend mit Hilbert sind daher seit der Jahrhundertwende die direkten Methoden der Variationsrechnung entwickelt worden, die dem oben besprochenen zweiten Zugang zur Auffindung von Minima von Funktionen auf dem $\mathbb{R}^n$ entsprechen. Man versucht, ein Minimum des Variationsfunktionales dadurch zu bestimmen, daß man aus einer Minimalfolge zum Variationsfunktional eine konvergente Teilfolge aussucht. Die Schwierigkeit ist nun, daß man für den Fall von unendlichdimensionalen Funktionenräumen kein so einfaches Kompaktheitskriterium zur Verfügung hat wie das von Bolzano-Weierstraß im endlichdimensionalen Fall. Aus diesem Grund arbeitet man mit Hilbert- und Banachräumen, und daher speziell mit Sobolevräumen. Allgemein wendet man Methoden der Funktionalanalysis an und ersetzt das Bolzano-Weierstraßsche Kompaktheitskriterium durch verschiedene Kompaktheitskriterien aus dieser Theorie. Dazu gehört der Kompaktheitssatz von Rellich für Sobolevräume oder die schwache Kompaktheit der Einheitskugel des Dualraumes eines Banachraumes. Wie erwähnt, hat die Funktionalanalysis ihren Ursprung in der Theorie der Hilberträume, die Hilbert gerade für den Zweck entwickelt hat, spezielle Variationsfunktionale, und insbesondere das Dirichletsche Integral, zu minimieren.

1.4 Inhalt der Vorlesung. Mein Interesse in dieser Vorlesung gilt mehr den modernen direkten Methoden als den klassischen. Die direkten Methoden spielen heute eine große Rolle in der Mathematik. Ich habe schon angedeutet, daß man heute zur Lösung von Randwertproblemen zu elliptischen partiellen Differentialgleichungen den umgekehrten Weg geht wie in der klassischen Methode der Variationsrechnung: Zu einer gegebenen elliptischen partiellen

Differentialgleichung sucht man ein Variationsfunktional, dessen Eulergleichung gerade die gegebene elliptische Differentialgleichung ist. Dann löst man das Variationsproblem mit direkten Methoden. Das Minimum des Variationsfunktionals löst dann gleichzeitig auch die gegebene elliptische Differentialgleichung. Mein Ziel mit dieser Vorlesung ist daher auch auf die Theorie der elliptischen Differentialgleichungen vorzubereiten, die ich im nächsten Semester in einer Vorlesung behandeln möchte.

Ich habe schon erwähnt, daß die Theorie der Sobolevräume für die direkten Methoden der Variationsrechnung und für die moderne Theorie der partiellen Differentialgleichungen eine zentrale Rolle spielt. Wie im Titel meiner Vorlesung angekündigt, möchte ich daher in den beiden anschließenden Kapiteln eine recht ausführliche Einführung in die Theorie der L_p– und Sobolevräume geben. Daraufhin will ich die klassischen Methoden der Variationsrechnung behandeln, mich dabei aber kurz fassen. Behandeln möchte ich dabei aber den Formalismus von Hamilton-Jacobi. Für die direkten Methoden der Variationsrechnung braucht man Kompaktheitskriterien. Eine große Rolle spielt dabei die Kompaktheit der Einheitskugel des Dualraumes eines Banachraumes bezüglich der schwachen Topologie. Zur Vorbereitung auf die direkten Methoden werde ich daher im sechsten Kapitel die schwache Topologie und im siebten konvexe Funktionale auf Hilberträumen behandeln. Die Ergebnisse dieser beiden Kapitel werde ich im achten Kapitel bei der Behandlung der direkten Methoden der Variationsrechnung anwenden.

2. L^p–Räume, Sobolevräume

2.1 Skalarprodukt. Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$. Eine Abbildung $(\cdot,\cdot) : V \times V \to \mathbb{K}$ heißt Skalarprodukt, falls gilt:

(i) Für alle $x, y \in V : (x,y) = \overline{(y,x)}$

(ii) Für alle $\lambda \in \mathbb{K}$, $x, y \in V : (\lambda x, y) = \lambda(x,y)$

(iii) Für alle $x_1, x_2, y \in V : (x_1 + x_2, y) = (x_1, y) + (x_2, y)$.

(iv) $(x,x) \geq 0$, $(x,x) = 0 \Longleftrightarrow x = 0$.

Aus (i) und (ii) zusammen folgt

$$(x, \lambda y) = \overline{(\lambda y, x)} = \overline{\lambda(y,x)} = \overline{\lambda}\,\overline{(y,x)} = \overline{\lambda}(x,y)$$

und aus (i) und (iii) folgt

$$\begin{aligned}(x, y_1 + y_2) &= \overline{(y_1 + y_2, x)} = \overline{(y_1, x) + (y_2, x)} = \overline{(y_1, x)} + \overline{(y_2, x)} \\ &= (x, y_1) + (x, y_2)\,,\end{aligned}$$

also ist das Skalarprodukt linear im ersten und antilinear im zweiten Argument.

2.2 Lemma. Sei $(\cdot,\cdot)$ ein Skalarprodukt auf einem Vektorraum V. Dann gilt für alle $x, y \in V$

(i) $|(x,y)| \leq (x,x)^{1/2}(y,y)^{1/2}$ (Cauchy-Schwarzsche Ungleichung)

(ii) $(x+y\,,\,x+y)^{1/2} \leq (x,x)^{1/2} + (y,y)^{1/2}$ (Dreiecksungleichung)

(iii) $(x+y\,,\,x+y) + (x-y\,,\,x-y) = 2((x,x) + (y,y))$

(Parallelogrammgleichung)

Beweis: (i) Sei $(x,x) + (y,y) > 0$. O.b.d.A. nehmen wir an, daß $(y,y) > 0$ ist. Setze $z = \frac{y}{(y,y)^{1/2}}$. Dann folgt

$$\begin{aligned}0 \;\leq\; &(x - (x,z)z\,,\, x - (x,z)z) = (x,x) - ((x,z)z, x) \\ &- (x, (x,z)z) + ((x,z)z\,,\,(x,z)z)\end{aligned}$$

$$
\begin{aligned}
&= (x,x) - (x,z)\overline{(x,z)} - \overline{(x,z)}(x,z) + (x,z)\overline{(x,z)}(z,z) \\
&= (x,x) - 2|(x,z)|^2 + |(x,z)|^2\Big(\frac{y}{(y,y)^{1/2}}\,,\ \frac{y}{(y,y)^{1/2}}\Big) \\
&= (x,x) - |(x,z)|^2\,,
\end{aligned}
$$

also

$$|(x,z)|^2 \le (x,x)\,,$$

oder

$$\frac{1}{(y,y)}\,|(x,z)|^2 \le (x,x)\,,$$

somit $|(x,y)|^2 \le (x,x)(y,y)$. Dies ist (i).

(ii) Aus (i) folgt

$$
\begin{aligned}
(x+y\,,\,x+y) &= (x,x)+(x,y)+(y,x)+(y,y) \\
&\le (x,x)+2(x,x)^{1/2}(y,y)^{1/2}+(y,y) \\
&= [(x,x)^{1/2}+(y,y)^{1/2}]^2\,.
\end{aligned}
$$

(iii) $\quad (x+y\,,\,x+y)+(x-y\,,\,x-y)$

$= (x,x) + 2\,\mathrm{Re}\,(\mathrm{x},\mathrm{y}) + (\mathrm{y},\mathrm{y}) + (\mathrm{x},\mathrm{x}) - 2\,\mathrm{Re}\,(\mathrm{x},\mathrm{y}) + (\mathrm{y},\mathrm{y})$

$= 2((x,x)+(y,y))\,.$

2.3 Prähilbertraum. Auf dem Vektorraum V sei ein Skalarprodukt $(\cdot,\cdot):$ $V\times V\to\mathbb{K}$ erklärt. Dann wird durch

$$\|x\| := (x,x)^{1/2}$$

eine Norm auf V erklärt. Denn aus 2.1 (ii), (iv) und 2.2 (ii) folgt

(i) $\quad \|x\|\ge 0\,,\quad \|x\|=0 \Longleftrightarrow x=0$

(ii) $\quad \|\lambda x\| = |\lambda|\,\|x\|$

(iii) $\quad \|x+y\| = (x+y\,,\,x+y)^{1/2} \le (x,x)^{1/2}+(y,y)^{1/2} = \|x\|+\|y\|\,.$

Mit dieser Norm wird V zum normierten Raum. Dieser normierte Raum zusammen mit dem Skalarprodukt heißt Prähilbertraum.

2.4 Konvergenz, Cauchyfolgen, Vollständigkeit, Hilbertraum. Sei V ein normierter Raum, $\{x_n\}_{n=1}^{\infty} \subseteq V$ eine Folge, $x_0 \in V$. Dann heißt die Folge $\{x_n\}_{n=1}^{\infty}$ konvergent gegen x_0, wenn

$$\lim_{n\to\infty} \|x_n - x_0\| = 0\,.$$

$\{x_n\}_{n=1}^{\infty}$ heißt Cauchyfolge, wenn zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert mit

$$\|x_n - x_m\| < \varepsilon$$

für alle $n, m \geq n_0$.

Jede konvergente Folge ist eine Cauchyfolge. Ein Raum, in dem jede Cauchyfolge konvergiert, heißt vollständig. Ein vollständiger normierter Raum heißt Banachraum. Ein vollständiger Prähilbertraum heißt Hilbertraum.

2.5 Lebesgue-Räume. Sei $M \subseteq \mathbb{R}^n$ eine meßbare Menge, und sei μ das Lebesgue-Maß auf $\mathbb{R}^n$ Sei $1 \leq p < \infty$ und sei $\mathcal{L}^p(M,\mu,\mathbb{K})$ (kürzer: $\mathcal{L}^p(M)$) die Menge aller μ–meßbaren Funktionen $f : M \to \mathbb{K}$ mit

$$\int_M |f|^p d\mu = \int_M |f(x)|^p d\mu(x) < \infty\,.$$

$\mathcal{L}^p(M,\mu,\mathbb{K})$ ist ein Vektorraum über $\mathbb{K}$. Setze

$$\|f\|_p = \Big(\int_M |f|^p d\mu\Big)^{1/p}\,.$$

Für $p = \infty$ sei $\mathcal{L}^\infty(M,\mu,\mathbb{K})$ die Menge aller μ–meßbaren Funktionen
$f : M \to \mathbb{K}$ mit

$$\|f\|_\infty = \operatorname*{ess\,sup}_M |f| := \inf_{\mu(N)=0} \; \sup_{x\in M\setminus N} |f(x)| < \infty\,.$$

Auch $\mathcal{L}^\infty(M,\mu,\mathbb{K})$ ist ein Vektorraum über $\mathbb{K}$. Da abzählbare Vereinigungen von Nullmengen wieder Nullmengen sind, gibt es dann eine μ–Nullmenge N, so daß

$$\|f\|_\infty = \sup_{x\in M\setminus N} |f(x)|\,.$$

Beachte, daß $\|\cdot\|_p$ keine Norm ist auf $\mathcal{L}^p(M,\mu,\mathbb{K})$, weil $\|f\|_p = 0$ ist für alle f, die nur auf einer μ–Nullmenge von Null verschieden sind. Aus $\|f\|_p = 0$ folgt also nicht $f = 0$. Andererseits gilt für $1 \le p < \infty$ und für alle $n \in \mathbb{N}$

$$\|f\|_p \ge \frac{1}{n}\left[\mu(\{x \in M \mid |f(x)| \ge \frac{1}{n}\})\right]^{1/p}.$$

Aus $\|f\|_p = 0$ folgt also

$$\mu(\{x \in M \mid |f(x)| \ge \frac{1}{n}\}) = 0$$

für alle $n \in \mathbb{N}$, somit

$$\begin{aligned} \mu(\{x \in M \mid |f(x)| > 0\}) &= \mu\Big(\bigcup_{n=1}^{\infty}\{x \in M \mid |f(x)| \ge \frac{1}{n}\}\Big) \\ &= \lim_{m\to\infty}\mu\Big(\bigcup_{n=1}^{m}\{x \in M \mid |f(x)| \ge \frac{1}{n}\}\Big) \\ &= \lim_{m\to\infty}\mu\Big(\{x \in M \mid |f(x)| \ge \frac{1}{m}\}\Big) = \lim_{m\to\infty} 0 = 0\,. \end{aligned}$$

Also folgt für $1 \le p < \infty$

$$\|f\|_p = 0 \Longleftrightarrow f = 0 \qquad \mu - \text{fast überall.}$$

Diese Relation gilt auch für $p = \infty$. Dann wenn $\|f\|_\infty = 0$ ist, gibt es eine Nullmenge N mit

$$0 = \|f\|_\infty = \sup_{x \in M\setminus N} |f(x)|\,.$$

Mit dieser Relation können wir nun einen normierten Raum folgendermaßen definieren:
Für $1 \le p \le \infty$ sei $L^p(M,\mu,\mathbb{K})$ (kürzer: $L^p(M)$) die Menge aller Äquivalenzklassen in $\mathcal{L}^p(M,\mu,\mathbb{K})$ zur Äquivalenzrelation

$$f \sim g \Longleftrightarrow f = g \qquad \mu - \text{fast überall.}$$

Auf $L^p(M)$ werden die Vektorraumoperationen folgendermaßen erklärt: Seien $[f], [g] \in L^p(M)$ Äquivalenzklassen zu den Repräsentanten f und g. Setze

$$\begin{aligned} [f] + [g] &= [f+g] \\ \lambda[f] &= [\lambda f]\,, \quad \lambda \in \mathbb{K}\,. \end{aligned}$$

Hiermit wird $L^p(M, \mu, \mathbb{K})$ zum Vektorraum über $\mathbb{K}$. Außerdem setze

$$\|[f]\|_p := \|f\|_p = \Big(\int_M |f|^p d\mu \Big)^{1/p} .$$

$\| \cdot \|_p : L^p(M) \to [0, \infty]$ ist eine sinnvoll definierte Funktion, weil $\|f\|_p = \|g\|_p$ gilt für alle $f, g \in \mathcal{L}^p(M)$ mit $[f] = [g]$.
Außerdem gilt nun

$$\begin{aligned} \|[f]\|_p = 0 \iff \|f\|_p = 0 \quad &\iff \quad f = 0 \qquad \mu - \text{fast überall.} \\ &\iff \quad f \sim 0 \iff [f] = [0] . \end{aligned}$$

Daß $\| \cdot \|_p : L^p(M) \to [0, \infty]$ auch die anderen Eigenschaften einer Norm hat, werden wir unten zeigen. Für $L^\infty(M, \mu, \mathbb{K})$ ist dies jedoch offensichtlich. Also ist $L^\infty(M, \mu, \mathbb{K})$ ein normierter Raum.

Zur Vereinfachung der Schreibweise läßt man überlicherweise die Klammern in $[f]$ weg und benutzt für die Äquivalenzklasse und ihren Repräsentanten dasselbe Zeichen.

2.6 Lemma (Vollständigkeit von L^∞). $L^\infty(M, \mu, \mathbb{K})$ ist vollständig, also ist $L^\infty(M, \mu, \mathbb{K})$ ein Banachraum.

Beweis: Sei $\{f_k\}_{k=1}^\infty$ eine Cauchy-Folge in $L^\infty(M, \mu, \mathbb{K})$. Dann gibt es zu jedem $n \in \mathbb{N}$ ein $k_0(n)$ mit

$$\|f_k - f_l\|_\mu \leq \frac{1}{n} , \quad \|f_k\|_\infty \leq C < \infty$$

für alle $k, l \geq k_0$, also gibt es für alle $k, l \geq k_0$ eine Nullmenge $N(l, k)$ mit

$$\sup_{x \in M \setminus N(l,k)} |f_k(x) - f_l(x)| = \|f_k - f_l\|_\infty \leq \frac{1}{n} , \qquad \sup_{x \in M \setminus N(l,k)} |f_k(x)| \leq C ,$$

also

$$(*) \qquad \sup_{x \in M \setminus N} |f_k(x) - f_l(x)| \leq \frac{1}{n} , \quad \sup_{x \in M \setminus N} |f_k(x)| \leq C$$

für alle $n \in N$ und alle $k, l \geq k_0(n)$, mit

$$\begin{aligned} N(n) &= \bigcup \{N(l,k) \mid l,k \geq k_0(n)\} \\ N &= \bigcup_{n=1}^{\infty} N(n)\,. \end{aligned}$$

N ist eine Nullmenge als abzählbare Vereinigung von Nullmengen. Also ist insbesondere $\{f_k(x)\}_{k=1}^{\infty}$ eine Cauchyfolge in $\mathbb{K}$ für alle $x \in M \backslash N$. Da $\mathbb{K}$ vollständig ist, hat diese Folge einen Grenzwert $f(x)$, und es gilt für alle $x \in M \backslash N$

$$\begin{aligned} |f(x)| &= |f(x) - f_k(x) + f_k(x)| \\ &= \lim_{k\to\infty} |f(x) - f_k(x) + f_k(x)| \\ &\leq \underbrace{\lim_{k\to\infty} |f(x) - f_k(x)|}_{=0} + \lim_{k\to\infty} |f_k(x)| \;\leq\; \lim_{k\to\infty} C \;=\; C\,, \end{aligned}$$

wegen $(*)$. Definiert man F auf N durch

$$f(x) = 0\,, \quad x \in N\,,$$

dann ist also f beschränkt und meßbar, also $f \in L^\infty(m, \mu, \mathbb{K})$. Weiter folgt aus $(*)$ für alle $x \in M \backslash N$

$$\begin{aligned} |f(x) - f_k(x)| &= \lim_{l\to\infty} |f(x) - f_l(x) + f_l(x) - f_k(x)| \\ &\leq \lim_{l\to\infty} |f(x) - f_l(x)| + \lim_{l\to\infty} |f_l(x) - f_k(x)| \\ &\leq \lim_{l\to\infty} |f_l(x) - f_k(x)| \;\leq\; \lim_{l\to\infty} \frac{1}{n} \;=\; \frac{1}{n}\,, \end{aligned}$$

für alle $k \geq k_0$, also

$$\|f - f_k\|_\infty \leq \sup_{x \in M\backslash N} |f(x) - f_k(x)| \leq \frac{1}{n}\,,$$

$k \geq k_0$. Hieraus folgt $\lim_{k\to\infty} f_k = f$ in $L^\infty(M, \mu, \mathbb{K})$, also ist dieser Raum vollständig.

2.7 Lemma (Höldersche Ungleichung). Seien $p, q \in [1, \infty]$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Ist dann $f \in L^p(M)$ und $g \in L^q(M)$, dann ist $fg \in L^1(M)$ und

$$\|fg\|_1 \leq \|f\|_p \, \|g\|_q \, .$$

Beweis: fg ist meßbar als Produkt meßbarer Funktionen. Für $p = 1$ ist $q = \infty$, also

$$|(fg)(x)| = |f(x)g(x)| \leq \|g\|_\infty |f(x)|$$

für fast alle $x \in M$, also ist $|fg|$ fast überall durch eine integrierbare Funktion beschränkt, also ist $fg \in L^1(M)$. Weiter folgt

$$\int_M |f(x)g(x)| d\mu(x) \leq \int_M \|g\|_\infty \, |f(x)| d\mu(x) = \|g\|_\infty \, \|f\|_1 \, ,$$

und dies ist die Höldersche Ungleichung.

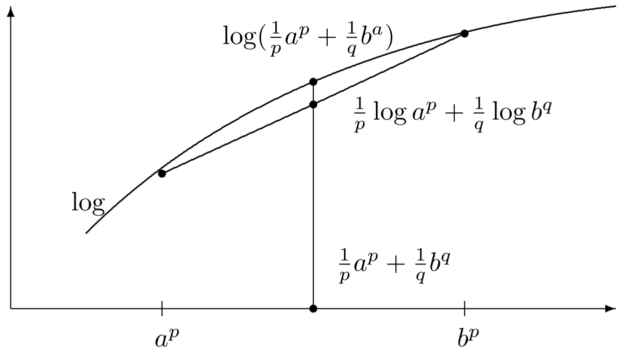
Die Behauptung ist auch klar für $\|f\|_p = 0$ oder $\|g\|_q = 0$, weil dann $f(x)\, g(x) = 0$ fast überall. Sei also $1 < p < \infty$ und $\|f\|_p > 0$ und $\|g\|_q > 0$. Für $a, b \geq 0$ gilt

$$ab \leq \frac{1}{p} a^p + \frac{1}{q} b^q \quad \text{(Youngsche Ungleichung).}$$

Für $ab = 0$ ist diese Ungleichung trivial. Zum Beweis für $a > 0$ und $b > 0$ bilde man den Logarithmus

$$\begin{aligned} \log ab &= \log a + \log b \; = \; \frac{p}{p} \log a + \frac{q}{q} \log b \\ &= \frac{1}{p} \log a^p + \frac{1}{q} \log b^q \; \leq \; \log \left(\frac{1}{p} a^p + \frac{1}{q} b^q \right) . \end{aligned}$$

Die letzte Ungleichung gilt wegen der Konkavität des Logarithmus. Denn wegen $\frac{1}{p}+\frac{1}{q}=1$ ist $\frac{1}{p}\,\log a^p+\frac{1}{q}\,\log b^q$ der in der Skizze eingezeichnete Wert auf der Sehne



Nun setze man $a=\frac{|f(x)|}{\|f\|_p}$, $b=\frac{|g(x)|}{\|g\|_q}$. Es folgt

$$\frac{|f(x)\,g(x)|}{\|f\|_p\,\|g\|_q}\leq\frac{|f(x)|^p}{p\|f\|_p^p}+\frac{|g(x)|^q}{q\|g\|_q^q}\,.$$

Die rechte Seite ist integrierbar, also auch die linke, und es folgt

$$\begin{aligned}\frac{\int|fg|d\mu}{\|f\|_p\,\|g\|_q} &\leq \frac{1}{p}\,\frac{\int|f|^p d\mu}{\|f\|_p^p}+\frac{1}{q}\,\frac{\int|g|^q d\mu}{\|g\|_q^q}\\ &= \frac{1}{p}\,\frac{\|f\|_p^p}{\|f\|_p^p}+\frac{1}{q}\,\frac{\|g\|_q^q}{\|g\|_q^q}=\frac{1}{p}+\frac{1}{q}=1\,.\end{aligned}$$

2.8 Lemma (Minkowski-Ungleichung). Sei $1\leq p\leq\infty$. Sind $f,g\in L^p(M)$, dann auch $f+g\in L^p(M)$ und

$$\|f+g\|_p\leq\|f\|_p+\|g\|_p\,.$$

Beweis: Für $p=1$ und $p=\infty$ ist diese Ungleichung trivial. Sei nun $1<p<\infty$. Dann folgt

$$(*)\qquad\begin{aligned}|f(x)+g(x)|^p &= |f(x)+g(x)|\,|f(x)+g(x)|^{p-1}\\ &\leq |f(x)|\,|f(x)+g(x)|^{p-1}+|g|\,|f(x)+g(x)|^{p-1}\end{aligned}$$

fast überall. Nach Voraussetzung sind $f,g\in L^p(M)$, und $|f+g|^{p-1}\in L^q(M)$ mit $\frac{1}{p}+\frac{1}{q}=1$, wegen

$$(|f+g|^{p-1})^q=|f+g|^{q(p-1)}=|f+g|^p\leq 2^{p-1}|f|^p+2^{p-1}|g|^p\,,$$

also sind nach der Hölderschen Ungleichung

$$|f|\,|f+g|^{p-1}\,,\ |g|\,|f+g|^{p-1} \in L^1(M)\,,$$

und damit wegen $(*)$ auch $|f+g|^p$. Dies bedeutet, daß

$$|f+g| \in L^p(M)\,.$$

Weiter folgt aus $(*)$ und aus der Hölderschen Ungleichung

$$\begin{aligned}
\int |f+g|^p d\mu &\leq \|f\|_p \,\|\,|f+g|^{p-1}\|_q + \|g\|_p \,\|\,|f+g|^{p-1}\|_q \\
&= (\|f\|_p + \|g\|_p)\Big(\int (|f+g|^{p-1})^q\, d\mu\Big)^{1/q} \\
&= (\|f\|_p + \|g\|_p)\Big(\int |f+g|^p\, d\mu\Big)^{1-\frac{1}{p}}\,.
\end{aligned}$$

Wenn $\int |f+g|^p\, d\mu = 0$ ist, ist die Behauptung klar. Im anderen Fall ergibt sie sich aus der letzten Ungleichung durch Kürzen.

Aus der Minkowskischen Ungleichung folgt nun, daß $\|\cdot\|_p : L^p(M,\mu,\mathbb{K}) \to [0,\infty)$ eine Norm ist für $1 \leq p \leq \infty$, und daß also $L^p(M,\mu,\mathbb{K})$ und auch l^p normierte Räume sind. Aus der Hölderschen Ungleichung und aus $f,g \in L^2(M,\mu,\mathbb{K})$ folgt $(f\overline{g}) \in L^1(M,\mu,\mathbb{K})$. Für $f,g \in L^2(M,\mu,\mathbb{K})$ existiert also das Integral

$$(f,g) = \int_M f(x)\,\overline{g(x)}\, d\mu(x)\,,$$

und stellt, wie man sofort nachprüft, ein Skalarprodukt dar mit

$$\|f\|_2 = (f,f)^{1/2}\,.$$

Also ist $L^2(M,\mu,\mathbb{K})$ ein Prähilbertraum. Im nächsten Satz wird gezeigt, daß $L^p(M,\mu,\mathbb{K})$ vollständig ist für alle p mit $1 \leq p \leq \infty$. Der Raum $L^p(M)$ ist also sogar ein Banachraum und $L^2(M)$ ist ein Hilbertraum.

2.9 Satz von Fischer-Riesz. $L^p(M,\mu,\mathbb{K})$ ist vollständig für $1 \leq p \leq \infty$.

Beweis: Sei $\{f_k\}_{k=1}^{\infty}$ eine Cauchy-Folge in $L^p(M)$. Es genügt zu zeigen, daß die Folge eine konvergente Teilfolge hat, weil der Grenzwert dann Häufungspunkt der Folge ist, und eine Cauchy-Folge gegen ihren Häufungspunkt konvergiert. Wähle eine Teilfolge $\{f_{k_i}\}_{i=1}^{\infty}$ aus mit

$$\sum_{i=1}^{\infty} \|f_{k_{i+1}} - f_{k_i}\|_p < \infty .$$

Um diese Teilfolge zu konstruieren, wähle $\tilde{k}_i$, so daß

$$\|f_k - f_l\|_p < 2^{-i}$$

gilt für $k, l \geq \tilde{k}_i$, und setze $k_i = \max(k_{i-1}+1, \tilde{k}_i)$. Im folgenden bezeichnen wir die Teilfolge $\{f_{k_i}\}_{i=1}^{\infty}$ wieder mit $\{f_k\}_{k=1}^{\infty}$. Setze

$$g_l = \sum_{k=1}^{l} |f_{k+1} - f_k| .$$

Wir zeigen nun, daß diese Folge punktweise μ–fast überall konvergiert. Für diejenigen x, für die $g_l(x)$ konvergiert, ist dann $\{f_k(x)\}_{k=1}^{\infty}$ eine Cauchy-Folge und besitzt also einen Grenzwert $f(x)$. Wir werden dann zeigen, daß $f \in L^p(M)$ ist, und daß $\{f_k\}_{k=1}^{\infty}$ auch in der Norm $\|\cdot\|_p$ gegen f konvergiert. Nach dem Lemma von Fatou und der Minkowskischen Ungleichung gilt wegen $g_l^p(x) \geq 0$

$$\begin{aligned}
\int_M \liminf_{l\to\infty} g_l^p(x)\, d\mu(x) &\leq \liminf_{l\to\infty} \int_M g_l^p(x)\, d\mu(x) \\
&= \liminf_{l\to\infty} \|g_l\|_p^p = \Big(\liminf_{l\to\infty} \|g_l\|_p\Big)^p \\
&\leq \Big(\liminf_{l\to\infty} \sum_{k=1}^{l} \|f_{k+1} - f_k\|_p\Big)^p \\
&= \Big(\sum_{k=1}^{\infty} \|f_{k+1} - f_k\|_p\Big)^p < \infty .
\end{aligned}$$

Aus dem Lemma von Fatou folgt dann, daß

$$\lim_{l\to\infty} g_l^p(x) = \liminf_{l\to\infty} g_l^p(x)$$

μ – fast überall. Also existiert auch

$$\lim_{k\to\infty} f_k(x) = f(x)$$

μ – fast überall. Da $|f(x) - f_l(x)|^p \geq 0$ ist, kann das Lemma von Fatou erneut angewendet werden, und es folgt

$$\begin{aligned}
\int_M |f(x) - f_l(x)|^p \, d\mu(x) &= \int_M \lim_{k\to\infty} |f_k(x) - f_l(x)|^p \, d\mu(x) \\
&= \int_M \liminf_{k\to\infty} |f_k(x) - f_l(x)|^p \, d\mu(x) \\
&\leq \liminf_{k\to\infty} \int_M |f_k(x) - f_l(x)|^p \, d\mu(x) \\
&= (\liminf_{k\to\infty} \|f_k - f_l\|_p)^p \\
&\leq \Big(\liminf_{k\to\infty} \sum_{i=l}^{k-1} \|f_{i+1} - f_i\|_p\Big)^p \\
&= \Big(\sum_{i=l}^{\infty} \|f_{i+1} - f_i\|_p\Big)^p \to 0 \quad \text{für } l \to \infty\,.
\end{aligned}$$

Hieraus folgt $f - f_l \in L^p(M)$, und damit nach der Minkowskischen Ungleichung auch $f \in L^p(M)$. Außerdem folgt hieraus, daß $\{f_k\}_{k=1}^{\infty}$ in $L^p(M)$ gegen f konvergiert. Damit ist der Satz bewiesen.

2.10 Schwache Ableitung. Sei $\Omega \subseteq \mathbb{R}^n$ eine offene Menge; sei $1 \leq p \leq \infty$, und sei

$$\mathring{C}_\infty(\Omega) = \{\varphi \in C^\infty(\mathbb{R}^n) \mid \operatorname{supp}\varphi \subseteq \Omega\,,\ \operatorname{supp}\varphi \text{ kompakt}\}\,.$$

Seien $u, v \in L^p(\Omega)$. v heißt schwache Ableitung von u nach x_i, wenn

$$\int_\Omega v(x)\,\varphi(x) dx = -\int_\Omega u(x)\,\frac{\partial}{\partial x_i}\,\varphi(x)\,dx$$

gilt für alle $\varphi \in \mathring{C}_\infty(\Omega)$.

2.11 Lemma: Seien v_1, v_2 schwache Ableitungen nach x_i von u. Dann gilt $v_1 = v_2$.

Beweis: Für alle $\varphi \in \overset{\circ}{C}_\infty(\Omega)$ gilt

$$\int_\Omega v_1(x)\,\varphi(x)\,dx = -\int_\Omega u(x)\,\frac{\partial}{\partial x_i}\,\varphi(x)dx = \int_\Omega v_2(x)\,\varphi(x)\,dx\,,$$

also

$$\int_\Omega (v_1 - v_2)\varphi\,dx = 0\,.$$

Die Behauptung folgt aus dieser Formel und aus dem später folgenden Lemma 3.7, falls $v_1 - v_2 \in L^{1,\mathrm{loc}}(\Omega)$ gilt. Also bleibt dies zu beweisen.
Hierzu sei $\Omega_1 \subseteq \Omega$ eine kompakte Menge. Dann folgt aus der Hölderschen Ungleichung für $v_1 - v_2 \in L^p(\Omega)$, $1 \le p \le \infty$,

$$\int_\Omega |v_1 - v_2|dx \le \Big[\int_{\Omega_1} 1^q\,dx\Big]^{\frac{1}{q}}\Big[\int_{\Omega_1} |v_1 - v_2|^p\,dx\Big]^{\frac{1}{p}} < \infty\,,$$

und somit ist $v_1 - v_2 \in L^{1,\mathrm{loc}}(\Omega)$. Dies beweist das Lemma.

Nach Definition ist für $u \in C_1(\Omega)$ die gewöhnliche Ableitung auch schwache Ableitung, also stimmen starke und schwache Ableitungen überein. Schwache Ableitungen sind also Verallgemeinerungen von gewöhnlichen (starken) Ableitungen und werden daher auch mit $\frac{\partial}{\partial x_i}u$, $\partial_i u$ oder mit u' bezeichnet. Entsprechend werden höhere schwache Ableitungen definiert. Hierzu sei $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$ ein Multiindex. Man setzt $|\alpha| = \sum_{i=1}^n \alpha_i$ und

$$D^\alpha\,\varphi(x) = \frac{\partial^{|\alpha|}}{\partial^{\alpha_1}x_1 \ldots \partial^{\alpha_n}x_n}\,\varphi(x)\,.$$

Seien $u, v \in L^p(\Omega)$. v heiße α–te schwache Ableitung von u, wenn für alle $\varphi \in \overset{\circ}{C}_\infty(\Omega)$ gilt

$$\int_\Omega v(x)\varphi(x)dx = (-1)^{|\alpha|}\int_\Omega u(x)D^\alpha\varphi(x)dx\,.$$

2.12 Sobolevräume $\boldsymbol{H^p_m(\Omega)}$. Für $1 \le p \le \infty$ und $m \in \mathbb{N}_0$ sei

$$H^p_m(\Omega) = \left\{ f \in L^p(\Omega) \;\middle|\; \begin{array}{l} f \text{ hat schwache Ableitungen } D^\alpha f \in L^p(\Omega) \\ \text{bis zur Ordnung } |\alpha| \le m \end{array} \right\} .$$

$H^p_m(\Omega)$ ist ein Vektorraum. Eine Norm auf $H^p_m(\Omega)$ ist

$$\|f\|_{p,m,\Omega} = \sum_{|\alpha| \le m} \|D^\alpha f\|_{p,\Omega} .$$

Für $p = 2$ ist ein Skalarprodukt auf $H_m(\Omega) = H^2_m(\Omega)$ definiert durch

$$(f, g)_{m,\Omega} = \sum_{|\alpha| \le m} (D^\alpha f, D^\alpha g)_\Omega ,$$

mit

$$(u, v)_\Omega = \int_\Omega u(x)\, \overline{v(x)}\, dx .$$

Die zugehörige Norm ist

$$\|f\|_{2,m,\Omega} = \sqrt{\sum_{|\alpha| \le m} \|D^\alpha f\|^2_{2,\Omega}} ,$$

die äquivalent ist zur oben definierten Norm $\|f\|_{2,m,\Omega}$, und die wir deshalb der Einfachheit halber mit demselben Symbol bezeichnen.

2.13 Satz (Vollständigkeit der Sobolevräume). Die Räume $H^p_m(\Omega)$ sind vollständig, also ist jeder der Räume $H^p_m(\Omega)$ ein Banachraum, und $H^2_m(\Omega)$ ist ein Hilbertraum.

Beweis: Sei $\{f_l\}_{l=1}^\infty \subseteq H^p_m(\Omega)$ eine Cauchy-Folge. Zu jedem $\varepsilon > 0$ gibt es dann ein l_0 mit

$$\|f_l - f_k\|_{p,m,\Omega} < \varepsilon$$

für alle $l, k \ge l_0$. Hieraus folgt für jeden Multiindex α mit $|\alpha| \le m$

$$\|D^\alpha f_l - D^\alpha f_k\|_{p,\Omega} \le \|f_l - f_k\|_{p,m,\Omega} < \varepsilon ,$$

also ist $\{D^\alpha f_l\}_{l=1}^\infty$ eine Cauchy-Folge in $L^p(\Omega)$. $L^p(\Omega)$ ist vollständig. Sei $f^{(\alpha)} \in L^p(\Omega)$ der Grenzwert von $\{D^\alpha f_l\}_{l=1}^\alpha$. Dann folgt für jedes $\varphi \in \overset{\circ}{C}_\infty(\Omega)$

$$\begin{aligned}
\int_\Omega f^\alpha(x)\,\varphi(x)\,dx &= \int_\Omega \Big[L^p - \lim_{l\to\infty} D^\alpha f_l(x)\Big]\,\varphi(x)\,dx \\
&= \lim_{l\to\infty} \int_\Omega D^\alpha f_l(x)\,\varphi(x)\,dx \\
&= \lim_{l\to\infty} (-1)^{|\alpha|} \int_\Omega f_l(x) D^\alpha \varphi(x)\,dx \\
&= (-1)^{|\alpha|} \int_\Omega \Big[L^p - \lim_{l\to\infty} f_l(x)\Big]\, D^\alpha\varphi(x)\,dx \\
&= (-1)^{|\alpha|} \int_\Omega f(x) D^\alpha \varphi(x)\,dx\,.
\end{aligned}$$

Also ist f^α die α–te Ableitung von $f, f^\alpha = D^\alpha f$, somit ist $f \in H_m^p(\Omega)$. Schließlich gilt

$$\lim_{l\to\infty} \|f - f_l\|_{p,m,\Omega} = \sum_{|\alpha|\le m} \lim_{l\to\infty} \|D^\alpha f - D^\alpha f_l\|_{p,m,\Omega} = 0\,,$$

d.h. $\{f_l\}_{l=1}^\infty$ konvergiert in $H_m^p(\Omega)$ gegen f. Also ist $H_m^p(\Omega)$ vollständig.

Wir setzen für alle $1 \le p \le \infty$ und $m \in \mathbb{N}_0 \cup \{\infty\}$

$$C_m^p(\Omega) = \left\{ f : \Omega \to \mathbb{K} \;\middle|\; \begin{array}{l} f \text{ ist } m\text{–fach stetig differenzierbar} \\ \text{mit } D^\alpha f \in L^p(\Omega) \text{ für alle} \\ \text{Multiindizes } \alpha \text{ mit } |\alpha| \le m\,. \end{array} \right\}$$

2.14 Satz (Dichtheit von $C_\infty^p(\Omega)$im Sobolevraum). Sei $1 \le p < \infty$. Dann gilt

$$\overline{C_\infty^p(\Omega)}^{\|\cdot\|_{m,p}} = H_m^p(\Omega)\,.$$

Das heißt, daß jedes $f \in H_m^p(\Omega)$ durch eine Folge $\{f_l\}_{l=1}^\infty \subseteq C_\infty^p(\Omega)$ bezüglich der Norm $\|\cdot\|_{p,m,\Omega}$ approximiert werden kann.

Ich werde diesen Satz später beweisen. Dieser Satz gilt nicht für $p = \infty$. Denn

$$\overline{C_m^\infty(\Omega)}^{\|\cdot\|_{m,\infty}} = C_m^\infty(\Omega)\,,$$

also
$$\overline{C_\infty^\infty(\Omega)}^{\|\cdot\|_{1,\infty}} = C_1^\infty(\Omega)\,.$$
Für $\Omega = (-1,1)$ ist
$$f(x) := |x|$$
in $H_1^\infty(\Omega)$ enthalten mit
$$f'(x) = \begin{cases} -1, x < 0 \\ +1, x \geq 0 \end{cases},$$
aber natürlich ist
$$f \notin C_1^\infty(\Omega)\,,$$
also
$$\overline{C_\infty^\infty}^{\|\cdot\|_{1,\infty}} \subseteq H_1^\infty(\Omega)\,.$$
Die Elemente von $H_m^p(\Omega)$ sind als Elemente von $L^p(\Omega)$ Äquivalenzklassen von meßbaren Funktionen. Etwas unpräzise sagt man, $f \in H_m^p(\Omega)$ sei stetig oder differenzierbar, falls die Äquivalenzklasse von f eine stetige oder differenzierbare Funktion enthält. Die Äquivalenzklasse von f enthält höchstens eine stetige oder differenzierbare Funktion. Alle anderen Funktionen aus dieser Äquivalenzklasse unterscheiden sich dann von der stetigen Funktion nur auf einer Menge vom Maß Null. Zwar ist $H_m^p(\Omega)$ eine Teilmenge von $L^p(\Omega)$, aber natürlich braucht nicht jedes Element f von $H_m^p(\Omega)$ stetig oder gar differenzierbar zu sein. Man kann aber beweisen, daß f stetig oder differenzierbar ist, falls m genügend groß ist (Sobolevscher Einbettungssatz).

Bei der Lösung von Randwertproblemen für partielle Differentialgleichungen möchte man haben, daß die Lösung, die man häufig in Sobolevräumen sucht, gewisse Werte auf dem Rand annimmt. Wenn der Rand einer offenen Menge Ω glatt ist, ist er eine Nullmenge, und somit ist nicht klar, was die Randwerte einer "Funktion" (Äquivalenzklasse) $u \in H_m^p(\Omega)$ sein sollen. Man kann aber unter gewissen Voraussetzungen an $\partial\Omega$ solche Randwerte definieren. Das folgende ist eine besonders elegante Definition von Funktionen in Sobolevräumen, die "im verallgemeinerten Sinn" auf dem Rand $\partial\Omega$ verschwinden.

2.15 Definition (Sobolevräume $\mathring{H}_m^p(\Omega)$). Für $1 \leq p < \infty$ und $m \in \mathbb{N}_0$ sei
$$\mathring{H}_m^p(\Omega) \;=\; \overline{\mathring{C}_\infty(\Omega)}^{H_m^p(\Omega)}$$

$$= \left\{ f \in H_m^p(\Omega) \;\middle|\; \begin{array}{l} \text{Es gibt eine Folge } \{\varphi_l\}_{l=1}^{\infty} \subseteq \overset{\circ}{C}_\infty(\Omega) \\ \text{mit } \|f - \varphi_l\|_{p,m,\Omega} \to 0 \text{ für } l \to \infty\,. \end{array} \right\}$$

$\overset{\circ}{H}{}_m^p(\Omega)$ ist ein abgeschlossener Unterraum des Banachraumes $H_m^p(\Omega)$, also selbst ein Banachraum. Für $p = 2$ insbesondere ist $\overset{\circ}{H}{}_m^2(\Omega) = \overset{\circ}{H}_m(\Omega)$ ein abgeschlossener Unterraum des Hilbertraumes $H_m(\Omega)$, also selbst ein Hilbertraum mit dem Skalarprodukt

$$(u, v)_{m,\Omega} = \sum_{|\alpha| \le m} (D^\alpha u, D^\alpha v)_\Omega\,.$$

3. Topologie der Soboleväume

3.1 Satz (Faltung). Sei $1 \leq p \leq \infty$, $\varphi \in L^1(\mathbb{R}^n)$, $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{K}$ mit $x \to f(y,x) \in L^p(\mathbb{R}^n)$ für alle $y \in \mathbb{R}^n$ und $\sup_{h \in \operatorname{supp}\varphi} \|f(\cdot, \cdot - h)\|_p < \infty$. Dann existiert für fast alle $x \in \mathbb{R}^n$ das Integral

$$F(x) = \int_{\mathbb{R}^n} \varphi(x-y)\, f(x,y)dy = \int_{\mathbb{R}^n} \varphi(y)\, f(x, x-y)dy\,,$$

und es gilt

$$\|F\|_p \leq \|\varphi\|_1 \sup_{h \in \operatorname{supp}\varphi} \|f(\cdot, \cdot - h)\|_p\,.$$

Beweis: Sei $p = \infty$. Dann folgt

$$\begin{aligned}
\operatorname*{ess\,sup}_{x \in \mathbb{R}^n} |F(x)| &= \operatorname*{ess\,sup}_{x \in \mathbb{R}^n} \Big| \int_{\mathbb{R}^n} \varphi(y)\, f(x, x-y)dy \Big| \\
&\leq \operatorname*{ess\,sup}_{x \in \mathbb{R}^n} \int_{\operatorname{supp}\varphi} |\varphi(y)|\, |f(x, x-y)|dy \\
&\leq \int_{\operatorname{supp}\varphi} |\varphi(y)| \sup_{h \in \operatorname{supp}\varphi} \operatorname*{ess\,sup}_{x \in \mathbb{R}^n} |f(x, x-h)|dy \\
&= \|\varphi\|_1 \sup_{h \in \operatorname{supp}\varphi} \|f(\cdot, \cdot - h)\|_\infty\,.
\end{aligned}$$

Hieraus folgt die Behauptung.

Sei $p < \infty$, und sei $1 < q \leq \infty$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Dann folgt aus der Hölderschen Ungleichung

$$\begin{aligned}
\int_{\mathbb{R}^n} |F(x)|^p dx &= \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} \varphi(x-y)\, f(x,y)dy \right|^p dx \\
&\leq \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} |(\varphi(x-y)|^{\frac{1}{q}}\, |\varphi(x-y)|^{\frac{1}{p}}\, |f(x,y)|\, dy \right|^p dx \\
&\leq \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} |\varphi(x-y)|dy \right]^{\frac{p}{q}} \left[\int_{\mathbb{R}^n} |\varphi(x-y)|\, |f(x,y)|^p\, dy \right] dx \\
&= \|\varphi\|_1^{\frac{p}{q}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |\varphi(y)|\, |f(x, x-y)|^p dx\, dy
\end{aligned}$$

$$\begin{aligned} &\leq \|\varphi\|_1^{\frac{p}{q}} \int_{\mathbb{R}^n} |\varphi(y)| \left[\sup_{y \in \operatorname{supp} \varphi} \int_{\mathbb{R}^n} |f(x, x-y)|^p dx \right] dy \\ &= \|\varphi\|_1^{\frac{p}{q}+1} \sup_{y \in \operatorname{supp} \varphi} \|f(\cdot, \cdot - y)\|_p^p . \end{aligned}$$

Also, wegen $\frac{1}{p} + \frac{1}{q} = 1$

$$\|F\|_p \leq \|\varphi\|_1 \sup_{y \in \operatorname{supp} \varphi} \|f(\cdot, \cdot - y)\|_p .$$

Wenn $f(y, x)$ nicht von y abhängt, schreibt man auch

$$F(x) = \int_{\mathbb{R}^n} \varphi(x-y)\, f(y) dy = (\varphi * f)(x) .$$

Der Operator $*$ heißt Faltung von φ und f . In diesem Fall nimmt die bewiesene Ungleichung die Form an

$$\|\varphi * f\|_p \leq \|\varphi\|_1 \, \|f\|_p .$$

Die Faltung ist also eine lineare und stetige Abbildung auf $L^p(\mathbb{R}^n)$ mit Norm $\leq \|\varphi\|_1$.

3.2 Lemma. Sei $f \in L^p(\mathbb{R}^n)$ und $\varphi \in \overset{\circ}{C}_\infty(\mathbb{R}^n)$. Dann ist $F = \varphi * f \in C^p_\infty(\mathbb{R}^n)$, und die partiellen Ableitungen erhält man durch Ableiten unter dem Integral.

Beweis: Übung.

Im folgenden sei $\Omega \subseteq \mathbb{R}^n$ eine offene Menge und

$$\overset{\circ}{C}(\Omega) = \{f \in C(\mathbb{R}^n) \mid \operatorname{supp} \mathrm{f} \subseteq \Omega\, ,\ \operatorname{supp} \mathrm{f} \text{ kompakt}\} .$$

3.3 Lemma. Sei $1 \leq p < \infty$ und $f \in L^p(\mathbb{R}^n)$. Dann gilt

$$\lim_{h \to 0} \|f(\cdot - h) - f\|_p = 0 .$$

Beweis. Aus der Lebesgueschen Integrationstheorie folgt, daß $\overset{\circ}{C}(\mathbb{R}^n)$ dicht ist in $L^p(\mathbb{R}^n)$ für $1 \leq p < \infty$. Also existiert zu $f \in L^p(\mathbb{R}^n)$ eine Folge $\{f_k\}_{k=1}^{\infty} \subseteq \overset{\circ}{C}(\mathbb{R}^n)$ mit $\|f - f_k\|_p \to 0$ für $k \to \infty$. Dann gilt

$$\begin{aligned}
\|f(\cdot - h) - f\|_p &\leq \|f(\cdot - h) - f_k(\cdot - h)\|_p + \|f_k(\cdot - h) - f_k\|_p + \|f_k - f\|_p \\
&\leq 2\,\|f - f_k\|_p + \left[\int_{\mathrm{supp\,f_k \cup supp\,f_k(\cdot - h)}} |f_k(x-h) - f_k(x)|^p dx\right]^{1/p} \\
&\leq 2\,\|f - f_k\|_p + \sup_{x \in R^n} |f_k(x-h) - f_k(x)| \left[\int_{\mathrm{supp\,f_k \cup supp\,f_k(\cdot - h)}} dx\right]^{1/p} .
\end{aligned}$$

Da f_k stetig ist auf $\mathbb{R}^n$ und außerhalb einer kompakten Menge verschwindet, ist f_k gleichmäßig stetig auf $\mathbb{R}^n$. Also gilt $\sup_{x \in \mathbb{R}^n} |f_k(x-h) - f_k(x)| \to 0$ für $h \to 0$. Zu $\varepsilon > 0$ wähle man k groß genug, so daß $\|f - f_k\|_p < \varepsilon/2$. Für dieses k wähle man h_0 so klein, daß

$$\sup_{x \in \mathbb{R}^n} |f_k(x-h) - f_k(x)| < \varepsilon/(2\,[2\int_{\mathrm{supp\,f_k}} dx]^{1/p})$$

gilt für alle h mit $|h| \leq h_0$. Es folgt

$$\|f(\cdot - h) - f\|_p < \varepsilon$$

für alle $|h| \leq h_0$.

3.4 Dirac-Familie. Eine Familie $\{\varphi_\varepsilon\}_{\varepsilon > 0} \subseteq L^1(\mathbb{R}^n)$ mit

$$\varphi_\varepsilon \geq 0\,, \quad \int_{\mathbb{R}^n} \varphi_\varepsilon(x)dx = 1\,, \quad \int_{\mathbb{R}^n \setminus B_\delta(0)} \varphi_\varepsilon(x)dx \to 0 \text{ für alle } \delta > 0$$

heißt Dirac-Familie.

Als Beispiel für eine solche Familie nehme man $\varphi \in L^1(\mathbb{R}^n)$ mit $\varphi \geq 0$ und mit $\int_{\mathbb{R}^n} \varphi(x)dx = 1$ und setze

$$\varphi_\varepsilon(x) = \varepsilon^{-n} \varphi\left(\frac{x}{\varepsilon}\right) .$$

3.5 Satz. Sei $1 \leq p < \infty$, und sei $\{\varphi_\varepsilon\}_{\varepsilon > 0}$ eine Dirac-Familie. Dann gilt für alle $f \in L^p(\mathbb{R}^n)$

$$\lim_{\varepsilon \to 0} \|\varphi_\varepsilon * f - f\|_p = 0 .$$

Beweis: Wegen $\int_{\mathbb{R}^n} \varphi_\varepsilon(x)dx = 1$ folgt

$$(\varphi_\varepsilon * f)(x) - f(x) = \int_{\mathbb{R}^n} \varphi_\varepsilon(x-y)\,[f(y)-f(x)]\,dy\,.$$

Für $\delta > 0$ sei

$$\varphi_{\varepsilon\delta}(x) = \begin{cases} \varphi_\varepsilon(x) & , \quad |x| < \delta \\ 0 & , \quad |x| \geq \delta \end{cases}\,, \quad \psi_{\varepsilon\delta}(x) = \begin{cases} 0 & , \quad |x| < \delta \\ \varphi_\varepsilon(x) & , \quad |x| \geq \delta \end{cases}\,.$$

Dann folgt

$$\begin{aligned} \|\varphi_\varepsilon * f - f\|_p = \| & \int_{\mathbb{R}^n} \varphi_{\varepsilon\delta}(x-y)\,[f(y)-f(x)]\,dy \\ & + \int_{\mathbb{R}^n} \psi_{\varepsilon\delta}(x-y)\,[f(y)-f(x)]\,dy\|_p \\ \leq \underbrace{\|\varphi_{\varepsilon\delta}\|_1}_{\leq 1} & \underbrace{\sup_{|h|\leq\delta} \|f(\cdot - h) - f\|_p}_{\substack{\to 0\,,\ \text{nach Lemma 4.3} \\ \text{für } \delta \to 0}} + \underbrace{\|\psi_{\varepsilon\delta}\|_1}_{\substack{\to 0 \\ \text{für } \varepsilon \to 0}} \underbrace{\sup_{h \in \mathbb{R}^n} \|f(\cdot - h) - f\|_p}_{\leq 2\|f\|_p}\,. \end{aligned}$$

3.6 Satz (Dichtheit von $\overset{\circ}{C}_\infty(\Omega)$ in $L^p(\Omega)$). Sei $\Omega \subseteq \mathbb{R}^n$ offen. Dann ist $\overset{\circ}{C}_\infty(\Omega)$ dicht in $L^p(\mathbb{R}^n)$ für $1 \leq p < \infty$.

Beweis: Sei $\psi \in \overset{\circ}{C}_\infty(\mathbb{R}^n)$ mit $\psi \geq 0$, $\psi \neq 0$, $\psi(x) = 0$ für $|x| \geq 1$.

(Beispiel:

$$\psi(x) = \begin{cases} \exp\left(-\dfrac{1}{1-|x|^2}\right) & , \quad |x| < 1 \\ 0 & , \quad |x| \geq 1\,. \quad) \end{cases}$$

Setze

$$\varphi(x) = \frac{1}{\|\psi\|_1}\,\psi(x)\,, \quad \varphi_\varepsilon(x) = \varepsilon^{-n}\varphi(\frac{x}{\varepsilon})\,.$$

Sei $f \in L^p(\mathbb{R}^n)$ und $\theta > 0$. Finde $f_\varepsilon \in \overset{\circ}{C}_\infty(\Omega)$ mit $\|f - f_\varepsilon\|_p \leq \theta$. Setzt man $f_\varepsilon = \varphi_\varepsilon * f$, dann ist $f_\varepsilon \in C_\infty(\Omega)$ mit $\|f - f_\varepsilon\|_p \to 0$ für $\varepsilon \to 0$, hat aber im allgemeinen keinen kompakten Träger in Ω. Daher setze für $\delta > 0$

$$\Omega_\delta = \left\{ x \in \Omega \mid \mathrm{dist}(\mathrm{x}, \partial\Omega) > \delta\,,\ |\mathrm{x}| \leq \frac{1}{\delta} \right\}$$

und

$$f_\delta(x) = \begin{cases} f(x) & , \quad x \in \Omega_\delta \\ 0 & , \quad x \in \mathbb{R}^n \backslash \Omega_\delta\,. \end{cases}$$

Es folgt

$$\|f - f_\delta\|_p^p = \int_\Omega |f(x) - f_\delta(x)|^p dx \to 0$$

für $\delta \to 0$, wegen $|f(x) - f_\delta(x)|^p \to 0$ für $\delta \to 0$ und alle $x \in \Omega$, und wegen $|f(x) - f_\delta(x)|^p \leq |f(x)|^p$ (Satz von Lebesgue). Für hinreichend kleines δ ist also $\|f - f_\delta\|_p < \theta/2$. Nun setze

$$f_\varepsilon = \varphi_\varepsilon * f_\delta\,.$$

Es existiert $\varepsilon_0 > 0$ mit

$$\|f_\delta - f_\varepsilon\|_p < \theta/2$$

für alle $\varepsilon \leq \varepsilon_0$. Außerdem gilt für $\varepsilon < \delta/2$ und $x \in \Omega\backslash\Omega_{\delta/2}$

$$\begin{aligned} |f_\varepsilon(x)| &= \Big| \int_{\mathbb{R}^n} \varphi_\varepsilon(y)\, f_\delta(x-y)dy \Big| \\ &\leq \int_{|y|\leq\varepsilon} |\varphi_\varepsilon(y)|\, |f_\delta(x-y)|dy = 0 \end{aligned}$$

wegen

$$\mathrm{dist}(\mathrm{x}-\mathrm{y}\,,\ \partial\Omega) \leq \varepsilon + \mathrm{dist}(\mathrm{x}, \partial\Omega) < \varepsilon + \frac{\delta}{2} < \delta\,,$$

für $|y| \leq \varepsilon$, also $f_\delta(x-y) = 0$. Also ist $f_\varepsilon \in \overset{\circ}{C}_\infty(\Omega)$ mit

$$\|f - f_\varepsilon\|_p < \theta\,.$$

3.7 Lemma (Fundamentallemma der Variationsrechnung). Sei $\Omega \subseteq \mathbb{R}^n$ offen und $g \in L^{1,loc}(\Omega)$ mit

$$\int_\Omega g(x)\, \varphi(x)dx = 0$$

für alle $\varphi \in \overset{\circ}{C}_\infty(\Omega)$. Dann folgt $g(x) = 0$ fast überall in Ω .

Beweis: Sei $E \subseteq \Omega$ meßbar und beschränkt, und sei $\overline{E} \subseteq \Omega$, also $\mathrm{dist(E, \mathbb{R}^n \backslash \Omega)} > 0$. Sei $\{\varphi_\varepsilon\}_{\varepsilon>0}$ wie im vorangehenden Beweis und setze

$$\phi_\varepsilon(x) = \int_E \varphi_\varepsilon(x-y)dy = \varphi_\varepsilon * \chi_E \,.$$

Dann folgt $\phi_\varepsilon \to \chi_E$ in $L^p(\mathbb{R}^n)$ nach Satz 4.5. Außerdem gilt

$$(*) \qquad |\phi_\varepsilon(x)| \le \int_E \varphi_\varepsilon(x-y)dy \le \int_{\mathbb{R}^n} \varphi_\varepsilon(x-y)dy = 1 \,.$$

Weil $\int_{\mathbb{R}^n} |\phi_\varepsilon(x) - \chi_E(x)|dx \to 0$ gilt für $\varepsilon \to 0$, also $|\phi_\varepsilon(x) - \chi_E(x)|$ in der L^1–Norm gegen Null konvergiert, gibt es nach einem Resultat aus der Lebesgueschen Integrationstheorie eine Folge ϕ_{ε_k} mit $\phi_{\varepsilon_k}(x) - \chi_E(x) \to 0$ fast überall in $\mathbb{R}^n$. Zusammen mit $(*)$ folgt aus dem Satz von Lebesgue somit

$$\int_\Omega (g\phi_{\varepsilon_k} - g\chi_E)dx \to 0 \,,$$

also, wegen $\phi_{\varepsilon_k} \in \overset{\circ}{C}_\infty(\Omega)$ für k genügend groß,

$$\int_E g\,dx = \lim_{k\to\infty} \int_\Omega g\,\phi_{\varepsilon_k}\,dx = 0 \,.$$

Nun sei Ω_δ wie im vorangehenden Beweis und setze

$$E_\pm = \{x \in \Omega_\delta \mid \pm g(x) > 0\}$$

(bei reellwertigem g; sonst betrachte $\mathrm{Re}\, g$ und $\mathrm{Im}\, g$ getrennt). Es folgt

$$\int_{\Omega_\delta \cap B_{\frac{1}{\delta}(0)}} |g|dx = \int_{E_+} g\,dx - \int_{E_-} g\,dx = 0 \,.$$

Da $\Omega = \bigcup\limits_{\delta>0} \Omega_\delta$ folgt $g = 0$ f.ü. auf Ω .

3.8 Approximation von Funktionen in Sobolevräumen. Sei $1 \le p < \infty$, sei $\Omega \subseteq \mathbb{R}^n$ offen und zu $\varphi \in \overset{\circ}{C}_\infty(\overline{B_1(0)})$ sei

$$\varphi_\varepsilon(x) = \varepsilon^{-n}\varphi(\frac{x}{\varepsilon})$$

eine Dirac-Familie. Für $u \in H_m^p(\Omega)$ setze

$$u_\varepsilon(x) = \varphi_\varepsilon * u = \int_\Omega \varphi_\varepsilon(x-y)\, u(y)dy\,.$$

Dann gilt $u_\varepsilon \in C_\infty(\mathbb{R}^n)$ und $\|u - u_\varepsilon\|_{p,\Omega} \to 0$ für $\varepsilon \to 0$. Außerdem gilt für $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| \le m$

$$D^\alpha u_\varepsilon(x) = \int_\Omega D_x^\alpha\, \varphi_\varepsilon(x-y)\, u(y)dy = (-1)^{|\alpha|} \int_\Omega \left[D_y^\alpha \varphi_\varepsilon(x-y)\right] u(y)dy\,.$$

Da $\operatorname{supp} \varphi_\varepsilon \subseteq \overline{\mathrm{B}_\varepsilon(0)}$ gilt, ist

$$y \mapsto \varphi_\varepsilon(x-y) \in \overset{\circ}{C}_\infty(\Omega)$$

für alle $x \in \Omega_\delta = \{x \in \Omega \mid \operatorname{dist}(\mathrm{x}, \partial\Omega) > \delta\}$ mit $\delta > \varepsilon$. Also gilt nach Definition der schwachen Ableitungen von u

$$\begin{aligned} D^\alpha u_\varepsilon(x) &= (-1)^{|\alpha|} \int_\Omega \left[D_y^\alpha\, \varphi_\varepsilon(x-y)\right] u(y)dy \\ &= \int_\Omega \varphi_\varepsilon(x-y)\, D^\alpha u(y)\, dy = (D^\alpha u)_\varepsilon(x)\,, \end{aligned}$$

also

$$[D^\alpha u_\varepsilon]\,|_{\Omega_\delta} = [(D^\alpha u)_\varepsilon]\,|_{\Omega_\delta}\,.$$

Nach Satz 4.5 gilt $\|D^\alpha u - (D^\alpha u)_\varepsilon\|_{p,\Omega} \to 0$ für $\varepsilon \to 0$, also resultiert für $0 < \varepsilon < \delta$

$$\begin{aligned} \|u - u_\varepsilon\|_{p,m,\Omega_\delta} &= \sum_{|\alpha|\le m} \|D^\alpha u - D^\alpha u_\varepsilon\|_{p,\Omega_\delta} \\ &= \sum_{|\alpha|\le m} \|D^\alpha u - (D^\alpha u)_\varepsilon\|_{p,\Omega_\delta} \to 0 \end{aligned}$$

für $\varepsilon \to 0$. Also kann $u \in H_m^p(\Omega)$ in jeder offenen Menge $D \subseteq \Omega$ mit $\operatorname{dist}(\mathrm{D}, \partial\Omega) > 0$ durch C_∞–Funktionen in der H_m^p–Norm approximiert werden. Als Folgerung erhalten wir hieraus:

3.8.1 Folgerung: Sei $\Omega \subseteq \mathbb{R}^n$ eine offene und beschränkte Menge, $1 \leq p < \infty$, $u \in H_m^p(\Omega)$. Falls $u(x) = 0$ ist für alle $x \in \Omega$ mit $\mathrm{dist}(\mathrm{x}, \partial\Omega) < \delta$, dann existiert eine Folge $\{u_k\}_{k=1}^{\infty} \subseteq \overset{\circ}{C}_{\infty}(\Omega)$ mit

$$\|u - u_k\|_{p,m,\Omega} \to 0\,, \quad k \to \infty\,.$$

Beweis: f kann durch Null zu einer Funktion in $H_m^p(\mathbb{R}^n)$ fortgesetzt werden, also gilt

$$D^\alpha u_{\varepsilon}{}_{|\Omega} = [(D^\alpha u)_\varepsilon]_{|\Omega}\,,$$

und $\|u - u_\varepsilon\|_{p,m,\Omega} \to 0$ für $\varepsilon \to 0$.
Es gilt $\mathrm{dist}(\mathrm{supp\,u}\,,\, \mathrm{R^n} \backslash \Omega) = \delta > 0$, also verschwindet

$$u_\varepsilon(x) = \int_{\mathbb{R}^n} \varphi_\varepsilon(x - y)\, u(y)\, dy$$

für alle x mit $\mathrm{dist}(\mathrm{x}, \mathrm{supp\,u}) > \frac{\delta}{2}$ falls $\varepsilon < \frac{\delta}{2}$ ist, also ist $\mathrm{supp\,u}_\varepsilon$ beschränkt, und damit kompakt, und in Ω enthalten. Hieraus folgt die Behauptung.

Die Approximation von beliebigem $u \in H_m^p(\Omega)$ auf Ω ist schwieriger und erfordert einige Vorbereitungen, denen wir uns jetzt zuwenden.

3.9 Überdeckungen. Sei $\Omega \subseteq \mathbb{R}^n$. Eine Familie $\{U_i\}_{i=1}^{\infty}$ von Teilmengen U_i von $\mathbb{R}^n$ heißt Überdeckung von Ω, wenn

$$\Omega \subseteq \bigcup_{i=1}^{\infty} U_i\,.$$

Wenn jedes U_i offen ist heißt $\{U_i\}_{i=1}^{\infty}$ offene Überdeckung von Ω. Gibt es zu jedem $x \in \Omega$ eine Umgebung V von x in $\mathbb{R}^n$ mit $U_i \cap \overline{V} \neq \emptyset$ für höchstens endlich viele $i \in \mathbb{N}$, dann heißt $\{U_i\}_{i=1}^{\infty}$ lokalendliche Überdeckung von Ω.

Sei $\Omega \subseteq \mathbb{R}^n$ offen. Wir benötigen im folgenden eine offene, lokalendliche Überdeckung $\{U_i\}_{i=1}^{\infty}$ von Ω mit U_i beschränkt und mit $\overline{U_i} \subseteq \Omega$ für jedes $i \in \mathbb{N}$.

Eine solche Überdeckung kann man folgendermaßen konstruieren:
Sei $\Omega_0 = \emptyset$ und für $i \in \mathbb{N}$

$$\Omega_i = \{x \in \Omega \;\Big|\; |x| < i\,,\ \mathrm{dist}(\mathrm{x}, \partial\Omega) > \frac{1}{\mathrm{i}}\}\,.$$

Dann ist Ω_i offen und beschränkt,

$$\overline{\Omega_i} \subseteq \Omega_{i+1} \subseteq \overline{\Omega_{i+1}} \subseteq \Omega \,, \quad \bigcup_{i=1}^{\infty} \Omega_i = \Omega \,.$$

Setze $U_i = \Omega_{i+1} \backslash \overline{\Omega_{i-1}}$. Dann ist U_i offen und beschränkt,

$$\overline{U_i} \subseteq \overline{\Omega_{i+1}} \subseteq \Omega \,,$$

$$\Omega \supseteq \bigcup_{i=1}^{\infty} U_i = \bigcup_{i=1}^{\infty} (\Omega_{i+1} \backslash \overline{\Omega_{i-1}}) \supseteq \bigcup_{i=2}^{\infty} (\Omega_{i+1} \backslash \Omega_i) \cup \Omega_2 = \bigcup_{i=1}^{\infty} \Omega_i = \Omega \,.$$

Schließlich ist $\{U_i\}_{i=1}^{\infty}$ lokalendlich, weil für $m \geq i+2$ gilt

$$\begin{aligned} U_m \cap U_i &= (\Omega_{m+1} \backslash \overline{\Omega_{m-1}}) \cap (\Omega_{i+1} \backslash \overline{\Omega_{i-1}}) \\ &\subseteq (\Omega_{m+1} \backslash \Omega_{m-1}) \cap \Omega_{m-1} = \emptyset \,. \end{aligned}$$

3.10 Zerlegung der Eins. Sei $\Omega \subseteq \mathbb{R}^n$ offen und $\{U_i\}_{i=1}^{\infty}$ eine lokalendliche, offene Überdeckung von Ω . Eine Folge $\{\eta_i\}_{i=1}^{\infty} \subseteq \overset{\circ}{C}_{\infty}(\mathbb{R}^n)$ heißt eine der Überdeckung $\{U_i\}_{i=1}^{\infty}$ untergeordnete Zerlegung der Eins, wenn gilt

$$\eta_i \in \overset{\circ}{C}_{\infty}(U_i) \,, \ \eta_i \geq 0 \,, \ \sum_{i=1}^{\infty} \eta_i(x) = 1 \quad \text{für alle } x \in \Omega \,.$$

Beachte, daß zu jedem $x \in \mathbb{R}^n$ in dieser Summe höchstens endlich viele Summanden von Null verschieden sind.

3.11 Lemma. Sei $\Omega \subseteq \mathbb{R}^n$ offen und $\{U_i\}_{i=1}^{\infty}$ eine offene, lokalendliche Überdeckung von Ω . Ist jedes der U_i beschränkt mit $\overline{U_i} \subseteq \Omega$, dann existiert eine dieser Überdeckung untergeordnete Zerlegung der Eins $\{\eta_i\}_{i=1}^{\infty}$.

Beweis: Durch Induktion können wir eine Folge offener Mengen $\{V_i\}_{i=1}^{\infty}$ mit $\overline{V_i} \subseteq U_i$ und mit $\bigcup_{i=1}^{\infty} V_i = \Omega$ finden. (Also ist auch $\{V_i\}_{i=1}^{\infty}$ eine offene, lokalendliche Überdeckung von Ω aus beschränkten Mengen.)

Dann seien $V_1, \ldots, V_m$ so gewählt, daß

$$\Omega = (\bigcup_{i=1}^{m} V_i) \cup (\bigcup_{i=m+1}^{\infty} U_i)$$

gilt. Für die abgeschlossene und beschränkte, also kompakte Menge ∂U_{m+1} gilt dann

$$\partial U_{m+1} \cap U_{m+1} = \emptyset \quad \text{und}$$
$$\partial U_{m+1} \subseteq \overline{U_{m+1}} \subseteq \Omega = (\bigcup_{i=1}^{m} V_i) \cup (\bigcup_{i=m+1}^{\infty} U_i)\,,$$

also sogar

$$\partial U_{m+1} \subseteq (\bigcup_{i=1}^{m} V_i) \cup (\bigcup_{i=m+2}^{\infty} U_i) = M$$

d.h.

$$\mathrm{dist}(\Omega \backslash \mathrm{M}\,,\, \partial \mathrm{U}_{\mathrm{m+1}}) \geq \mathrm{dist}(\mathbb{R}^{\mathrm{n}} \backslash \mathrm{M}\,,\, \partial \mathrm{U}_{\mathrm{m+1}}) = \delta > 0\,.$$

Setze $V_{m+1} = \{x \in U_{m+1} \mid \mathrm{dist}(\mathrm{x}, \partial \mathrm{U}_{\mathrm{m+1}}) > \frac{\delta}{2}\}$. Dann folgt

$$\Omega = (\bigcup_{i=1}^{m+1} V_i) \cup (\bigcup_{i=m+2}^{\infty} U_i)\,.$$

Für die so konstruierte Folge $\{V_i\}_{i=1}^{\infty}$ gilt nun

$$\Omega = \bigcup_{i=1}^{\infty} V_i\,.$$

Denn für $x \in \Omega$ gibt es höchstens endlich viele $i_1, \ldots, i_k$ mit $x \in U_{i_l}$, $l = 1, \ldots, k$. Also folgt

$$x \notin \bigcup_{j=i_k+1}^{\infty} U_j\,, \quad \text{somit} \quad x \in \bigcup_{j=1}^{i_k} V_j \subseteq \bigcup_{j=1}^{\infty} V_j\,,$$

d.h.

$$\Omega = \bigcup_{j=1}^{\infty} V_j\,.$$

Wegen $\overline{V_i} \subseteq U_i$ ist $\mathrm{dist(V_i\,,\,\partial U_i) = \delta_i > 0}$. Deswegen kann man eine offene Zwischenmenge E_i wählen (etwa $E_i = \bigcup\limits_{x \in V_i} B_{\delta_i/2}(x)$) mit

$$\overline{V_i} \subseteq E_i \subseteq \overline{E_i} \subseteq U_i \,,$$

also

$$\mathrm{dist(\partial E_i\,,\,V_i \cup (\mathbb{R}^n \backslash U_i)) = \varepsilon_i > 0}\,.$$

Für eine Dirac-Folge $\{\varphi_\varepsilon\}_{\varepsilon>0}$ zu $\varphi \in \overset{\circ}{C}_\infty(\overline{B_1(0)}\,)$ setze dann

$$\widetilde{\eta}_i(x) = (\varphi_{\varepsilon_i} * \chi_{E_i})(x) = \int_{E_i} \varphi_{\varepsilon_i}(x-y)dy \,.$$

Wegen $\varphi_{\varepsilon_i} \in \overset{\circ}{C}_\infty(B_{\varepsilon_i}(0))\,,\ \displaystyle\int_{|z|<\varepsilon_i} \varphi_{\varepsilon_i}(z)dz = 1$ ist dann $\widetilde{\eta}_i \in \overset{\circ}{C}_\infty(U_i)$ mit

$$\chi_{V_i} \leq \widetilde{\eta}_i \leq 1 \,,$$

also

$$\sum_{i=1}^{\infty} \widetilde{\eta}_i \geq \sum_{i=1}^{\infty} \chi_{V_i} \geq \chi_\Omega \,,$$

wobei jeweils höchstens endlich viele Summanden von Null verschieden sind. Setze nun für $x \in \Omega$

$$\eta_i(x) = \frac{\widetilde{\eta}_i(x)}{\sum_{j=1}^{\infty} \widetilde{\eta}_j(x)} \in \overset{\circ}{C}_\infty(U_i) \,.$$

Damit ist Lemma 3.11 bewiesen.

3.12 Lemma (Leibnizsche Regel für schwache Ableitungen). Sei $f \in H^p_m(\Omega)\,,\ 1 \leq p \leq \infty$ und $\eta \in \overset{\circ}{C}_\infty(\Omega)$. Dann ist $\eta f \in H^p_m(\Omega)$, und die schwachen Ableitungen erhält man aus der Leibnizschen Regel:

$$D^\alpha(\eta f) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D^\beta \eta \, D^{\alpha-\beta} f \,,$$

wobei $\beta \leq \alpha \Leftrightarrow \beta_i \leq \alpha_i$ und $\displaystyle\binom{\alpha}{\beta} = \binom{\alpha_1}{\beta_1} \cdots \binom{\alpha_n}{\beta_n}$.

Beweis: (Durch Induktion) Für $\varphi \in \overset{\circ}{C}_\infty(\Omega)$ gilt

$$\begin{aligned}\int_\Omega \eta f \frac{\partial}{\partial x_i}\varphi \, dx &= \int_\Omega f \frac{\partial}{\partial x_i}(\eta\varphi) dx - \int_\Omega f (\frac{\partial}{\partial x_i}\eta)\varphi \, dx \\ &= -\int_\Omega \left[(\frac{\partial}{\partial x_i} f)\eta + f \frac{\partial}{\partial_{x_i}}\eta \right] \varphi \, dx \, .\end{aligned}$$

Also hat ηf die schwache partielle Ableitung

$$\frac{\partial}{\partial x_i} f \, \eta + f \frac{\partial}{\partial x_i} \eta \in L^p(\Omega) \, .$$

Für die zweiten Ableitungen ergibt sich die Aussage durch Anwendung dieser Schlüsse auf $(\frac{\partial}{\partial x_i} f)\eta$ und $f \frac{\partial}{\partial x_i} \eta$, und entsprechend geht man bei den höheren Ableitungen vor.

3.13 Beweis von Satz 2.14. Sei $1 \le p < \infty$, $\Omega \subseteq \mathbb{R}^n$ sei eine offene Menge, $f \in H_m^p(\Omega)$ und sei $\varepsilon > 0$. Zum Beweis des Satzes muß eine Funktion $g \in C_\infty^p(\Omega)$ konstruiert werden mit

$$\|f - g\|_{p,m,\Omega} < \varepsilon \, .$$

Sei $\{U_i\}_{i=1}^\infty$ eine offene, lokalendliche Überdeckung von Ω, und sei $\{\eta_i\}_{i=1}^\infty$ eine dieser Überdeckung untergeordnete Zerlegung der Eins auf Ω. Nach 3.9 und Lemma 3.11 gibt es solche Überdeckungen und eine Zerlegung der Eins.

Da $\eta_i \in \overset{\circ}{C}_\infty(U_i)$ ist, ist nach Lemma 3.12 $\eta_i f \in H_m^p(U_i)$ mit $\operatorname{supp}(\eta_i f) \subseteq \operatorname{supp} \eta_i \subset\subset U_i$. Nach Folgerung 3.8.1 gibt es eine Funktion $f_i \in \overset{\circ}{C}_\infty(U_i)$ mit

$$\|\eta_i f - f_i\|_{p,m,\Omega} < \frac{\varepsilon}{2^i} \, .$$

Setze $g(x) = \sum_{i=1}^\infty f_i(x)$, wobei für jedes $x \in \Omega$ nur eine endliche Anzahl der Summanden von Null verschieden ist. Also gilt $g \in C_\infty(\Omega)$, und

$$\|f - g\|_{p,m,\Omega} = \| \sum_{i=1}^\infty \eta_i f - \sum_{i=1}^\infty f_i \|_{p,m,\Omega}$$

$$
\begin{aligned}
&= \|\sum_{i=1}^{\infty}(\eta_i f - f_i)\|_{p,m,\Omega} \\
&\leq \sum_{i=1}^{\infty}\|\eta_i f - f_i\|_{p,m,\Omega} \\
&< \sum_{i=1}^{\infty}\frac{\varepsilon}{2^i} = \varepsilon \,.
\end{aligned}
$$

Als zunächst letztes Ergebnis aus der Sobolevraumtheorie zeige ich:

3.14 Satz. (Poincarésche Ungleichung). Sei $1 \leq p < \infty$ und sei $\Omega \subseteq \mathbb{R}^n$ eine offene, beschränkte Menge. Sei

$$d = \dim\Omega = \sup_{x,y\in\Omega} |x-y| \,.$$

Dann gilt für alle Funktionen $u \in \overset{\circ}{H}{}^p_1(\Omega)$

$$\|u\|_{p,\Omega} \leq p^{-1/p} d\, |u|_{p,1,\Omega}$$

mit

$$|u|_{p,1,\Omega} = \sum_{|\alpha|=1} \|D^\alpha u\|_{p,\Omega} \,.$$

Beweis: Sei $\varphi \in \overset{\circ}{C}_\infty(\Omega)$, sei $x = (x_1, \dots, x_n) \in \Omega$, sei $y = (y_1, x_2, \dots, x_n) \in \partial\Omega$. Dann folgt

$$\varphi(x) = \int_{y_1}^{x_1} \frac{\partial}{\partial x_1}\varphi(\xi, x_2, \dots, x_n) d\xi \,,$$

also, wegen der Hölderschen Ungleichung

$$
\begin{aligned}
|\varphi(x)|^p &\leq \Big(\int_{y_1}^{x_1} \Big|\frac{\partial}{\partial x_1}\varphi(\xi, x_2, \dots, x_n)\Big|\, d\xi\Big)^p \\
&\leq |x_1 - y_1|^{\frac{p}{q}} \int_{y_1}^{x_1} \Big|\frac{\partial}{\partial x_1}\varphi\Big|^p\, d\xi \\
&\leq |x_1 - y_1|^{\frac{p}{q}} \int_{-\infty}^{\infty} \sum_{|\alpha|=1} |D^\alpha\varphi(\xi, x_2, \dots, x_n)|^p\, d\xi \,.
\end{aligned}
$$

Integration bezüglich x_1 ergibt

$$\int_{y_1}^{y_1+d} |\varphi(x)|^p \, dx_1 \le \frac{1}{1+\frac{p}{q}} \, d^{1+\frac{p}{q}} \int_{-\infty}^{\infty} \sum_{|\alpha|=1} |D^\alpha \varphi(\xi, x_2, \ldots, x_n)|^p \, d\xi \, .$$

Nach Integration bezüglich der anderen Variablen ergibt sich

$$\int_\Omega |\varphi(x)|^p dx \le \frac{1}{p} \, d^p \int_\Omega \sum_{|\alpha|=1} |D^\alpha \varphi(x)|^p \, dx$$

und somit

$$\|\varphi\|_{p,\Omega} \le p^{-1/p} \, d \, |\varphi|_{p,1,\Omega} \, .$$

Also ist die behauptete Ungleichung richtig für $\varphi \in \overset{\circ}{C}_\infty(\Omega)$. Nun sei $u \in \overset{\circ}{H}{}_1^p(\Omega)$. Wähle eine Folge $\{\varphi_k\}_{k=1}^\infty \subseteq \overset{\circ}{H}{}_1^p(\Omega)$ mit

$$\lim_{k\to\infty} \|u - \varphi_k\|_{p,1,\Omega} = 0 \, .$$

Dann folgt

$$\begin{aligned}
\|u\|_{p,\Omega} &\le \lim_{k\to\infty} (\|u - \varphi_k\|_{p,\Omega} + \|\varphi_k\|_{p,\Omega}) \\
&\le \lim_{k\to\infty} \|\varphi_k\|_{p,\Omega} \le p^{-1/p} \, d \lim_{k\to\infty} |\varphi|_{p,1,\Omega} \\
&\le p^{-1/p} \, d \lim_{k\to\infty} (|\varphi - u|_{p,1,\Omega} + |u|_{p,1,\Omega}) \\
&= p^{-1/p} \, d \, |u|_{p,1,\Omega} \, .
\end{aligned}$$

4. Klassische Methoden. Eulergleichung und Beispiele.

In diesem Abschnitt soll eine Einführung in die klassischen Methoden der Variationsrechnung gegeben werden. Zur Vorbereitung benötige ich ein Resultat für konvexe Funktionen.

4.1 Definition. (Konvexe Mengen und Funktionen). Sei X ein Vektorraum über $\mathbb{K}$.

(i) Eine Menge $M \subseteq X$ heißt konvex, wenn für alle $x, y \in M$ und alle $t \in [0,1]$ gilt

$$tx + (1-t)y \in M\,.$$

(ii) Sei $M \subseteq X$ eine konvexe Menge und $f : M \to (-\infty, \infty]$. Die Funktion f heißt konvex, wenn für alle $x, y, \in M$ und alle $t \in [0,1]$ gilt

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)\,.$$

(iii) f heißt strikt konvex, wenn für alle $x, y \in M$ mit $x \neq y$ und alle $t \in (0,1)$ gilt

$$f(tx + (1-t)y) < tf(x) + (1-t)f(y)\,.$$

4.2 Lemma (Eigenschaften konvexer Abbildungen). Sei $f : \mathbb{R}^n \to \mathbb{R}$.

(i) Falls $f \in C_1(\mathbb{R}^n, \mathbb{R})$ ist, gilt: f ist konvex genau dann wenn die Ungleichung

$$f(x) \geq f(y) + \nabla f(y) \cdot (x-y)$$

gültig ist für alle $x, y \in \mathbb{R}^n$.

(ii) Falls $f \in C_2(\mathbb{R}^n, \mathbb{R})$ ist, gilt: f ist konvex genau dann wenn die Hessesche Matrix

$$\nabla^2 f(x) = \left(D^\alpha f(x)\right)_{|\alpha|=2}$$

positiv semidefinit ist für alle $x \in \mathbb{R}^n$.

Beweis: (i) Sei f konvex. Aus der Differenzierbarkeit von f folgt für alle $x, y \in \mathbb{R}^n$

$$f(y + t(x - y)) = f(y) + \nabla f(y) \cdot t(x - y) + o(t)$$

für $t \to 0$, also

$$\begin{aligned}
\nabla f(y) \cdot (x - y) &= \lim_{\substack{t \to 0 \\ t > 0}} \nabla f(y) \cdot (x - y) \\
&= \lim_{\substack{t \to 0 \\ t > 0}} \frac{1}{t} \Big[f(y + t(x - y)) - f(y) \Big] \\
&= \lim_{\substack{t \to 0 \\ t > 0}} \frac{1}{t} \Big[f(tx + (1 - t)y) - f(y) \Big] \\
&\leq \lim_{\substack{t \to 0 \\ t > 0}} \frac{1}{t} [tf(x) - tf(y)] \\
&= f(x) - f(y) \,,
\end{aligned}$$

also

$$f(x) \geq f(y) + \nabla f(y) \cdot (x - y) \,.$$

Umgekehrt sei diese Ungleichung für alle $x, y \in \mathbb{R}^n$ gültig. Für $x, y \in \mathbb{R}^n$ und $t \in [0, 1]$ setze

$$z = tx + (1 - t)y \,.$$

Dann folgt

$$\begin{aligned}
f(y) &\geq \nabla f(z)(y - z) \quad + f(z) \\
f(x) &\geq \nabla f(z)(x - z) \quad + f(z) \,,
\end{aligned}$$

also

$$\begin{aligned}
tf(x) + (1 - t)f(y) &\geq t\nabla f(z)(x - z) + tf(z) \\
&\quad + (1 - t)\nabla f(z)(y - z) + (1 - t)f(z) \\
&= f(z) + \nabla f(z) \cdot (t(x - z) + (1 - t)(y - z)) \\
&= f(z) + \nabla f(z) \cdot \underbrace{(tx + (1 - t)y - z)}_{=0} \\
&= f(z) \\
&= f(tx + (1 - t)y)
\end{aligned}$$

Also ist f konvex.

(ii) Wenn $f \in C_2(\mathbb{R}^n, \mathbb{R})$ ist, gilt für $x, y \in \mathbb{R}^n$ und $t \in \mathbb{R}$

$$f(tx + y) = f(y) + \nabla f(y) \cdot tx + tx[\nabla^2 f(y)]tx + o(t^2)\,,$$

falls $t \to 0$. Falls f konvex ist, folgt hieraus und aus der eben bewiesenen Aussage (i), daß

$$\begin{aligned} x[\nabla^2 f(y)]x &= \lim_{t\to 0} x[\nabla^2 f(y)]x \\ &= \lim_{t\to 0} \frac{1}{t^2}[f(tx+y) - f(y) - \nabla f(y) \cdot tx - o(t^2)] \\ &= \lim_{t\to 0} \frac{1}{t^2}[f(tx+y) - f(y) - \nabla f(y) \cdot tx] \\ &\geq 0\,. \end{aligned}$$

Also ist $\nabla^2 f(y)$ positiv semidefinit. Wenn umgekehrt $\nabla^2 f(y)$ positiv semidefinit ist, folgt für alle $x, y \in \mathbb{R}^n$ aus der Taylorformel

$$f(x) = f(y) + \nabla f(y) \cdot (x - y) + (x - y)[\nabla^2 f(y^*)](x - y)$$

für ein geeignetes y^* aus der Verbindungsstrecke von x und y. Da $(x - y)[\nabla^2 f(y^*)](x - y) \geq 0$ ist, folgt

$$f(x) \geq f(y) + \nabla f(y) \cdot (x - y)$$

für alle $x, y \in \mathbb{R}^n$. Aus der Aussage (i) ergibt sich also, daß f konvex ist.

Ich komme nun zur Ableitung der Eulergleichung für eindimensionale Variationsprobleme. Mit der Bezeichnung des Abschnitts 1.1 gilt also $n = N = 1$.

4.3 Satz (Eulergleichung). Seien $a < b$, $f \in C_2([a, b] \times \mathbb{R} \times \mathbb{R})$, $\alpha, \beta \in \mathbb{R}$,

$$D = \Big\{ v \in C_1[[a, b], \mathbb{R}) : v(a) = \alpha, v(b) = \beta \Big\}\,,$$

und $I : D \to \mathbb{R}$ mit

$$I(v) = \int_a^b f(x, v(x), v'(x))dx\,.$$

(i) Es existiere ein Minimum u von I in D, und für dieses Minimum gelte zusätzlich $u \in C_2([a,b],\mathbb{R})$. Dann gilt für dieses Minimum

$$\begin{aligned} \tfrac{d}{dx}\Big[f_\xi(x,u(x),u'(x))\Big] &= f_u(x,u(x),u'(x))\,, \quad x \in [a,b] \\ u(a) = \alpha\,, &\quad u(b) = \beta \end{aligned} \tag{E}$$

wobei $f_\xi(x,u,\xi) = \frac{\partial}{\partial \xi} f(x,u,\xi)$, $f_u = \frac{\partial}{\partial u} f(x,u,\xi)$ bedeute.

(ii) Sei umgekehrt $u \in C_2([a,b])$ eine Lösung des Randwertproblems (E) und sei

$$(u,\xi) \mapsto f(x,u,\xi) : \mathbb{R}^2 \to \mathbb{R}$$

eine konvexe Funktion für jedes $x \in [a,b]$. Dann gilt

$$I(u) = \min_{v \in D} I(v)\,.$$

(iii) Sei $(u,\xi) \mapsto f(x,u,\xi)$ strikt konvex für jedes $x \in [a,b]$. Dann besitzt I höchstens ein lokales Minimum. Wenn ein lokales Minimum existiert, ist dies gleichzeitig ein globales Minimum (das dann natürlich auch eindeutig ist).

Die Differentialgleichung im Randwertproblem heißt **Eulergleichung** zum Variationsfunktional $I(v)$. Die Eulergleichung ist eine Differentialgleichung zweiter Ordnung.

Beweis: (i) Es sei $u \in D \cap C_2([a,b],\mathbb{R})$ ein Minimum des Variationsfunktionals $I(v)$. Dann ist für jedes $\varphi \in \mathring{C}_\infty((a,b),\mathbb{R})$ und $t \in \mathbb{R}$ die Funktion $u + t\varphi \in D$ wegen $(u+t\varphi)(a) = u(a) + 0 = \alpha$, $(u+t\varphi)(b) = u(b) + 0 = \beta$, sowie

$$I(u) \leq I(u + t\varphi)\,.$$

Also hat die reelle Funktion $t \mapsto I(u+t\varphi) : \mathbb{R} \to \mathbb{R}$ ein Minimum an der Stelle $t = 0$, und somit folgt

$$
\begin{aligned}
0 &= \frac{d}{dt} I(u+t\varphi)_{|t=0} \\
&= \frac{d}{dt} \int_a^b f(x, u(x)+t\varphi(x), u'(x)+t\varphi'(x))dx \\
&= \int_a^b \frac{d}{dt} f(x, u(x)+t\varphi(x), u'(x)+t\varphi'(x))_{|t=0} dx \\
&= \int_a^b f_\xi(x, u(x), u'(x))\varphi'(x) + f_u(x, u(x), u'(x))\varphi(x)dx\,.
\end{aligned}
$$

Diese Gleichung, die für alle $\varphi \in \overset{\circ}{C}_\infty((a,b))$ gilt, heißt **schwache Form der Eulerschen Differentialgleichung**. Zur Herleitung dieser Gleichung benötigt man nur $f \in C_1([a,b] \times \mathbb{R} \times \mathbb{R})$ und $u \in C_1([a,b])$.

Zur Ableitung der Eulergleichung benutzt man nun, daß wegen der Voraussetzungen $f \in C_2([a,b] \times \mathbb{R} \times \mathbb{R})$ und $u \in C_2([a,b])$ partiell integriert werden darf. Es folgt

$$
\begin{aligned}
0 &= \int_a^b \Big[-\frac{d}{dx}\Big[f_\xi(x,u(x),u'(x))\Big]\varphi(x) + f_u(x,u(x),u'(x))\varphi(x)\Big] dx \\
&\quad + f_\xi(b,u(b),u'(b))\varphi(b) - f_\xi(a,u(a),u'(a))\varphi(a) \\
&= \int_a^b \Big[-\frac{d}{dx} f_\xi(x,u(x),u'(x)) + f_u(x,u(x),u'(x))\Big]\varphi(x)dx\,,
\end{aligned}
$$

weil $\varphi(a) = \varphi(b) = 0$. Wegen

$$
-\frac{d}{dx} f_\xi(x,u(x),u'(x)) + f_u(x,u(x),u'(x)) \in C_0([a,b]) \subseteq L^{1,loc}((a,b))
$$

folgt aus dem Fundamentallemma der Variationsrechnung (Lemma 3.7), daß

$$
\frac{d}{dx} f_\xi(x,u(x),u'(x)) - f_u(x,u(x),u'(x)) = 0
$$

gilt für alle $x \in [a,b]$. Also ist das Minimum u eine Lösung des Randwertproblems (E).

(ii) Sei $u \in C_2([a,b])$ eine Lösung von (E) . Wenn $(\mu,\xi) \mapsto f(x,\mu,\xi)$ konvex ist, gilt nach Lemma 4.2 (i) für alle $\mu, \xi \in \mathbb{R}$

$$
\begin{aligned}
f(x,\mu,\xi) \quad \geq \quad & f(x,u(x),u'(x)) \\
& + f_u(x,u(x),u'(x))(\mu - u(x)) \\
& + f_\xi(x,u(x),u'(x))(\xi - u'(x))\,.
\end{aligned}
$$

Also folgt für alle $v \in D$

$$
\begin{aligned}
I(v) \quad &= \quad \int_a^b f(x,v(x)v'(x))dx \\
&\geq \quad \int_a^b f(x,u(x),u'(x))dx \\
&\qquad + \int_a^b \Big[f_u(x,u(x),u'(x))(v(x)-u(x)) \\
&\qquad + f_\xi(x,u(x),u'(x))(v'(x)-u'(x)) \Big] dx \\
&= \quad I(u) + \int_a^b \Big(f_u(x,u(x)u'(x)) \\
&\qquad - \frac{d}{dx}\Big[f_\xi(x,u(x),u'(x)) \Big] \Big) (v(x)-u(x))dx \\
&\qquad + f_\xi(b,u(b),u'(b))(v(b)-u(b)) \\
&\qquad - f_\xi(a,u(a),u'(a))(v(a)-u(a)) \\
&= \quad I(u)\,,
\end{aligned}
$$

wegen $v(b) = u(b) = \beta$, $v(a) = u(a) = \alpha$. Also ist u ein Minimum des Variationsfunktionals.

(iii) Seien $u \in D$ eine lokale Minimalstelle von I, wobei ich als Topologie auf $C_1([a,b])$ die Norm

$$
\|w\|_{\infty,1} = \sup_{x\in[a,b]} |w(x)| + \sup_{x\in[a,b]} |w'(x)|
$$

benutze. Dann existiert eine Umgebung U von u in $C_1([a,b])$ mit $I(w) \geq I(u)$ für alle $w \in D \cap U$. Ich zeige, daß für alle $v \in D$ mit $v \neq u$ gilt

$$
I(v) > I(u)\,.
$$

Hieraus folgt, daß u ein globales Minimum ist, und daß es nur ein einziges Minimum gibt.
Zum Beweis setze ich für $v \in D$ mit $v \neq u$ und für $t \in (0,1)$

$$w_t = tv + (1-t)u\,.$$

Aus der strikten Konvexität von $(\mu, \xi) \mapsto f(x, \mu, \xi)$ folgt dann

$$\begin{aligned} f(x, w_t, (x), w_t'(x)) &= f(x, tv(x) + (1-t)u(x), tv'(x) + (1-t)u'(x)) \\ &\leq tf(x, v(x), v'(x)) + (1-t)f(x, u(x), u'(x))\,, \end{aligned}$$

wobei das echte Ungleichheitszeichen steht falls
$(u(x), u'(x)) \neq (v(x), v'(x))$ ist. Da nach Voraussetzung $u \neq v$ gilt, gibt es eine offene Teilmenge von $[a, b]$, auf der sich $u(x)$ und $v(x)$ unterscheiden. Also folgt durch Integration

$$\begin{aligned} I(w_t) &= \int_a^b f(x, w_t(x), w_t'(x))dx \\ &< t\int_a^b f(x, v(x), v'(x))dx + (1-t)\int_a^b f(x, u(x), u'(x))dx \\ &= tI(v) + (1-t)I(u)\,. \end{aligned}$$

Wäre nun $I(v) \leq I(u)$, dann folgte für alle $t \in (0,1)$

$$\begin{aligned} I(w_t) &< tI(v) + (1-t)I(u) \\ &\leq tI(u) + (1-t)I(u) \\ &= I(u)\,. \end{aligned}$$

Wegen $w_t \to u$ für $t \to 0$ gibt es aber ein $t_0 > 0$ mit $w_t \in U$ für alle $0 < t < t_0$. Für diese t muß somit $I(w_t) \geq I(u)$ gelten, im Widerspruch zur eben hergeleiteten Ungleichung. Also muß $I(v) > I(u)$ gelten.

4.4 Beispiel (Der Fall $\boldsymbol{f(x, u, \xi) = f(\xi)}$). In diesem Fall lautet die Eulergleichung

$$\frac{d}{dx}[f'(u')] = 0\,,$$

also $f'(u'(x)) = \text{ const}$. Eine Lösung des Randwertproblems (E) ist also

$$u(x) = \frac{\beta - \alpha}{b - a}(x - a) + \alpha\,,$$

und wenn f konvex ist, folgt aus Satz 4.3 (ii), daß dieses u auch eine Lösung des Variationsproblems

$$I(u) = \min_{v \in D} I(v)$$

ist mit

$$I(v) = \int_a^b f(v'(x))dx\,.$$

Wenn f nicht konvex ist, braucht dieses u aber keine Lösung des Variationsproblems zu sein. Dies zeigen die folgenden beiden Beispiele.

a) Es sei

$$f(\xi) = e^{-\xi^2}$$

und

$$D = \Big\{ v \in C_1([-1,1]) : v(-1) = v(1) = 0 \Big\}\,.$$

Ich zeige zunächst, daß

$$\inf_{v \in D} I(v) = 0$$

gilt für $I(v) = \int_{-1}^1 f(v'(x))dx = \int_{-1}^1 e^{-(v'(x))^2}dx$. Hierzu sei $m \in \mathbb{N}$ und

$$v_m(x) = m(x^2 - 1)\,.$$

Dann gilt $v_m'(x) = 2mx$ und

$$\begin{aligned} I(v_m) &= \int_{-1}^1 e^{(2mx)^2}dx \\ &= \frac{1}{2m}\int_{-2m}^{2m} e^{-y^2}dy \\ &\leq \frac{1}{2m}\int_{-\infty}^{\infty} e^{-y^2}dy \to 0 \end{aligned}$$

für $m \to \infty$. Da $I(v) \geq 0$ ist für alle $v \in D$, folgt also $\inf_{v \in D} I(v) = 0$. I hat aber kein Minimum u auf D, denn sonst müßte $I(u) = 0$ und folglich $e^{-(u'(x))^2} = 0$ sein für fast alle $x \in [-1,1]$, was unmöglich ist. Die Lösung u von (E) ist also nicht Lösung des Variationsproblems.

b) Sei $f(\xi) = (\xi^2 - 1)^2$ und

$$D = \Big\{ v \in C_1([-1,1]) : v(-1) = v(1) = 0 \Big\}\,.$$

Die Eulergleichung zum Variationsfunktional

$$I(v) = \int_{-1}^{1} f(v'(x))dx = \int_{-1}^{1} (v'(x)^2 - 1)^2 dx$$

ist dann

$$\frac{d}{dx} f'(u'(x)) = 4\frac{d}{dx}[u'(u'^2 - 1)] = 0\,,$$

und $u = 0$ ist eine Lösung des Randwertproblems (E). Ich zeige aber, daß auch in diesem Fall I kein Minimum auf D besitzt. Hierzu beweise ich, daß

$$\inf_{v \in D} I(v) = 0$$

ist. Für das Minimum müßte dann gelten

$$I(u) = \int_{-1}^{1} (u'(x)^2 - 1)^2 dx = 0\,,$$

also

$$u'(x) = \pm 1$$

für alle $x \in [-1, 1]$. Wenn u stetig differenzierbar wäre, würde hieraus $u'(x) =$ const mit $|\text{const}| = 1$ folgen, was der Bedingung $u(-1) = u(1) = 0$ widerspräche.
Um zu zeigen, daß $\inf_{v \in D} I(v) = 0$ gilt, sei

$$\psi(x) = \begin{cases} e^{-\frac{1}{1-x^2}} & , \quad |x| < 1 \\ 0 & , \quad |x| \geq 1 \end{cases}$$

und

$$\varphi_\varepsilon(x) = \frac{1}{\|\varphi\|_1}\, \frac{1}{\varepsilon} \varphi(\frac{x}{\varepsilon})\,.$$

$\{\varphi_\varepsilon\}_{\varepsilon > 0}$ ist also die Dirac–Familie aus 3.6. Sei

$$H(x) = \begin{cases} 1, & x \geq 0 \\ -1, & x < 0 \end{cases}\,,$$

und $H_\varepsilon(x) = (\varphi_\varepsilon * h)(x) = \int_{-\infty}^{\infty} \varphi_\varepsilon(x-y)H(y)dy$. Für diese Funktion gilt $H_\varepsilon \in C_\infty(\mathbb{R})$, und aus $\varphi_\varepsilon(-z) = \varphi_\varepsilon(z)$ und $H(-z) = -H(z)$ folgt

$$\begin{aligned}
H_\varepsilon(x) &= \int_{-\infty}^{\infty} \varphi_\varepsilon(x-y)H(y)dy \\
&= \int_{-\infty}^{\infty} \varphi_\varepsilon(y-x)H(y)dy \\
&= -\int_{\infty}^{-\infty} \varphi_\varepsilon(-z-x)H(-z)dz \\
&= \int_{-\infty}^{\infty} \varphi_\varepsilon(-x-z)H(-z)dz \\
&= -\int_{-\infty}^{\infty} \varphi_\varepsilon(-x-z)H(z)dz \\
&= -H_\varepsilon(-x)\,.
\end{aligned}$$

Außerdem folgt

$$|H_\varepsilon(x)| = |\int_{-\infty}^{\infty} \varphi_\varepsilon(x-y)H(y)dy| \le \int_{-\infty}^{\infty} \varphi_\varepsilon(x-y)dy = 1\,,$$

sowie

$$H_\varepsilon(x) = \begin{cases} -1, & x \le -\varepsilon \\ 1, & x \ge \varepsilon\,. \end{cases}$$

Ich setze nun

$$u_\varepsilon(x) = \int_{-1}^{x} H_\varepsilon(y)dy\,,\ -1 \le x \le 1\,.$$

Dann gilt $u_\varepsilon(-1) = 0$ und

$$u_\varepsilon(1) = \int_{-1}^{1} H_\varepsilon(y)dy = 0\,,$$

also $u_\varepsilon \in D$, und

$$\begin{aligned}
I(u_\varepsilon) &= \int_{-1}^{1} (u_\varepsilon'(x)^2 - 1)^2 dx \\
&= \int_{-1}^{1} (H_\varepsilon(x)^2 - 1)^2 dx
\end{aligned}$$

$$
\begin{aligned}
&= \int_{-\varepsilon}^{\varepsilon} (H_\varepsilon(x)^2 - 1)^2 dx \\
&\leq \int_{-\varepsilon}^{\varepsilon} dx = 2\varepsilon \to 0
\end{aligned}
$$

für $\varepsilon \to 0$, also $\inf_{v\in D} I(v) = 0$, wie behauptet, und I hat kein Minimum auf D.
Jedoch nimmt I auf einem größeren Raum $\tilde{D}$ das Minimum an. Zur Konstruktion von $\tilde{D}$ beachte man, daß die Familie $\{u_\varepsilon\}_{\varepsilon>0}$ für $\varepsilon \to 0$ im Raum $H_1^p((-1,1))$ mit $1 \leq p < \infty$ gegen die Funktion

$$
u(x) = \begin{cases} -(x+1), & -1 \leq x \leq 0 \\ x-1, & 0 \leq x \leq 1 \end{cases}
$$

konvergiert. Wegen $u \in H_1^p((-1,1))$ mit $u'(x) = H(x)$ gilt $I(u) = 0$. Also ist u ein Minimum von I auf dem Raum

$$
\tilde{D} = \overline{D}^{H_1^p((-1,1))} \subseteq H_1^p((-1,1)) .
$$

Dieses Minimum ist aber nicht eindeutig, vielmehr gibt es unendlich viele Minima in $\tilde{D}$. Ein weiteres Minimum ist zum Beispiel $-u$.

4.5 Beispiel (Der Fall $f(x,u,\xi) = f(x,\xi)$). In diesem Fall lautet die Eulergleichung

$$
\frac{d}{dx}[f_\xi(x,u')] = 0 ,
$$

also

$$
f_\xi(x,u') = \text{ const} .
$$

Das folgende Beispiel für diesen Fall stammt von Weierstraß: Seien $f(x,\xi) = x\xi^2$ und

$$
D = \Big\{ v \in C_1([0,1]) : v(0) = 1,\ v(1) = 0 \Big\} .
$$

Ich werde zeigen, daß das Funktional

$$
I(v) = \int_0^1 f(x,v'(x))dx = \int_0^1 x(v'(x))^2 dx
$$

kein Minimum auf D besitzt. Die Eulergleichung zu diesem Funktional ist

$$
xu'(x) = c = \text{ const} ,
$$

also

$$u'(x) = \frac{c}{x}\,.$$

Die allgemeine Lösung lautet

$$u(x) = c\,\log x + d$$

mit beliebigen Konstanten c und d. Das Randwertproblem (E) ist also nicht lösbar, weil es nicht möglich ist, c und d so zu bestimmen, daß

$$u(0) = 1, \quad u(1) = 0$$

gilt. Dies bedeutet, daß I auf D kein Minimum $u \in D \cap C_2([0,1])$ besitzt. Ich zeige, daß I überhaupt kein Minimum $u \in D$ besitzt. Hierzu beweise ich wieder, daß

$$\inf_{v \in D} I(v) = 0 \tag{$*$}$$

ist. Für ein Minimum würde hieraus folgen $x(u'(x))^2 = 0$ für fast alle $x \in [0,1]$, also $u' \equiv 0$. Natürlich gibt es keine Funktion $u \in C_1([0,1])$ mit $u(0) = 1, u(1) = 0$ und $u' \equiv 0$.

Zum Beweis, daß $\inf_{v \in D} I(v) = 0$ ist, sei $m \in N$ und

$$v_m(x) = \begin{cases} 1, & 0 \le x \le \frac{1}{m} \\ \frac{\log x}{\log \frac{1}{m}}, & \frac{1}{m} \le x \le 1 \end{cases}$$

Es ist $v_m \in H_1^p((0,1))$ für $1 \le p < \infty$ mit

$$v_m'(x) = \begin{cases} 0, & 0 \le x \le \frac{1}{m} \\ \frac{1}{x \log \frac{1}{m}}, & \frac{1}{m} \le x \le 1\,, \end{cases}$$

also

$$\begin{aligned} I(v_m) &= \int_0^1 x(v_m'(x))^2 dx = \int_{\frac{1}{m}}^1 \frac{1}{x(\log m)^2} dx \\ &= \frac{1}{(\log \frac{1}{m})^2}[\log x]_{\frac{1}{m}}^1 = -\frac{1}{(\log \frac{1}{m})^2} \log \frac{1}{m} \\ &= \frac{1}{\log m} \end{aligned}$$

Hieraus folgt $\lim_{m\to\infty} I(v_m) = 0$. Dies genügt nicht zum Beweis von $(*)$ wegen $v_m \notin D$. Deswegen konstruiere ich zu jedem $m \in \mathbb{N}$ eine Folge $\{\varphi_{mk}\}_{k=1}^{\infty}$ mit $\varphi_{mk} \in D$ und

$$\lim_{k\to\infty} \|v_m - \varphi_{mk}\|_{2,1,(0,1)} = 0\,.$$

Es folgt dann

$$\begin{aligned}
\lim_{k\to\infty} I(\varphi_{mk}) &= \lim_{k\to\infty} I(\varphi_{mk} - v_m + v_m) \\
&= \lim_{k\to\infty} \int_0^1 x((\varphi_{mk} - v_m)' + v_m')^2 dx \\
&= \lim_{k\to\infty} \Big[\int_0^1 x(\varphi_{mk}' - v_m')^2 dx \\
&\qquad + \int_0^1 2x(\varphi_{mk}' - v_m')v_m' dx + I(v_m)\Big] \\
&= I(v_m)\,,
\end{aligned}$$

wegen

$$\begin{aligned}
\lim_{k\to\infty} \int_0^1 x(\varphi_{mk}' - v_m')^2 dx &\le \lim_{k\to\infty} \int_0^1 |\varphi_{mk}' - v_m'|^2 dx \\
&= \lim_{k\to\infty} \|\varphi_{mk}' - v_m'\|_2 \le \lim_{k\to\infty} \|\varphi_{mk} - v_m\|_{2,1} = 0
\end{aligned}$$

und

$$\begin{aligned}
\lim_{k\to\infty} \int_0^1 2x(\varphi_{mk}' - v_m')v_m' dx &\le \lim_{k\to\infty} \int_0^1 |(\varphi_{mk}' - v_m')v_m'| dx \\
&\le \lim_{k\to\infty} \|\varphi_{mk}' - v_m'\|_2 \|v_m'\|_2 \le \lim_{k\to\infty} \|\varphi_{mk} - v_m\|_{2,1} \|v_m'\|_2 = 0\,.
\end{aligned}$$

Aus $\lim_{k\to\infty} I(\varphi_{mk}) = I(v_m) = \frac{1}{\log m}$ folgt $(*)$. Also bleibt zu zeigen, daß die Folge $\{\varphi_{mk}\}_{k=1}^{\infty}$ existiert. Dies ergibt sich aus folgendem

4.5.1 Lemma. Sei $u \in H_1^2((a,b))$.

(i) Dann ist u hölderstetig auf $[a,b]$ mit $|u(x) - u(y)| \le |x-y|^{1/2} \|u\|_{2,1,(a,b)}$.

(ii) Sei $\{\psi_k\}_{k=1}^{\infty}$ eine Folge mit $\psi_k \in C_\infty((a,b)) \cap H_1^2((a,b))$ und mit $\lim_{k\to\infty} \|u - \psi_k\|_{2,1,(0,1)} = 0$.

Dann gilt

$$\lim_{k\to\infty} \psi_k(y) = u(y)$$

für alle $y \in [a,b]$.

(iii) Es gibt eine Folge $\{\varphi_k\}_{k=1}^{\infty}$ mit

$$\varphi_k \in C_\infty((a,b)) \cap C([a,b]) \cap H_1^2((a,b)), \varphi_k(a) = u(a), \varphi_k(b) = u(b)$$

und mit

$$\lim_{k\to\infty} \|u - \varphi_k\|_{2,1,(0,1)} = 0\,.$$

Beweis: Sei $a < y < b$. Sei

$$\chi_y(x) = \begin{cases} \alpha_1 \sinh(x) + \beta_1 \cosh(x)\,, & a \le x < y \\ \alpha_2 \sinh(x) + \beta_2 \cosh(x)\,, & y < x \le b\,, \end{cases}$$

wobei die Koeffizienten $\alpha_1, \beta_1, \alpha_2, \beta_2$ so bestimmt werden, daß

$$\chi_y'(a) = 0, \quad \chi_y'(b) = 0\,, \quad \chi_y(y-) = \chi_y(y+), \quad \chi_y'(y-) - \chi_y'(y+) = 1\,.$$

Dies führt auf ein lineares Gleichungssystem für $\alpha_1\,, \ldots\,, \beta_2\,,$ das immer lösbar ist. Für $x \neq y$ gilt dann auch $\chi_y''(x) = \chi_y(x)$.
Für $y = a$ setze

$$\chi_a(x) = \alpha \sinh(x) + \beta \cosh(x)$$

mit

$$-\chi_a'(a) = 1, \quad \chi_a'(b) = 0$$

und entsprechend für $y = b$.
Für $v \in H_1^2((0,1)) \cap C([0,1])$ ergibt sich nun

$$\begin{aligned} (v, \chi_y)_{1,(a,b)} &= (v, \chi_y)_{(a,b)} + (v', \chi_y')_{(a,b)} \\ &= (v, \chi_y)_{(a,y)} + (v', \chi_y')_{(a,y)} \\ &\quad + (v, \chi_y)_{(y,b)} + (v', \chi_y')_{(y,b)} \\ &= (v, \chi_y)_{(a,y)} - (v, \chi'')_{(a,y)} + (v, \chi_y)_{(y,b)} - (v, \chi_y'')_{(y,b)} \\ &\quad + v(y)\chi_y'(y) - v(a)\chi_y'(a) + v(b)\chi_y'(b) - v(y)\chi_y'(y+) \\ &= v(y)[\chi_y'(y-) - \chi_y'(y+)] = v(y)\,, \end{aligned} \qquad (*)$$

und somit

$$\begin{aligned}|\psi_k(y) - \psi_l(y)| &= |(\psi_k, \chi_y)_{1,(a,b)} - (\psi_l, \chi_y)_{1,(a,b)}| \\ &= |\psi_k - \psi_l, \chi_y)_{1,(a,b)}| \\ &\leq \|\psi_k - \psi_l\|_{2,1,(a,b)} \|\chi_y\|_{2,1,(a,b)} \\ &\to 0 \text{ für } k, l \to \infty\end{aligned}$$

Also ist $\{\psi_k(y)\}_{k=1}^{\infty}$ eine Cauchyfolge und somit konvergent in $\mathbb{K}$. Da es eine Nullmenge N und eine Teilfolge $\{\psi_{k_l}\}_{l=1}^{\infty}$ gibt mit

$$\lim_{l\to\infty} \psi_{k_l}(y) = u(y)$$

für alle $y \in [a, b]\backslash N$, gilt für diese y sogar

$$\lim_{k\to\infty} \psi_k(y) = u(y) .$$

Aus der Hölderschen Ungleichung folgt für alle $x, y \in [a, b]$

$$\begin{aligned}|\psi_k(x) - \psi_k(y)| &= |\int_x^y \psi_k'(z)dz| \\ &\leq |x - y|^{1/2} (\int_x^y |\psi_k'(z)|^2 dz)^{1/2} \\ &\leq |x - y|^{1/2} \|\psi_k\|_{2,1,(a,b)} ;\end{aligned}$$

für $x, y \in [a, b]\backslash N$ ergibt sich also

$$\begin{aligned}|u(x) - u(y)| &= \lim_{k\to\infty} |\psi_k(x) - \psi_k(y)| \\ &\leq \lim_{k\to\infty} |x - y|^{1/2} \|\psi_k\|_{2,1(a,b)} \\ &= |x - y|^{1/2} \|u\|_{2,1(a,b)}\end{aligned}$$

Folglich ist u gleichmäßig hölderstetig auf $[a, b]\backslash N$, und kann damit in eindeutiger Weise zu einer hölderstetigen Funktion auf $[a, b]$ fortgesetzt werden. Dies beweist (i). Wegen $u \in H_1^2((a, b)) \bigcap C([a, b])$ folgt aus $(*)$ für alle $y \in [a, b]$:

$$\lim_{k\to\infty} \psi_k(y) = \lim_{k\to\infty} (\psi_k, \chi_y)_{1,(a,b)} = (u, \chi_y)_{1,(a,b)}$$

Hiermit ist (ii) gezeigt.

Zum Beweis von (iii) setze

$$h_k(x) = [u(b) - \psi_k(b)]\frac{x-a}{b-a} + [u(a) - \psi_k(a)]\frac{b-x}{b-a}$$

Dann gilt $h_k \in C_\infty([a,b]), h_k(b) = u(b) - \psi_k(b), h_k(a) = u(a) - \psi_k(a)$, sowie

$$\begin{aligned}
\|h_k\|_{2,1(a,b)} &\leq \|[u(b) - \psi_k(b)]\frac{x-a}{b-a}\|_{2,1} \\
&\quad + \|[u(a) - \psi_k(a)]\frac{b-x}{b-a}\|_{2,1} \\
&\leq |u(b) - \psi_k(b)|\,\|\frac{x-a}{x-b}\|_{2,1} + |u(a) - \psi_k(a)|\,\|\frac{b-x}{b-a}\|_{2,1} \to 0
\end{aligned}$$

für $k \to \infty$, nach Aussage (i) des Lemmas. Für $\varphi_k = \psi_k + h_k$ ergibt sich folglich

$$\varphi_k \in C_\infty((a,b)) \cap C([a,b]), \varphi_k(b) = u(b), \varphi_k(a) = u(a)$$

und

$$\begin{aligned}
\|u - \varphi_k\|_{2,1,(a,b)} &= \|u - \psi_k - h_k\|_{2,1} \\
&\leq \|u - \psi_k\|_{2,1} + \|h_k\|_{2,1} \to 0 \text{ für } k \to \infty\,.
\end{aligned}$$

4.6 Beispiel (Der Fall $f(x,u,\xi) = f(u,\xi)$). In diesem Fall lautet die Eulergleichung

$$\frac{d}{dx}[f_\xi(u,u')] = f_u(u,u')\,,$$

also

$$\begin{aligned}
\frac{d}{dx}\Big[f(u(x),u'(x)) - u'(x)f_\xi(u(x),u'(x))\Big] &= f_u u' + f_\xi u'' - u'' f_\xi - u'\frac{d}{dx}f_\xi \\
&= -u'\Big(\frac{d}{dx}[f_\xi(u,u')] - f_u(u,u')\Big) = 0\,.
\end{aligned}$$

Also ist $u \in C_2([a,b])$ eine Lösung der Eulergleichung genau dann wenn

$$f(u(x),u'(x)) - u'(x)f_\xi(u(x),u'(x)) = \text{ const} \qquad (*)$$

gilt. Wie in den anderen bisher betrachteten Fällen konnte die Eulergleichung auf eine Differentialgleichung erster Ordnung zurückgeführt werden. (Man sagt, es existiere ein **erstes Integral** der Differentialgleichung.)
In den folgenden beiden Beispielen 4.6.1 und 4.6.2 ist f von dieser Form:

4.6.1 Brachistochronenproblem. Ich betrachte das Brachistochronenproblem aus Abschnitt 1.2.2. Sei $\beta < 0$,

$$D = \Big\{ v \in C_1([a,b],\mathbb{R}) : v(a) = 0, v(b) = -\beta \Big\} ,$$

und sei

$$I(v) = \int_a^b \sqrt{\frac{1+v'(x)^2}{2gv(x)}} dx = \frac{1}{\sqrt{2g}} \int_a^b v^{-1/2}(1+v'^2)^{1/2} dx \, .$$

Gesucht ist $u \in D$ mit $I(u) = min_{v\in D} I(v)$. Dann ist $w = -u$ Lösung des Brachistochronenproblems, und u stellt die Fallhöhe dar. Die Eulergleichung zu diesem Funktional lautet

$$\frac{d}{dx}\left[\frac{u'}{u^{1/2}(1+u'^2)^{1/2}}\right] = -\frac{1}{2} u^{-3/2}(1+u'^2)^{1/2} ,$$

und die Gleichung $(*)$ ist

$$\begin{aligned} u^{-1/2}(1+u'^2)^{1/2} - u'\frac{u'}{u^{1/2}(1+u'^2)^{1/2}} &= \frac{(1+u'^2)-u'^2}{u^{1/2}(1+u'^2)^{1/2}} \\ &= u^{-1/2}(1+u'^2)^{-1/2} = c \, , \end{aligned}$$

wobei c eine Konstante sei. Um diese Gleichung zu lösen, sei

$$\varphi(x) := \arg\tan u'(x) \, ,$$

also

$$u'(x) = \tan\varphi(x) \, .$$

Die Eulergleichung ergibt dann

$$u^{-1/2}(1+u'^2)^{-1/2} = u^{-1/2}(1+\tan^2\varphi)^{-1/2} = c \, ,$$

also

$$u(x)^{-1/2}\cos\varphi(x) = c \, ,$$

folglich

$$u(x) = \frac{1}{c^2}\big(\cos\varphi(x)\big)^2 \,.$$

Dies hilft zunächst nicht weiter, da weder $u(x)$ noch $\varphi(x)$ bekannt sind, aber man kennt nun u als Funktion von φ. Mit $\chi(h) = \frac{1}{c^2}(\cos h)^2$ und $\psi = \varphi^{-1}$ folgt somit wegen $\varphi^{-1} = u^{-1} \circ \chi$, daß

$$\begin{aligned}
\frac{d}{dh}\psi(h) &= \frac{d}{dh}(u^{-1} \circ \chi)(h) = \frac{1}{u'\Big(u^{-1}(\chi(h))\Big)}\chi'(h) \\
&= \frac{-1}{\tan\Big[\varphi(\varphi^{-1}(h))\Big]}\frac{2}{c^2}(\cos h \sin h) \\
&= -\frac{2}{c^2}\frac{\cos h \sin h}{\tan h} = -\frac{2}{c^2}\cos^2 h \,,
\end{aligned}$$

folglich

$$\psi'(h) = -\frac{2}{c^2}\,\frac{1}{2}\big(\cos(2h) + 1\big) \,,$$

oder

$$\begin{aligned}
x(h) &= \varphi^{-1}(h) = \psi(h) = -\frac{1}{2c^2}\sin(2h) - \frac{1}{c^2}h + k \\
u(h) &= \chi(h) = \frac{1}{c^2}\cos^2 h = \frac{1}{2c^2}\cos(2h) + \frac{1}{2c^2}
\end{aligned}$$

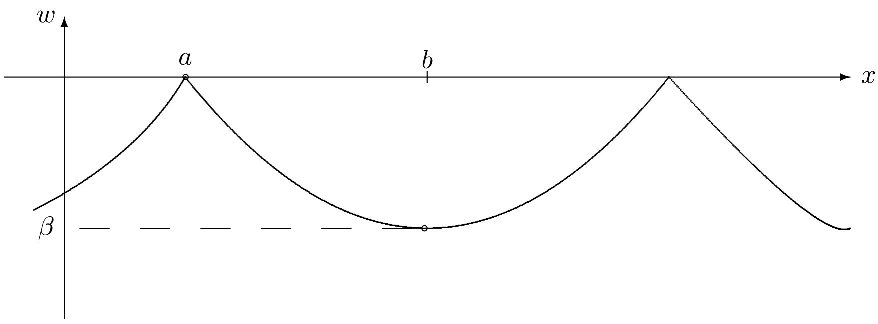
und somit mit $k' = -\frac{1}{c^2}$

$$\begin{aligned}
x(h) &= \frac{1}{2}k'\sin(2h) + k'h + k \\
w(h) &= -\,u(h) = \frac{1}{2}k'\cos(2h) + \frac{1}{2}k'
\end{aligned}$$

Dies ist die Parameterdarstellung einer Zykloide. Die beiden Konstanten k, k' müssen aus

$$\begin{aligned}
x(h_0) &= a, & w(h_0) &= \alpha = 0 \\
x(h_1) &= b, & w(h_1) &= \beta < 0
\end{aligned}$$

bestimmt werden.



Es folgt: Falls das Brachistochronenproblem eine zweimalig stetig differenzierbare Lösung hat, muß es eine Zykloide sein.

4.6.2 Wirtingersche Ungleichung. Sei $\lambda \geq 0$ und seien

$$\begin{aligned} f_\lambda(u,\xi) &= \frac{1}{2}(\xi^2 - \lambda^2 u^2)\,; \\ D &= \{v \in C_1([0,1]) :\ v(0) = v(1) = 0\} \end{aligned}$$

und

$$\begin{aligned} I(v) &= \int_0^1 f_\lambda(u(x), u'(x))dx = \frac{1}{2}\int_0^1 (|u'(x)|^2 - \lambda^2 |u(x)|^2)dx \\ &= \frac{1}{2}(|u|^2_{2,1,(0,1)} - \lambda^2 \|u\|^2_{2,(0,1)})\,. \end{aligned}$$

Die Eulergleichung lautet in diesem Fall wegen $\frac{\partial}{\partial \xi} f_\lambda(u,\xi) = \xi$ und $\frac{\partial}{\partial u} f_\lambda(u,\xi) = -\lambda^2 u$:

$$u'' + \lambda^2 u = 0\,, \quad u(0) = u(1) = 0\,.$$

Das erste Integral dieser Gleichung ist $-u'(x)^2 + \frac{1}{2}[u'(x)^2 - \lambda^2 u(x)^2] = c$, also

$$\frac{1}{2}u'(x)^2 + \frac{\lambda^2}{2}u(x)^2 = c\,.$$

Für jedes λ ist $u = 0$ eine Lösung des Eulergleichung. Dies ist jedoch nicht in jedem Fall ein Minimum des Funktionals I, wie ich gleich zeigen werde. Falls eine Funktion $u \in D$ existiert mit $I(u) < 0$, dann besitzt I kein Minimum, wegen $\alpha u \in D$ für $\alpha \in \mathbb{R}$ und

$$\begin{aligned} I(\alpha u) &= \frac{1}{2}\int_0^1 (|\alpha u'(x)|^2 - \lambda^2 |\alpha u(x)|^2)dx \\ &= \alpha^2 I(u) \to -\infty \end{aligned}$$

für $|\alpha| \to \infty$.

Wenn also I ein Minimum auf D besitzt, folgt $I(u) \geq 0$ für alle $u \in D$, also

$$\int_0^1 |u'(x)|^2 dx \geq \lambda^2 \int_0^1 |u(x)|^2 dx$$

für alle $u \in C_1([0,1])$ mit $u(0) = u(1) = 0$. Natürlich ist in diesem Fall die Funktion $u \equiv 0$ ein Minimum von I auf D, wegen

$$0 = I(u) \leq I(v)\,, v \in D\,.$$

Aus der Poincaréschen Ungleichung (Satz 3.14) folgt, daß die obenstehende Ungleichung gilt für $\lambda^2 \leq 2$, wegen

$$\begin{aligned} \int_0^1 |u'(x)|^2 dx &= |u|^2_{2,1,(0,1)} \\ &\geq 2\|u\|^2_{2,(0,1)} \geq \lambda^2 \int_0^1 |u(x)|^2 dx\,. \end{aligned}$$

Es gilt aber sogar

$$\int_0^1 |u'(x)|^2 dx \geq \pi^2 \int_0^1 |u(x)|^2 dx\,.$$

Diese Ungleichung heißt Wirtingersche Ungleichung. Bevor ich diese Ungleichung beweise, zeige ich noch, daß diese Ungleichung für $\lambda^2 > \pi^2$ nicht gilt: Denn für

$$u(x) = \sin \pi x \in D$$

ist

$$\begin{aligned} I(u) &= \frac{1}{2}\int_0^1 u'(x)^2 - \lambda^2 u(x)^2 dx \\ &= -\frac{1}{2}\int_0^1 u(x)u''(x) + \lambda^2 u(x)^2 dx \\ &= -\frac{1}{2}\int_0^1 [u''(x) + \lambda^2 u(x)]u(x) dx \\ &= \frac{1}{2}(\pi^2 - \lambda^2)\int_0^1 (\sin \pi x)^2 dx < 0\,. \end{aligned}$$

Beweis der Wirtingerschen Ungleichung. Ich zeige zuerst, daß $\{u_m\}_{m=1}^\infty$ mit $u_m(x) = \sqrt{\frac{2}{1+(m\pi)^2}}\sin(m\pi x)$ ein vollständiges Orthonormalsystem ist im Hilbertraum $\mathring{H}_1((0,1)) = \mathring{H}_1^2((0,1))$. Hierzu muß zuerst gezeigt werden, daß $u_m \in \mathring{H}_1((0,1))$ ist für $m \in \mathbb{N}$. Sei für $\varepsilon > 0$ und $x \in \mathbb{R}$

$$v_{m,\varepsilon}(x) = \begin{cases} u_m((1+\varepsilon)(x-\frac{1}{2})+\frac{1}{2}), & \frac{\varepsilon}{2(1+\varepsilon)} \leq x \leq \frac{2+\varepsilon}{2(1+\varepsilon)} \\ 0, & \text{sonst} \end{cases}$$

Wegen $v_{m,\varepsilon}(x) = 0$ für $x = \frac{\varepsilon}{2(1+\varepsilon)}$ und $x = \frac{2+\varepsilon}{2(1+\varepsilon)}$ ist $v_{m,\varepsilon} \in H_1(\mathbb{R})$ mit der schwachen Ableitung

$$v'_{m,\varepsilon}(x) = \begin{cases} (1+\varepsilon)u'_m((1+\varepsilon)(x-\frac{1}{2})+\frac{1}{2}), & \frac{\varepsilon}{2(1+\varepsilon)} \leq x \leq \frac{2+\varepsilon}{2(1+\varepsilon)} \\ 0, & \text{sonst,} \end{cases}$$

und verschwindet in einer Umgebung der Randpunkte $x = 0$ und $x = 1$. Nach Folgerung 3.8 gibt es eine Folge $\{\varphi_k\}_{k=1}^\infty$ mit $\varphi_k \in \mathring{C}_\infty((0,1))$ und $\|v_{m,\varepsilon} - \varphi_k\|_{2,1,\Omega} \to 0$ für $k \to \infty$, also ist $v_{m,\varepsilon} \in \mathring{H}_1((0,1))$. Außerdem gilt $\|u_m - v_{m,\varepsilon}\|_{2,1,(0,1)} \to 0$ für $\varepsilon \to 0$, also folgt

$$u_m \in \overline{\mathring{H}_1((0,1))} = \mathring{H}_1((0,1)).$$

Weiter muß gezeigt werden, daß $\{u_m\}_{m=1}^\infty$ ein Orthonormalsystem ist. Es gilt

$$\begin{aligned} (u_k, u_m)_{1,(0,1)} &= (u'_k, u'_m)_{(0,1)} + (u_k, u_m)_{(0,1)} \\ &= \begin{cases} -(u''_k, u_m) + (u_k, u_m) = [(k\pi)^2+1](u_k, u_m)_{(0,1)} \\ -(u_k, u''_m) + (u_k, u_m) = [(m\pi)^2+1](u_k, u_m)_{[0,1)}, \end{cases} \end{aligned}$$

und somit

$$[(k\pi)^2 - (m\pi)^2](u_k, u_m)_{(0,1)} = 0,$$

was für $k \neq m$ nur sein kann wenn $(u_k, u_m)_{(0,1)} = 0$ und somit auch

$$(u_k, u_m)_{1,(0,1)} = 0.$$

Für $k = m$ erhält man

$$\begin{aligned}(u_m, u_m)_{1,(0,1)} &= [(m\pi)^2 + 1](u_m, u_m)_{0,1} \\ &= [(m\pi)^2 + 1]\frac{2}{(m\pi)^2 + 1}\int_0^1 \sin(m\pi x)^2 dx \\ &= 1\,.\end{aligned}$$

Also ist $\{u_m\}_{m=1}^{\infty}$ ein Orthonormalsystem. Es bleibt zu zeigen, daß dieses Orthonormalsystem vollständig ist.

Es sei X der von $\{u_m\}_{m=1}^{\infty}$ aufgespannte Teilraum von $\overset{\circ}{H}_1((0,1))$, d.h. die Menge aller endlichen Linearkombinationen der u_m's. Es sei $\overline{X}$ der Abschluß von X. Bekanntlich gilt

$$\overline{X} = \left\{ \sum_{m=1}^{\infty} \alpha_m u_m : \sum_{m=1}^{\infty} \alpha_m u_m \text{ konvergiert bezüglich der } H_1^2\text{–Norm} \right\}.$$

Wenn für diesen abgeschlossenen Unterraum $\overline{X} = \overset{\circ}{H}_1((0,1))$ gilt, dann ist $\{u_m\}_{m=1}^{\infty}$ vollständig. Ich zeige, daß $\overset{\circ}{C}_\infty((0,1)) \subseteq \overline{X}$ ist. Wegen $\overline{\overset{\circ}{C}_\infty((0,1))} = \overset{\circ}{C}_1((0,1))$, folgt dann

$$\overset{\circ}{H}_1((0,1)) = \overline{\overset{\circ}{C}_\infty((0,1))} \subseteq \overline{X} \subseteq \overset{\circ}{H}_1((0,1))\,,$$

und somit muß $\{u_m\}_{m=1}^{\infty}$ vollständig sein. Sei $\varphi \in \overset{\circ}{C}_\infty(0,1)$ und $\alpha_m = (\varphi, u_m)_{1,(0,1)}$. Es folgt mit partieller Integration

$$\begin{aligned}\alpha_m &= (\varphi, u_m)_{1,(0,1)} = (\varphi, u_m)_{(0,1)} + (\varphi', u_m')_{(0,1)} \\ &= (\varphi, u_m)_{(0,1)} - (\varphi, u_m'')_{(0,1)} = (\varphi, u_m)_{(0,1)} + (\pi m)^2(\varphi, u_m)_{(0,1)} \\ &= (1 + (\pi m)^2)\sqrt{\frac{2}{1 + (\pi m)^2}}(\varphi, \sin(m\pi x))_{(0,1)} \\ &= \sqrt{2(1 + (\pi m)^2)}(\varphi, \sin(m\pi x))_{(0,1)}\,.\end{aligned}$$

Die Reihe $\sum_{m=1}^{\infty} \alpha_m u_m$ konvergiert in $\overset{\circ}{H}_1((0,1))$. Die Grenzfunktion bezeichne ich mit ψ. Insbesondere konvergiert dann $\sum_{m=1}^{\infty} \alpha_m u_m$ bezüglich der L^2–Norm gegen ψ. Andererseits gilt

$$\sum_{m=1}^{\infty} \alpha_m u_m = \sum_{m=1}^{\infty} \sqrt{2(1 + (m\pi)^2)}\sqrt{\frac{2}{1 + (m\pi)^2}}(\varphi, \sin(m\pi x))_{(0,1)} \sin(m\pi x)$$

$$= \sum_{m=1}^{\infty} (\varphi, \sqrt{2}\sin(m\pi x))_{(0,1)} \, [\sqrt{2}\sin(m\pi x)] \,,$$

und diese Reihe konvergiert bezüglich der L^2–Norm gegen φ, weil $\{\sqrt{2}\sin(m\pi x)\}_{m=1}^{\infty}$ ein vollständiges Orthonormalsystem ist in $L^2((0,1))$, also gilt $\varphi = \psi$ fast überall. Folglich ist $\varphi = \sum_{m=1}^{\infty} \alpha_m u_m \in \overline{X}$. Da $\varphi \in \overset{\circ}{C}_\infty((0,1))$ beliebig war, folgt $\overset{\circ}{C}_\infty((0,1)) \subseteq \overline{X}$, also ist $\{u_m\}_{m=1}^{\infty}$ vollständig.

Nun kann die Wirtingersche Ungleichung bewiesen werden. Für $u \in \overset{\circ}{H}_1((0,1))$ sei

$$\alpha_m = (u, u_m)_{1,(0,1)} \,.$$

Dann gilt

$$u = \sum_{m=1}^{\infty} \alpha_m u_m \,,$$

und

$$\int_0^1 |u'(x)|^2 dx = \|u'\|_{0,1}^2 = \| \sum_{m=1}^{\infty} \alpha_m u_m' \|_{(0,1)}^2 \,.$$

Die Reihe $\sum_{m=1}^{\infty} \alpha_m u_m'$ konvergiert bezüglich der L^2–Norm. Wegen der Stetigkeit der Norm folgt somit

$$\begin{aligned}
\int_0^1 |u'(x)|^2 dx &= \| \lim_{k\to\infty} \sum_{m=1}^{k} \alpha m u_m' \|_{(0,1)}^2 \\
&= \lim_{k\to\infty} \| \sum_{m=1}^{k} \alpha_m u_m' \|_{(0,1)}^2 = \lim_{k\to\infty} (\sum_{m=1}^{k} \alpha_m u_m', \sum_{l=1}^{k} \alpha_l u_l')_{(0,1)} \\
&= \lim_{k\to\infty} \sum_{m=1}^{k} \sum_{l=1}^{k} (\alpha_m u_m', \alpha_l u_l')_{(0,1)} = \lim_{k\to\infty} \sum_{m,l=1}^{k} \alpha_m \overline{\alpha}_l (u_m', u_l')_{(0,1)} \\
&= -\lim_{k\to\infty} \sum_{m=1}^{k} \alpha_m \overline{\alpha}_l (u_m'', u_l)_{(0,1)} = \lim_{k\to\infty} \sum_{m=1}^{k} \alpha_m \overline{\alpha}_l (\pi m)^2 (u_m, u_l)_{(0,1)} \\
&= \lim_{k\to\infty} \sum_{m=1}^{k} |\alpha_m|^2 (\pi m)^2 (u_m, u_m)_{(0,1)} \geq \pi^2 \lim_{k\to\infty} \sum_{m=1}^{\infty} |\alpha_m|^2 (u_m, u_m)_{(0,1)}
\end{aligned}$$

$$
\begin{aligned}
&= \pi^2 \lim_{k\to\infty} \sum_{m,l=1}^{k} \alpha_m \overline{\alpha}_l (u_m, u_l)_{(0,1)} = \pi^2 \lim_{k\to\infty} \Big(\sum_{m=1}^{k} \alpha_m u_m, \sum_{l=1}^{k} \alpha_l u_l \Big)_{(0,1)} \\
&= \pi^2 \lim_{k\to\infty} \Big\| \sum_{m=1}^{k} \alpha_m u_m \Big\|_{(0,1)}^2 = \pi^2 \|u\|_{(0,1)}^2 \\
&= \pi^2 \int_0^1 |u(x)|^2 dx \,,
\end{aligned}
$$

weil $\sum_{m=1}^{\infty} \alpha_m u_m$ bezüglich der L^2–Norm gegen u konvergiert, und $(u_m, u_l)_{(0,1)} = 0$ ist für $m \neq l$.
Dies beweist die Wirtingersche Ungleichung.

4.7 Beispiel (Der allgemeine Fall. Das Fermatsche Prinzip). Bei der Bestimmung des kürzesten Lichtweges (Beispiel 1.2.1) ist das Funktional

$$
I(v) = \int_a^b f(x, v(x), v'(x)) dx
$$

zu minimieren mit

$$
f(x, u, \xi) = n(x, u) \sqrt{1 + \xi^2} \,.
$$

In diesem Fall lautet die Eulersche Differentialgleichung

$$
\frac{d}{dx} \left(n(x, u(x)) \frac{u'(x)}{\sqrt{1 + u'(x)^2}} \right) = \frac{\partial}{\partial u} n(x, u(x)) \sqrt{1 + u'(x)^2} \,.
$$

Diese Gleichung ist sehr unübersichtlich. Eine übersichtlichere Form erhält man später mit der Hamilton–Jacobi–Theorie.

4.8 Lemma (Zweite Form der Eulergleichung) Sei $f \in C_2([a, b] \times \mathbb{R} \times \mathbb{R})$. Wenn $u \in C_2([a, b])$ eine Lösung der Eulergleichung ist, dann gilt

$$
\frac{d}{dx} \Big[f(x, u(x), u'(x)) - u'(x) f_\xi(x, u(x), u'(x)) \Big] = f_x(x, u(x), u'(x))
$$

Beweis: Es gilt

$$\begin{aligned}
&\frac{d}{dx}[f(x,u,u') - u'f_\xi(x,u,u')] - f_x(x,u,u') \\
&\qquad = f_x + f_u u' + f_\xi u'' - f_\xi u'' - u'\frac{d}{dx}f_\xi - f_x = u'(f_u - \frac{d}{dx}f_\xi) = 0\,,
\end{aligned}$$

weil $f_u - \frac{d}{dx}f_\xi = 0$ ist. Damit ist das Lemma bewiesen.

5. Klassische Methoden. Hamilton-Jacobi Theorie

In diesem Abschnitt werde ich die Hamilton-Jacobi Theorie darstellen, die auch zu den klassischen Methoden der Variationsrechnung gehört. Ich diskutiere die kanonische Form der Eulergleichung und die Hamilton-Jacobi Gleichung. Neben anderen Beispielen für die Anwendung der Theorie betrachte ich das Fermatsche Prinzip des kürzesten Lichtweges, leite die kanonische Form der Eulergleichung zu diesem Variationsproblem her und zeige, daß in diesem Beispiel die Hamilton-Jacobi Gleichung mit der Eikonalgleichung der geometrischen Optik übereinstimmt.

Als Hilfsmittel benötige ich die Legendretransformation, die ich zunächst studiere.

Nach Lemma 4.2 (i) gilt für jede konvexe Funktion $f \in C_1(\mathbb{R}^n, \mathbb{R})$, daß

$$f(x) \geq f(y) + \nabla f(y) \cdot (x - y)$$

für alle $x, y \in \mathbb{R}^n$. Im folgenden Satz wird dieses Ergebnis auf beliebige konvexe Funktionen verallgemeinert:

5.1 Satz. Sei $f : \mathbb{R}^n \to \mathbb{R}$ konvex. Dann gibt es zu jedem $y \in \mathbb{R}^n$ ein $z \in \mathbb{R}^n$ mit

$$f(x) \geq f(y) + z \cdot (x - y)$$

für alle $x \in \mathbb{R}^n$.

Beweis: Im Beweis brauche ich die beiden folgenden Resultate:
(i) Jede konvexe Funktion $f : \mathbb{R}^n \to \mathbb{R}$ ist stetig. (Diese Aussage gilt nicht für konvexe Funktionen $f : X \to \mathbb{R}$ mit beliebigem Vektorraum X als Definitionsbereich.)
(ii) (Trennungssatz.) Sei C eine nichtleere, konvexe und offene Teilmenge des $\mathbb{R}^m$ und sei ξ ein Punkt des $\mathbb{R}^m$, der nicht zu C gehört. Dann gibt es eine Hyperebene im $\mathbb{R}^m$, die C von ξ trennt.

Ich will beide Aussagen nicht beweisen. Die erste ist ein Ergebnis aus der Infinitesimalrechnung für Funktionen mehrerer Variablen, und die zweite ist ein einfacher Spezialfall eines funktionalanalytischen Resultates, das für allgemeine topologische Vektorräume gilt (siehe H.H. Schäfer: Topological Vector Spaces, Springer 1966, S. 64).

Es sei nun $f : \mathbb{R} \to \mathbb{R}$ eine konvexe Funktion und

$$\operatorname{epi}(f) = \{(x, \lambda) \in \mathbb{R}^n \times \mathbb{R} : \lambda \geq f(x)\}$$

der Epigraph von f. Weil f nach (i) stetig ist, ist das Innere der Menge $\operatorname{epi}(f)$ nicht leer, also

$$\overset{\circ}{\widehat{\operatorname{epi}}}\,(f) \neq \emptyset\,.$$

Aus der Konvexität von f folgt sofort, daß $\operatorname{epi} f$ eine konvexe Teilmenge von $\mathbb{R}^{n+1}$ ist. Sei nun $y \in \mathbb{R}^n$. Dann gehört der Punkt $(y, f(y))$ zum Rand von $\operatorname{epi}(f)$, also ist $(y, f(y)) \notin \overset{\circ}{\widehat{\operatorname{epi}}}\,(f)$, und somit gibt es nach (ii) eine Hyperebene in $\mathbb{R}^{n+1}$, die die nichtleere, konvexe und offene Menge $\overset{\circ}{\widehat{\operatorname{epi}}}\,(f)$ von der einpunktigen, also konvexen und nichtleeren Menge $\{(y, f(y))\}$ trennt.

Diese Hyperebene muß den Punkt $(y, f(y))$ enthalten. Also kann man diese Hyperebene als Graph einer Funktion

$$x \mapsto z \cdot (x - y) + f(y)$$

darstellen mit geeignetem $z \in \mathbb{R}^n$. Da die Hyperebene keinen Punkt von $\overset{\circ}{\widehat{\operatorname{epi}}}\,(f)$ enthält, gilt für alle $x \in \mathbb{R}^n$

$$z \cdot (x - y) + f(y) \leq f(x)\,.$$

Damit ist der Satz bewiesen.

5.2 Beispiel. Die Funktion $x \mapsto f(x) := |x| : \mathbb{R}^n \to \mathbb{R}$ ist konvex. Für $y \neq 0$ ist f differenzierbar mit

$$\nabla f(y) = \frac{y}{|y|}$$

und es gilt in diesem Fall

$$f(x) \geq f(y) + \nabla f(y) \cdot (x - y) = f(y) + \frac{y}{|y|}(x - y)\,.$$

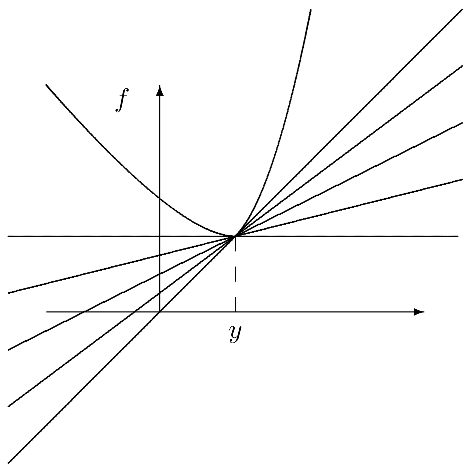
Für $y = 0$ gilt für alle $z \in \mathbb{R}^n$ mit $|z| \leq 1$

$$z \cdot x \leq |z|\,|x| \leq |x|\,,$$

also

$$f(x) \geq z \cdot x = f(0) + z \cdot (x - 0)\,,$$

Falls $f(x) \geq f(y) + z \cdot (x - y)$ für alle $x \in \mathbb{R}^n$ gilt, dann ist der Graph der Funktion $x \mapsto z \cdot (x - y) + f(y)$ eine Hyperebene durch den Punkt $(y, f(y))$, die ganz unter dem Graphen der Funktion f liegt, und z ist der Gradient der Funktion $x \mapsto z \cdot (x - y) + f(y)$.



Der Satz 5.1 motiviert die folgende

5.3 Definition (Subdifferentiale für Funktionen von n Variablen). Sei $f : \mathbb{R}^n \to (-\infty, \infty]$. Das Subdifferential ∂f von f ist eine Funktion $\partial f : \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^{\backslash})$ (= Potenzmenge von $\mathbb{R}^n$), die folgendermaßen definiert ist:

$$z \in \partial f(y) \Longleftrightarrow \forall x \in \mathbb{R}^n : f(x) \geq f(y) + z \cdot (x - y)\,.$$

Satz 5.1 sagt also, daß bei einer konvexen Funktion f das Subdifferential an keiner Stelle leer ist: $\quad \partial f(y) \neq \emptyset$ für alle $y \in \mathbb{R}^n$.

Ich werde das Subdifferential später genauer untersuchen.

5.4 Definition (Legendretransformation). Sei $f : \mathbb{R}^n \to [-\infty, \infty]$ und sei $f^* : \mathbb{R}^n \to [-\infty, \infty]$ definiert durch

$$f^*(y) = \sup_{x \in \mathbb{R}^n} \{x \cdot y - f(x)\} .$$

f^* heißt Legendretransformierte von f .

5.5 Lemma (Eigenschaften der Legendretransformation). Sei $f : \mathbb{R}^n \to [-\infty, \infty]$. Dann gilt:

(i) f^* ist konvex.

(ii) Es gilt $(f^*)^* \leq f$. Ist $f : \mathbb{R}^n \to \mathbb{R}$ konvex, dann gilt $(f^*)^* = f$.

(iii) Sei $f \in C_1(\mathbb{R}^n, \mathbb{R})$ konvex. Dann gilt

$$f^*(\nabla f(x)) = \nabla f(x) \cdot x - f(x)$$

für alle $x \in \mathbb{R}^n$.

(iv) Sei f strikt konvex und

$$R(\partial f) = \bigcup_{x \in \mathbb{R}^n} \partial f(x) .$$

Dann ist f^* stetig differenzierbar auf $R(\partial f)$.

(v) Sei $f \in C_1(\mathbb{R}^n, \mathbb{R})$ strikt konvex. Dann ist die Abbildung $x \mapsto \nabla f(x) : \mathbb{R}^n \to \mathbb{R}^n$ injektiv. Gilt außerdem

$$\lim_{|x| \to \infty} \frac{f(x)}{|x|} = \infty ,$$

dann ist diese Abbildung sogar bijektiv, und es gilt für ihre Inverse g

$$f^*(y) = y \cdot g(y) - f(g(y))$$

sowie

$$g(y) = \nabla f^*(y) .$$

Die Aussage (iv) will ich hier nicht beweisen.

Beweis: (i) Seien $y, z \in \mathbb{R}^n$ und $t \in [0,1]$. Dann gilt

$$\begin{aligned} f^*(ty + (1-t)z) &= \sup_{x \in \mathbb{R}^n} \Big[(ty + (1-t)z) \cdot x - f(x)\Big] \\ &= \sup_{x \in \mathbb{R}^n} \Big[t(y \cdot x - f(x)) + (1-t)(z \cdot x - f(x))\Big] \\ &\leq t \sup_{x \in \mathbb{R}^n} (y \cdot x - f(x)) + (1-t) \sup_{x \in \mathbb{R}^n} (z \cdot x - f(x)) \\ &= t f^*(y) + (1-t) f^*(z) \end{aligned}$$

(ii) Es ist $f^*(y) = \sup_{x \in \mathbb{R}^n} (x \cdot y - f(x))$, also $f^*(y) \geq x \cdot y - f(x)$, und somit $f(x) \geq x \cdot y - f^*(y)$ für alle $x, y \in \mathbb{R}^n$. Wegen

$$(f^*)^*(x) = \sup_{y \in \mathbb{R}^n} (x \cdot y - f^*(y))$$

folgt

$$(f^*)^*(x) \leq f(x).$$

Sei $f : \mathbb{R}^n \to \mathbb{R}$ konvex. Es gilt nach Definition für $z \in \mathbb{R}^n$

$$\begin{aligned} (f^*)^*(z) &= \sup_{y \in \mathbb{R}^n} (z \cdot y - f^*(y)) \\ &= \sup_{y \in \mathbb{R}^n} \Big(z \cdot y - \sup_{x \in \mathbb{R}^n} (y \cdot x - f(x))\Big) = \sup_{y \in \mathbb{R}^n} \inf_{x \in \mathbb{R}^n} [z \cdot y - y \cdot x + f(x)] \\ &= \sup_{y \in \mathbb{R}^n} \inf_{x \in \mathbb{R}^n} [y \cdot (z - x) + f(x)]. \end{aligned}$$

Da f konvex ist, ist nach Satz 5.1 die Menge $\partial f(z) \neq \emptyset$. Sei $y_0 \in \partial f(z)$. Dann folgt

$$f(x) \geq f(z) + y_0 \cdot (x - z),$$

also

$$y_0 \cdot (z - x) + f(x) \geq f(z)$$

und somit

$$\inf_{x \in \mathbb{R}^n} [y_0 \cdot (z - x) + f(x)] \geq f(z),$$

also

$$
\begin{aligned}
(f^*)^*(z) &= \sup_{y\in\mathbb{R}^n} \inf_{x\in\mathbb{R}^n} \; [y\cdot(z-x)+f(x)] \\
&\geq \inf_{x\in\mathbb{R}^n} [y_0\cdot(z-x)+f(x)] \geq f(z)\,.
\end{aligned}
$$

Wegen $(f^*)^*(z) \leq f(z)$ folgt

$$
(f^*)^*(z) = f(z)
$$

für alle $z\in\mathbb{R}^n$, also $(f^*)^* = f$.

(iii) Sei $x\in\mathbb{R}^n$. Das Funktional

$$
h(z) = f(z) - \nabla f(x)\cdot z
$$

ist konvex wegen

$$
\begin{aligned}
&f(ty+(1-t)z) - \nabla f(x)\cdot(ty+(1-t)z) \\
&\quad \leq tf(y)+(1-t)f(z) - t\nabla f(x)y - (1-t)\nabla f(x)\cdot z \\
&\quad = t[f(y)-\nabla f(x)\cdot y] + (1-t)[f(z)-\nabla f(x)\cdot z]\,.
\end{aligned}
$$

Außerdem gilt

$$
\nabla h(z) = \nabla f(z) - \nabla f(x)\,,
$$

also $\nabla h(x) = 0$. Aus Lemma 4.2 folgt somit

$$
h(z) \geq h(x) + \nabla h(x)\cdot(z-x) = h(x)\,,
$$

also hat h ein Minimum im Punkt x. Somit hat $-h$ ein Maximum im Punkt x, also gilt

$$
\begin{aligned}
f^*(\nabla f(x)) &= \sup_{z\in\mathbb{R}^n} [\nabla f(x)\cdot z - f(z)] \\
&= \sup_{z\in\mathbb{R}^n} (-h(z)) = \max_{z\in\mathbb{R}^n} (-h(z)) \\
&= -h(x) = \nabla f(x)\cdot x - f(x)\,.
\end{aligned}
$$

(v) Zum Beweis, daß $x\mapsto\nabla f(x):\mathbb{R}^n\to\mathbb{R}^n$ injektiv ist, seien $y,z\in\mathbb{R}^n$ mit $\nabla f(y) = \nabla f(z)$. Nach Lemma 4.2 (i) gilt für alle $x\in\mathbb{R}^n$

$$
f(x) \geq f(y) + \nabla f(y)(x-y)\,.
$$

Gäbe es ein $x \in \mathbb{R}^n$ mit $x \neq y$ und mit

$$f(x) = f(y) + \nabla f(y)(x-y)\,,$$

dann folgte wegen der strikten Konvexität von f für $0 < t < 1$

$$\begin{aligned} f(tx + (1-t)y) &< tf(x) + (1-t)f(y) \\ &= t[f(y) + \nabla f(y)(x-y)] + (1-t)f(y) \\ &= f(y) + \nabla f(y)t(x-y) \leq f(y + t(x-y)) \\ &= f(tx + (1-t)y)\,. \end{aligned}$$

Folglich muß für alle $x \in \mathbb{R}^n$ mit $x \neq y$ gelten, daß

$$f(x) > f(y) + \nabla f(y)(x-y)$$

ist, und ebenso für alle $x \neq z$

$$f(x) > f(z) + \nabla f(z)(x-z)\,.$$

Wäre nun $y \neq z$, dann folgte

$$\begin{aligned} f(z) &> f(y) + \nabla f(y)(z-y) \\ &> f(z) + \nabla f(z)(y-z) + \nabla f(y)(z-y) \\ &= f(z) + \nabla f(y)[(y-z) + (z-y)] = f(z)\,. \end{aligned}$$

Also muß $y = z$ sein, also ist $x \mapsto \nabla f(x)$ injektiv.

Nun zeige ich, daß diese Abbildung surjektiv ist falls $\lim_{|x| \mapsto \infty} \frac{f(x)}{|x|} = \infty$ gilt. Sei $y \in \mathbb{R}^n$. Nach Voraussetzung gibt es dann ein $C > 0$ mit

$$\begin{aligned} \frac{f(x)}{|x|} &\geq |y| + 1 \\ &\geq |y \cdot \frac{x}{|x|}| + 1 \end{aligned}$$

für alle $|x| \geq C$, wegen $\lim_{|x| \to \infty} \frac{f(x)}{|x|} = \infty$. Also gilt

$$\begin{aligned} \lim_{|x| \to \infty} [y \cdot x - f(x)] &= \lim_{|x| \to \infty} \Big(y \cdot \frac{x}{|x|} - \frac{f(x)}{|x|} \Big) \cdot |x| \\ &\leq \lim_{|x| \to \infty} (-|x|) = -\infty\,. \end{aligned}$$

Wegen der Stetigkeit von $x \mapsto y \cdot x - f(x)$ gibt es also ein $z \in \mathbb{R}^n$ mit

$$y \cdot z - f(z) = \max_{x \in \mathbb{R}^n} (y \cdot x - f(x)) .$$

Da $f \in C_1(\mathbb{R}^n, \mathbb{R})$ ist, folgt

$$0 \;=\; \nabla_x (y \cdot x - f(x))\big|_{x=z} = y - \nabla f(z) ,$$

also $\nabla f(z) = y$. Somit ist $x \mapsto \nabla f(x)$ surjektiv, und folglich bijektiv. Sei g die Inverse dieser Abbildung. Dann gilt mit (iii)

$$\begin{aligned} f^*(y) &= f^*\Big(\nabla f(g(y))\Big) \\ &= \nabla f(g(y)) \cdot g(y) - f(g(y)) \\ &= y \cdot g(y) - f(g(y)) . \end{aligned}$$

Es bleibt zu zeigen, daß $g(y) = \nabla f^*(y)$ ist. Nach (i) und (iv) ist $f^* \in C_1(\mathbb{R}^n)$ und konvex, und nach (ii) und (iii) gilt für alle $y \in \mathbb{R}^n$

$$\begin{aligned} f(\nabla f^*(y)) &= (f^*)^*(\nabla f^*(y)) \\ &= \nabla f^*(y) \cdot y - f^*(y) \\ &= \nabla f^*(y) \cdot y - g(y) \cdot y + f(g(y)) \end{aligned}$$

Da g die Inverse von $x \mapsto \nabla f(x)$ ist, gilt $y = \nabla f(g(y))$, und somit folgt aus dieser Gleichung

$$f(\nabla f^*(y)) = f(g(y)) + \nabla f(g(y)) \cdot [\nabla f^*(y) - g(y)] .$$

Oben wurde gezeigt, daß wegen der strikten Konvexität von f diese Gleichung nur gelten kann wenn $\nabla f^*(y) = g(y)$ gilt. Damit ist Lemma 5.5 bewiesen.

5.6 Definition (Hamiltonfunktion). Sei $a < b, f : [a, b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. Dann heißt die Funktion $H : [a, b] \times \mathbb{R} \times \mathbb{R} \to [-\infty, \infty]$

$$H(x, u, v) = \sup_{\xi \in \mathbb{R}^n} (v\xi - f(x, u, \xi))$$

die Hamiltonfunktion von f . H ist die Legendretransformierte von $f(x, u, \xi)$ bezüglich der Variablen ξ .

5.7 Satz (Eigenschaften der Hamiltonfunktion). Sei $f \in C_2([a,b] \times \mathbb{R} \times \mathbb{R})$ so daß die Funktion $\xi \mapsto f(x,u,\xi)$ strikt konvex ist für alle $(x,u) \in [a,b] \times \mathbb{R}$. Für $(x,u) \in [a,b] \times \mathbb{R}$ sei

$$W(x,u) = \{f_\xi(x,u,\xi) : \xi \in \mathbb{R}\},$$
$$W = \{(x,u,v) : (x,u) \in [a,b] \times \mathbb{R}, v \in W(x,u)\},$$

und sei H die Hamiltonfunktion von f. Dann ist $H \in C_1(W,\mathbb{R})$, und die Funktion $v \mapsto H(x,u,v) : W(x,u) \to \mathbb{R}$ ist konvex. Sei $(u,v) \in C_2([a,b]) \times C_1([a,b])$ mit $(x,u(x),v(x)) \in W$ für alle $x \in [a,b]$ eine Lösung von

$$\begin{aligned} u'(x) &= H_v(x,u(x),v(x) \\ v'(x) &= -H_u(x,u(x),v(x)) \end{aligned} \tag{1}$$

in $[a,b]$. Dann ist u eine Lösung der Eulergleichung

$$\frac{d}{dx}\Big[f_\xi(x,u(x),u'(x))\Big] = f_u(x,u(x),u'(x)) \tag{2}$$

in $[a,b]$. Wenn umgekehrt $u \in C_2([a,b],\mathbb{R})$ eine Lösung von (2) ist, dann ist die Funktion (u,v) mit

$$v(x) = f_\xi(x,u(x),u'(x))$$

eine Lösung von (1). Gilt außerdem

$$\lim_{|\xi|\to\infty} \frac{f(x,u,\xi)}{|\xi|} = \infty$$

für alle $(x,u) \in [a,b] \times \mathbb{R}$, dann ist $W(x,u) = \mathbb{R}$ für alle $(x,u) \in [a,b] \times \mathbb{R}$ und somit $W = [a,b] \times \mathbb{R} \times \mathbb{R}$.

Beweis: Ich beweise den Satz nur unter der Voraussetzung, daß $f_{\xi\xi}(x,u,\xi) > 0$ sei für alle $(x,u,\xi) \in [a,b] \times \mathbb{R} \times \mathbb{R}$. Diese Voraussetzung ist etwas stärker als die strikte Konvexität von $\xi \mapsto f(x,u,\xi)$. Aus Lemma 5.5 (i) folgt, daß $v \mapsto H(x,u,v)$ konvex ist. Nach Lemma 5.5 (v) ist $\xi \mapsto f_\xi(x,u,\xi) : \mathbb{R} \to W(x,u)$ injektiv. Sei $v \mapsto g(x,u,v) : W(x,u) \to \mathbb{R}$ die Inverse. Dann gilt

$$v = f_\xi(x,u,g(x,u,v)).$$

Wegen $f_{\xi\xi}(x,u,g(x,u,v)) > 0$ folgt hieraus und aus dem Satz über inplizite Funktionen, daß $g : W \to \mathbb{R}$ stetig differenzierbar ist. Nach Lemma 5.5 (iv) ist $v \mapsto H(x,u,v) : W(x,u) \to \mathbb{R}$ stetig differenzierbar, und nach Lemma 5.5 (iii) gilt

$$H(x,u,f_\xi(x,u,\xi)) = f_\xi(x,u,\xi) \cdot \xi - f(x,u,\xi)\,.$$

Durch Ableiten dieser Gleichung nach ξ folgt

$$H_v(x,u,f_\xi(x,u,\xi))f_{\xi\xi}(x,u,\xi) = f_{\xi\xi}(x,u,\xi) \cdot \xi + f_\xi(x,u,\xi) - f_\xi(x,u,\xi)\,,$$

also

$$\Big[H_v(x,u,f_\xi(x,u,\xi)) - \xi\Big] \cdot f_{\xi\xi}(x,u,\xi) = 0\,,$$

und folglich

$$H_v(x,u,f_\xi(x,u,\xi)) = \xi$$

für alle $\xi \in \mathbb{R}^n$, weil $f_{\xi\xi}(x,u,\xi) \neq 0$ ist. Für $\xi = g(x,u,v)$ folgt

$$H_v(x,u,v) = g(x,u,v) \in C_1(W)\,.$$

Durch Integration folgt hieraus $H \in C_1(W)$. Falls

$$\lim_{|\xi|\to\infty} \frac{f(x,u,\xi)}{|\xi|} = \infty$$

gilt, folgt aus Lemma 5.5 (v), daß $\xi \mapsto f_\xi(x,u,\xi) : \mathbb{R} \to \mathbb{R}$ bijektiv ist, also $W(x,u) = \mathbb{R}$ gilt. Um die verbleibenden Behauptungen zu beweisen benütze ich, daß nach Lemma 5.5 (v) gilt

$$H(x,u,v) = v \cdot g(x,u,v) - f(x,u,g(x,u,v))\,.$$

Wegen $g \in C_1(W)$ resultiert also für alle $(x,u,v) \in W$

$$\begin{aligned} H_u(x,u,v) &= v \cdot g_u(x,u,v) - f_u(x,u,g(x,u,v)) \\ &\quad - f_\xi(x,u,g(x,u,v))g_u(x,u,v) \\ &= \Big[v - f_\xi(x,u,g(x,u,v))\Big]\, g_u(x,u,v) \\ &\quad - f_u(x,u,g(x,u,v))\,. \end{aligned}$$

Wegen $f_\xi(x,u,g(x,u,v)) = v$ folgt also

$$H_u(x,u,v) = -f_u(x,u,g(x,u,v))$$

sowie
$$H_v(x,u,v) = g(x,u,v)\,.$$
Sei nun $u \in C_2([a,b])$ eine Lösung von (2). Man definiert
$$v(x) = f_\xi(x,u,(x),u'(x))\,.$$
Dann ist $v(x) \in W(x,u(x))$, also $(x,u(x),v(x)) \in W$ für alle $x \in [a,b]$, und es folgt durch Anwendung der Inversen auf diese Gleichung
$$\begin{aligned}
u'(x) &= g(x,u(x),v(x)) \\
&= H_v(x,u,(x),v(x)) \\
v'(x) &= \frac{d}{dx}\Big[f_\xi(x,u(x),u'(x))\Big] \\
&= f_u(x,u(x),u'(x)) \\
&= f_u\Big(x,u(x),g(x,u(x),v(x))\Big) \\
&= -H_u(x,u(x),v(x))\,.
\end{aligned}$$
Also ist (u,v) eine Lösung von (1). Sei umgekehrt $(u,v) \in C_2([a,b]) \times C_1([a,b])$ eine Lösung von (1). Dann folgt
$$u'(x) = H_v(x,u(x),v(x)) = g(x,u(x),v(x))\,,$$
also durch Anwendung der Inversen
$$f_\xi(x,u(x),u'(x)) = f_\xi\Big(x,u(x),g(x,u(x),v(x))\Big) = v(x)\,,$$
und
$$\begin{aligned}
v'(x) &= -H_u(x,u(x),v(x)) \\
&= f_u\Big(x,u(x),g(x,u(x),v(x))\Big) \\
&= f_u(x,u(x),u'(x))\,.
\end{aligned}$$
Durch Kombination der beiden letzten Gleichungen ergibt sich
$$\frac{d}{dx}\Big[f_\xi(x,u(x),u'(x))\Big] = \frac{d}{dx}v(x) = f_u(x,u(x),u'(x))\,.$$

Also ist u eine Lösung von (2). Damit ist der Satz bewiesen.

5.8 Hamiltonsche Differentialgleichungen, kanonische Form der Eulergleichung. Für $(u,v) \in C_1([a,b]) \times C([a,b])$ sei

$$J(u,v) = \int_a^b \Big[u'(x)v(x) - H(x,u(x),v(x)) \Big] dx\,.$$

Um die Eulergleichung zu diesem Variationsfunktional auszurechnen, sei $w = (u,v) \in, C_2([a,b]) \times C_1([a,b])$ ein stationärer Punkt von J, d.h. für alle $\varphi \in \overset{\circ}{C}_\infty((a,b),\mathbb{R}^2)$ sei

$$\frac{d}{dt} J(w+t\varphi)_{|_{t=0}} = 0\,,$$

also

$$\begin{aligned} 0 &= \frac{d}{dt} J(w+t\varphi)_{|_{t=0}} \\ &= \int_a^b \frac{\partial}{\partial t} \Big[(u'(x) + t\varphi_1'(x))[v(x) + t\varphi_2(x)) \\ &\quad - H(x, u(x) + t\varphi_1(x), v(x) + t\varphi_2(x)) \Big]_{|_{t=0}} dx \\ &= \int_a^b \Big[v(x)\varphi_1'(x) + u'(x)\varphi_2(x) \\ &\quad - H_u(x,u(x),v(x))\varphi_1(x) - H_v(x,u(x),v(x))\varphi_2(x) \Big] dx \\ &= \int_a^b \Big(-v'(x) - H_u(x,u(x),v(x)) \Big) \varphi_1(x) \\ &\quad + \Big(u'(x) - H_v(x,u(x),v(x)) \Big) \varphi_2(x) dx \end{aligned}$$

Insbesondere gilt dies für alle $\varphi \in \overset{\circ}{C}_\infty((a,b),\mathbb{R}^2)$ der Form $\varphi = (\varphi_1, 0)$ und der Form $\varphi = (0, \varphi_2)$, also folgt

$$\begin{aligned} u'(x) &= H_v(x,u(x),v(x)) \\ v'(x) &= -H_u(x,u(x),v(x))\,. \end{aligned}$$

Die Eulergleichungen zum Funktional J stimmen als mit dem Gleichungssystem (1) aus Satz 5.7 überein. Unter gewissen Bedingungen an f ist die Bestimmung der stationären Punkte von I äquivalent zur Bestimmung der stationären Punkte von J. Dabei versteht man unter einem stationären Punkt

von I beziehungsweise J eine Lösung der entsprechenden Eulergleichung. Man sagt, das Gleichungssystem (1) von Satz 5.7 sei die **kanonische Form der Eulergleichungen.** Man nennt es auch das **Hamiltonsche Differentialgleichungssystem.**
Wegen $H \in C_1(W)$ gilt nach Lemma 5.5

$$\begin{aligned} f(x,u,H_v(x,u,v)) &= H^*(x,u,H_v(x,u,v)) \\ &= H_v(x,u,v)\cdot v - H(x,u,v)\,. \end{aligned}$$

Falls $u'(x) = H_v(x,u(x)v(x))$ erfüllt ist, gilt also

$$f(x,u(x),u'(x)) = u'(x)v(x) - H(x,u(x),v(x))\,,$$

und somit

$$\begin{aligned} J(u,v) &= \int_a^b \Big[u'(x)v(x) - H(x,u(x),v(x))\Big]dx \\ &= \int_a^b f(x,u(x),u'(x))dx \\ &= I(u)\,. \end{aligned}$$

Da $u'(x) = H_v(x,u(x),v(x))$ genau dann gilt, wenn $v(x) = f_\xi(x,u(x),u'(x))$ ist, gilt

$$\begin{aligned} &\Big\{(u,v) \in C_2([a,b]) \times C_1([a,b]) : u'(x) = H_v(x,u(x),v(x))\Big\} \\ &\quad = \Big\{(u,v) \in C_2([a,b]) \times C_1([a,b]) : v(x) = f_\xi(x,u(x),u'(x))\Big\}\,. \end{aligned}$$

Dies ist eine Untermannigfaltigkeit von $C_2([a,b]] \times C_1([a,b])$. Die obenstehende Gleichung zeigt, daß auf dieser Untermannigfaltigkeit $J(u,v)$ mit $I(u)$ übereinstimmt. Dies zeigt, daß das Minimieren von $J(u,v)$ unter der Nebenbedingung $v(x) = f_\xi(x,u(x),u'(x))$ äquivalent ist zum Minimieren von $I(u)$.

5.9 Beispiele.

5.9.1 Das Hamiltonsche Prinzip. Nach 1.2.3 bewegt sich ein Massenpunkt vom Ort $\alpha \in \mathbb{R}$ zur Zeit a nach dem Ort $\beta \in R$ zur Zeit b, so daß

$$\int_a^b T(u'(t)) - U(t,u(t))dt = \int_a^b \Big[\frac{1}{2}mu'(t)^2 - k(t)u(t)\Big]dt$$

stationär ist, wobei $-k(t)$ die zur Zeit t auf den Massenpunkt wirkende Kraft ist. Also ist

$$f(t,u,\xi) = \frac{1}{2}m\xi^2 - k(t)u\,.$$

Die Eulergleichung ist

$$\frac{d}{dt}f_\xi(t,u(t),u'(t)) = f_u(t,u(t),u'(t))\,,$$

also

$$mu''(t) = -k(t)\,.$$

Dies ist das Newtonsche Bewegungsgesetz. Die Hamiltonfunktion H ist

$$H(t,u,v) = v\cdot g(t,u,v) - f(t,u,g(t,u,v))\,,$$

wobei $v \mapsto g(t,u,v)$ die Inverse von

$$\xi \to f_\xi(t,u,\xi) = m\xi$$

ist, also

$$g(t,u,v) = \frac{1}{m}v\,,$$

und somit

$$H(t,u,v) \quad = \quad \frac{1}{m}v^2 - \frac{1}{2}m\Big(\frac{v}{m}\Big)^2 + k(t)u = \frac{v^2}{2m} + k(t)u\,.$$

In kanonischer Form lauten die Eulergleichungen

$$\begin{aligned} u'(t) &= H_v(t,u(t),v(t)) &&= \frac{v(t)}{m} \\ v'(t) &= -H_u(t,u(t),v(t)) &&= -\,k(t)\,. \end{aligned}$$

Das zugehörige Variationsfunktional ist

$$J(u,v) = \int_a^b \Big[u'(t)v(t) - \frac{v(t)^2}{2m} - k(t)u(t)\Big]dt$$

Das Hamiltonsche Differentialgleichungssystem kann man auch folgendermaßen interpretieren: $v(t)$ ist der Impuls des Teilchens zur Zeit t, $u'(t)$ die Geschwindigkeit. Durch Angabe der Geschwindigkeit zu jeder Zeit wird die

Bewegung des Teilchens vorgeschrieben. Die Gleichung $u'(t) = \frac{1}{m}v(t)$ sagt nun, daß die Bewegung des Teilchens nicht unabhängig vom Impuls ist, den das Teilchen trägt, sondern daß der Impuls die Geschwindigkeit bestimmt. Dagegen ist $v'(t) = -k(t)$ die Newtonsche Bewegungsgleichung in ihrer allgemeinen Form.

5.9.2 Das Fermatsche Prinzip. Nach 4.7 ist zur Bestimmung des kürzesten Lichtweges zwischen (a, α) und (b, β) das Funktional

$$I(u) = \int_a^b f(x, u(x), u'(x))dx$$

zu minimieren mit

$$f(x, u, \xi) = n(x, u)\sqrt{1 + \xi^2}\,.$$

Wegen $f_{\xi\xi}(x, u, \xi) = \frac{n(x,u)}{(1+\xi^2)^{3/2}} > 0$ ist f strikt konvex. Wegen $\xi \mapsto f_\xi(x, u, \xi) = n(x, u)\frac{\xi}{\sqrt{1+\xi^2}}$ ist der Wertbereich dieser Abbildung das offene Intervall $(-n(x, u), n(x, u))$. Die Hamiltonfunktion ist

$$H(x, u, v) = \sup_{\xi\in\mathbb{R}} \,[v \cdot \xi - f(x, u, \xi)] = \sup_{\xi\in\mathbb{R}} \,[v \cdot \xi - n(x, u)\sqrt{1 + \xi^2}]$$

Also gilt

$$\begin{array}{rcccl} H(x, u, n(x, u)) & = & n(x, u) \sup_{\xi\in\mathbb{R}} (\xi - \sqrt{1 + \xi^2}) & = & 0\,, \\ H(x, u, -n(x, u)) & = & -\, n(x, u) \inf_{\xi\in\mathbb{R}} (\xi + \sqrt{1 + \xi^2}) & = & 0 \end{array}$$

und für $|v| > n(x, u)$

$$H(x, u, v) = n(x, u) \sup_{\xi\in\mathbb{R}} \left(\frac{v}{n(x, u)}\xi - \sqrt{1 + \xi^2}\right) = +\infty\,.$$

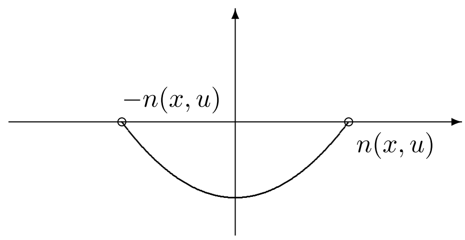
Die Abbildung

$$\xi \mapsto f_\xi(x, u, \xi) = n(x, u)\frac{\xi}{\sqrt{1 + \xi^2}} : \mathbb{R} \to (\,-\,n(x, u), n(x, u))$$

ist invertierbar. Mit der Inversen

$$v \mapsto g(x, u, v) : (\,-\,n(x, u), n(x, u)) \to \mathbb{R}$$

gilt
$$H(x,u,v) = v \cdot g(x,u,v) - f(x,u,g(x,u,v))\,,$$
wobei $v \mapsto \xi = g(x,u,v)$ die Inverse ist von
$$\xi \mapsto v = f_\xi(x,u,\xi) = n(x,u)\frac{\xi}{\sqrt{1+\xi^2}}\,,$$
also $(1+\xi^2)v^2 = n(x,u)^2\xi^2$, oder
$$(n(x,u)^2 - v^2)\xi^2 = v^2\,,$$
und somit, da ξ und v immer dasselbe Vorzeichen haben und $|v| < n(x,u)$ gilt
$$\xi = \frac{v}{\sqrt{n(x,u)^2 - v^2}} = g(x,u,v)\,.$$
Also ist
$$\begin{aligned} H(x,u,v) &= \frac{v^2}{\sqrt{n(x,u)^2 - v^2}} - n(x,u)\sqrt{1 + \frac{v^2}{n(x,u)^2 - v^2}} \\ &= \frac{v^2}{\sqrt{n(x,u)^2 - v^2}} - \frac{n(x,u)^2}{\sqrt{n(x,u)^2 - v^2}} \\ &= -\frac{n(x,u)^2 - v^2}{\sqrt{n(x,u)^2 - v^2}} \\ &= -\sqrt{n(x,u)^2 - v^2} \end{aligned}$$

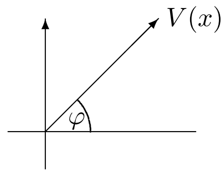


Für das Hamiltonsche Differentialgleichungssystem ergibt sich nun
$$\begin{aligned} u'(x) &= H_v(x,u(x),v(x)) = \frac{v(x)}{\sqrt{n(x,u(x))^2 - v(x)^2}} \\ v'(x) &= -H_u(x,u(x),v(x)) = \frac{n(x,u(x))}{\sqrt{n(x,u(x))^2 - v(x)^2}} n_u(x,u(x))\,. \end{aligned}$$

Man kann diese Gleichungen folgendermaßen interpretieren: Da $|v| < n$ ist, gibt es $\varphi(x) \in \big(-\frac{\pi}{2}, \frac{\pi}{2}\big)$ mit

$$v(x) = n(x, u(x)) \sin \varphi(x)\,.$$

Dies bedeutet, daß $v(x)$ die zweite Komponente eines Vektors $V(x) \in \mathbb{R}^2$ mit $|V(x)| = n(x, u(x))$ ist:



$$V(x) = n(x, u(x))\big(\cos \varphi(x), \sin \varphi(x)\big)\,.$$

Es gilt nun

$$\begin{aligned}
\frac{n(x,u(x))}{\sqrt{n(x,u(x))^2 - v(x)^2}} &= \frac{n(x,u(x))}{\sqrt{n(x,u(x))^2 - n(x,u(x))^2 \sin^2 \varphi(x)}} \\
&= \frac{1}{\sqrt{1 - \sin^2 \varphi(x)}} \\
&= \frac{1}{\cos \varphi(x)}\,,
\end{aligned}$$

also kann die zweite der Hamiltonschen Differentialgleichungen in der Form

$$v'(x) = \frac{1}{\cos \varphi(x)} n_u(x, u(x))\,,$$

oder

$$\cos \varphi(x) \frac{d}{dx} v(x) = n_u(x, u(x))$$

geschrieben werden. Da v nicht von u abhängt, kann man dies auch schreiben als

$$\frac{V}{|V|} \cdot \nabla_{(x,u)} v(x) = \cos \varphi(x) \frac{d}{dx} v(x) + \sin \varphi(x) \frac{d}{du} v(x) = n_u(x, u(x))\,.$$

Da $\frac{V}{|V|} \cdot \nabla_{(x,u)}$ die Ableitung in Richtung des Vektors V ist, sagt diese Gleichung, daß die Ableitung der zweiten Komponente des Vektors $V(x)$ in Richtung von $V(x)$ mit der Ableitung von $n(x,u)$ in Richtung der zweiten Variablen übereinstimmt.

Falls u bekannt ist, bestimmt somit die zweite Hamiltonsche Gleichung die Funktion v und damit auch V. “Bei der Bewegung entlang des Graphen von u trägt das Licht den Vektor V mit sich.“
Die erste Hamiltonsche Gleichung ist

$$\begin{aligned} u'(x) &= \frac{v(x)}{\sqrt{n(x,u(x))^2 - v(x)^2}} \\ &= \frac{n \sin\varphi(x)}{\sqrt{n^2 - n^2 \sin^2\varphi(x)}} \\ &= \frac{\sin\varphi(x)}{\cos\varphi(x)} \\ &= \tan\varphi(x) \end{aligned}$$

Diese Gleichung bedeutet, daß der Tangentenvektor $(1, u'(x))$ an den Graphen von u im Punkt $(x, u(x))$ die Richtung des Vektors $V(x)$ hat. Diese erste Hamiltonsche Gleichung erzwingt also, daß die Bewegung des Lichtes nicht unabhängig vom “mitgetragenen“ Vektor $V(x)$ ist, sondern daß sich das Licht in Richtung des Vektors $V(x)$ bewegt.

5.9.3 Der Fall $\boldsymbol{f(x,u,\xi) = f(\xi)}$. Die Eulergleichung ist in diesem Fall

$$\frac{d}{dx} f'(u') = 0\,,$$

also $f'(u'(x)) = \lambda = \text{ const}$. Die Hamiltonfunktion H ist gegeben durch

$$H(x,u,v) = H(v) = f^*(v)\,,$$

also

$$\begin{aligned} u' &= H_v(v) &&= {f^*}'(v) \\ v' &= -H_u(v) &&= 0\,. \end{aligned}$$

Hieraus folgt $v(x) = \mu = \text{const}$ und

$$u' = {f^*}'(\mu)\,,$$

also

$$u(x) = {f^*}'(\mu)x + \nu\,.$$

5.9.4 Der Fall $f(x,u,\xi) = f(x,\xi)$. In diesem Fall ist die Eulergleichung

$$\frac{d}{dx} f_\xi(x, u'(x)) = 0\,,$$

also

$$f_\xi(x, u'(x)) = \text{ const}$$

Für die Hamiltonfunktion gilt

$$H(x,u,v) = \sup_{\xi \in \mathbb{R}} [v\xi - f(x,\xi)] = H(x,v)\,.$$

Also sind die Hamiltonschen Differentialgleichungen

$$\begin{aligned} u'(x) &= \quad H_v(x, v(x)) \\ v'(x) &= -H_u(x, v(x)) = 0\,, \end{aligned}$$

also $v(x) = \mu$ und

$$u'(x) = H_v(x, \mu)\,,$$

und somit

$$u(x) = \int_a^x H_v(y,\mu)dy + u(a)\,.$$

Das Hamiltonsche Differentialgleichungssystem heißt in diesem Fall vollständig integrabel.

5.9.5 Der Fall $f(x,u,\xi) = f(u,\xi)$. Nach 4.6 gilt in diesem Fall für die Eulergleichung

$$f(u(x), u'(x)) - u'(x) f_\xi(u(x), u'(x)) = \text{ const}\,.$$

Für die Hamiltonfunktion gilt

$$H(x,u,v) = \sup_{\xi \in \mathbb{R}} [v \cdot \xi - f(u,\xi)] = H(u,v)\,,$$

also

$$\begin{aligned} u'(x) &= H_v(u(x), v(x)) \\ v'(x) &= -H_u(u(x), v(x))\,. \end{aligned}$$

Hieraus folgt

$$\frac{d}{dx}H(u(x),v(x)) = H_u(u(x),v(x))u'(x) + H_v(u(x),v(x))v'(x) = 0\,,$$

also

$$x \mapsto H(u(x),v(x)) = \text{ const}\,.$$

Wenn die Funktion f nicht von x abhängt, ist die Hamiltonfunktion entlang von Lösungskurven $x \mapsto (u(x),v(x))$ des Hamiltonschen Differentialgleichungssystems konstant.

5.10 Definition (Hamilton–Jacobi Gleichung) . Die partielle Differentialgleichung

$$S_x(x,u) + H(x,u,S_u(x,u)) = 0\,.$$

für die Funktion $S : [a,b] \times \mathbb{R} \to \mathbb{R}$ heißt Hamilton–Jacobi Gleichung.

Man beachte, daß u hier die Bedeutung einer Variablen und nicht einer Funktion hat. Die Hamilton–Jacobi Gleichung steht im engen Zusammenhang mit dem Hamiltonschen Differentialgleichungssystem, wie der folgende Satz zeigt.

5.11 Satz (Eigenschaften der Lösung der Hamilton–Jacobi Gleichung). Sei $\Gamma \subseteq \mathbb{R}$ ein Intervall und sei $S \in C_2([a,b] \times \mathbb{R} \times \Gamma, \mathbb{R})$ für jedes $\gamma \in \Gamma$ eine Lösung der Hamilton–Jacobi Gleichung:

$$S_x(x,u,\gamma) + H(x,u,S_u(x,u,\gamma)) = 0\,, \quad (x,u) \in [a,b] \times \mathbb{R}\,.$$

Sei $\gamma \in \mathbb{R}$ und sei (u,v) eine Lösung von

$$\begin{aligned} u'(x) &= H_v(x,u(x),v(x)) \\ v'(x) &= -\,H_u(x,u(x),v(x)) \end{aligned} \tag{H}$$

für $x \in [a,b]$ und mit $v(a) = S_u(a,u(a),\gamma)$. Dann existiert eine Konstante c mit

$$\begin{aligned} S_\gamma(x,u(x),\gamma) &= c \\ S_u(x,u(x),\gamma) &= v(x) \end{aligned}$$

für alle $x \in [a,b]$. Entlang des Graphen von u ist also S_γ konstant.

Sei umgekehrt $S_{\gamma u}(x,u,\gamma) \neq 0$ für alle $(x,u,\gamma) \in [a,b] \times \mathbb{R} \times \Gamma$. Sei $\gamma \in \Gamma$, sei $c \in \mathbb{R}$ und sei $u \in C_2([a,b])$ mit

$$S_\gamma(x,u(x),\gamma) = c$$

für alle $x \in [a,b]$. Setze

$$v(x) = S_u(x,u(x),\gamma) \, .$$

Dann ist (u,v) eine Lösung von (H).

Beweis: Sei (u,v) eine Lösung von (H). Dann gilt

$$\begin{aligned} \frac{d}{dx} S_u(x,u(x),\gamma) &= S_{ux} + u' S_{uu} \\ &= S_{ux} + H_v(x,u(x),v(x)) S_{uu} \, . \end{aligned}$$

Es gilt

$$\begin{aligned} S_{ux}(x,u(x),\gamma) &= S_{xu}(x,u(x),\gamma) \\ &= \frac{\partial}{\partial u} S_x(x,u,\gamma)|_{u=u(x)} \\ &= - \frac{\partial}{\partial u} H\Big(x,u(x),S_u(x,u(x),\gamma)\Big) \\ &= - H_u\Big(x,u(x),S_u(x,u(x),\gamma)\Big) \\ &\quad - H_v\Big(x,u(x),S_u(x,u(x),\gamma)\Big) S_{uu} \, . \end{aligned}$$

Kombination mit der vorangehenden Gleichung ergibt

$$\begin{aligned} \frac{d}{dx} S_u(x,u(x),\gamma) = &- H_u\Big(x,u(x),S_u(x,u(x),\gamma)\Big) \\ &+ \Big[H_v(x,u(x),v(x)) - H_v\Big(x,u(x),S_u(x,u(x),\gamma)\Big)\Big] \, S_{uu} \, . \end{aligned}$$

Dies kann man auffassen als Differentialgleichung erster Ordnung für die Funktion $w(x) = S_u(x,u(x),\gamma)$. $S_u(x,u(x),\gamma)$ erfüllt diese Differentialgleichung und die Anfangsbedingung $S_u(a,u(a),\gamma) = v(a)$. Wegen (H) ist auch

$v(x)$ eine Lösung dieses Anfangswertproblems. Die Lösung ist aber eindeutig, also folgt

$$v(x) = S_u(x, u(x), \gamma)\,.$$

Hiermit folgt

$$\begin{aligned}
\frac{d}{dx}\Big[S_\gamma(x, u(x), \gamma)\Big] &= S_{x\gamma} + S_{u\gamma} u' \\
&= S_{x\gamma} + H_v(x, u(x), v(x)) S_{u\gamma} \\
&= S_{x\gamma} + H_v\Big(x, u(x), S_u(x, u(x), \gamma)\Big) S_{u\gamma} \\
&= \frac{d}{d\gamma}\left[S_x(x, u(x), \gamma) + H\Big(x, u(x), S_u(x, u(x), \gamma)\Big)\right] \\
&= 0\,,
\end{aligned}$$

also folgt

$$S_\gamma(x, u(x), \gamma) = \text{ const}$$

für alle $x \in [a, b]$.
Sei umgekehrt $S_\gamma(x, u(x), \gamma) = c$ für $x \in [a, b]$. Es folgt dann

$$S_{x\gamma} + S_{u\gamma} u' = \frac{d}{dx}\Big[S_\gamma(x, u(x), \gamma)\Big] = 0\,.$$

Außerdem folgt

$$S_{x\gamma} + H_v S_{u\gamma} = \frac{d}{d\gamma}\left[S_x(u, u(x), \gamma) + H\Big(x, u, S_u(x, u(x), \gamma)\Big)\right] = 0\,.$$

Durch Kombination beider Gleichungen folgt

$$\left[H_v\Big(x, u(x), S_u(x, u(x), \gamma)\Big) - u'(x)\right] S_{u\gamma}(x, u(x), \gamma) = 0\,.$$

Wegen $v = S_u$ und wegen $S_{u\gamma} \neq 0$ folgt

$$u'(x) = H_v(x, u(x), v(x))\,.$$

Außerdem folgt

$$v'(x) = \frac{d}{dx}\Big[S_u(x, u(x), \gamma)\Big]$$

$$
\begin{aligned}
&= S_{xu} + S_{uu}u' \\
&= S_{xu} + S_{uu}H_v \\
&= [S_x + H(x,u,S_u)]_u - H_u \\
&= -H_u\Big(x, u(x), S_u(x, u(x), \gamma)\Big) \\
&= -H_u(x, u(x), v(x))\,.
\end{aligned}
$$

Also ist (u, v) eine Lösung von (H).

Dieser Satz erlaubt aus der Lösung der Hamilton–Jacobi Gleichung die Lösung des Hamiltonschen Differentialgleichungssystems zu bestimmen. Es bleibt aber die Frage offen, ob überhaupt eine Lösung der Hamilton–Jacobi Gleichung existiert. Dies ergibt sich aus folgendem Satz:

5.12 Satz (Existenz von Lösungen der Hamilton–Jacobi Gleichung). Mit den Bezeichnungen von Satz 5.7 sei $H \in C_3(W, \mathbb{R})$. Sei $S_0 \in C_2(\mathbb{R})$ eine Funktion mit

$$v_0(u) := S_0'(u) \in W(a, u)$$

für alle $u \in \mathbb{R}$. Dann gibt es eine in $[a, b] \times \mathbb{R}$ relativ offene Teilmenge U von $[a, b] \times \mathbb{R}$ mit $\{a\} \times \mathbb{R} \subseteq U$ und eine Lösung $S \in \mathbb{C}_2(U, \mathbb{R})$ der Hamilton–Jacobischen Differentialgleichung

$$S_x(x, u) + H(x, u, S_u(x, u)\Big) = 0$$

in U zur Anfangsbedingung

$$S(a, u) = S_0(u)\,.$$

Beweis: Aus der Theorie der gewöhnlichen Differentialgleichungen folgt, daß es eine in $[a, b] \times \mathbb{R}$ relativ offene Teilmenge U von $[a, b] \times \mathbb{R}$ gibt mit $\{a\} \times \mathbb{R} \subseteq U$, die in eineindeutiger Weise von den Kurven $x \mapsto (x, u, (x, \alpha))$ überdeckt wird, wobei $u(x, \alpha)$ die erste Komponente der Lösung $(u, v) = (u(x, \alpha), v(x, \alpha))$ von

$$
\begin{aligned}
u'(x, \alpha) &= H_v(x, u(x, \alpha), v(x, \alpha))\,, & u(a, \alpha) &= \alpha \\
v'(x, \alpha) &= -H_u(x, u(x, \alpha), v(x, \alpha))\,, & v(a, \alpha) &= v_0(\alpha)
\end{aligned}
$$

ist. Hierbei sei

$$u'(x,\alpha) = \frac{\partial}{\partial x}\, u(x,\alpha), \quad v'(x,\alpha) = \frac{\partial}{\partial x}\, v(x,\alpha)\,.$$

Es sei $V \subseteq [a,b] \times \mathbb{R}$ die Menge alle (x,α) mit $(x,u(x,\alpha)) \in U$. Wenn nötig durch Verkleinerung von U und V kann erreicht werden, daß U von der Form

$$U = \{(x,y) : a \le x < c(y)\}$$

ist mit $a < c(y) \le b$, und daß die Funktionaldeterminante der Abbildung $(x,\alpha) \mapsto (x,u(x,\alpha))$ in V von Null verschieden ist:

$$\det\left[\frac{\partial(x,u(x,\alpha))}{\partial(x,\alpha)}\right] = \begin{vmatrix} 1 & 0 \\ \frac{\partial u}{\partial x} & \frac{\partial u}{\partial \alpha} \end{vmatrix} = \frac{\partial u}{\partial \alpha}(x,\alpha) > 0\,,$$

wegen $\frac{\partial u}{\partial \alpha}(a,\alpha) = \frac{\partial \alpha}{\partial \alpha} = 1,$. Es existiert dann die Inverse $(x,u) \mapsto (x,\alpha(x,u)) : U \to V$ von $(x,\alpha) \mapsto (x,u(x,\alpha))$, die dieselbe Differenzierbarkeitsordnung hat wie $(x,\alpha) \mapsto u(x,\alpha)$. Wegen $H \in C_3$ ist $u(x,\alpha) \in C_2$, also auch $\alpha(x,u)$. Ich definiere die Abbildung $\tilde{v} : U \to \mathbb{R}$ durch

$$\tilde{v}(x,u) = v(x,\alpha(x,u))\,.$$

Sei nun S die Lösung von

$$(*) \qquad S_x(x,u) = -H(x,u,\tilde{v}(x,u))\,, \quad S(a,u) = S_0(u)\,,$$

in U , also

$$S(x,u) = S_0(u) - \int_0^x H(y,u,\tilde{v}(y,u))dy$$

für alle $(x,u) \in U$. Ich werde zeigen, daß

$$S_u(x,u) = \tilde{v}(x,u)$$

gilt. Wenn dies gezeigt ist, folgt direkt aus der Definition $(*)$ von S , daß S eine Lösung der Hamilton–Jacobischen Differentialgleichung zur Anfangsbedingung $S(a,u) = S_0(u)$ ist.

Zum Beweis beachte man, daß durch Einsetzen von $u = u(x,\alpha)$ in die Definition von $\tilde{v}$ folgt

$$v(x,\alpha) = \tilde{v}(x,u(x,\alpha))\,,$$

also

$$\begin{aligned}
\frac{\partial}{\partial x} v(x,\alpha) &= \frac{d}{dx}\tilde{v}(x,u(x,\alpha)) \\
&= \tilde{v}_x(x,u(x,\alpha)) + \tilde{v}_u(x,u(x,\alpha))\frac{\partial u}{\partial x}(x,\alpha) \\
&= \tilde{v}_x(x,u(x,\alpha)) + \tilde{v}_u(x,u(x,\alpha))u'(x,\alpha)\,,
\end{aligned}$$

d. h.

$$v_x(x,\alpha(x,u)) = \tilde{v}_x(x,u) + \tilde{v}_u(x,u)u'(x,\alpha(x,u))\,.$$

Aus $(*)$ ergibt sich somit

$$\begin{aligned}
S_{ux}(x,u) &= S_{xu}(x,u) \\
&= -\frac{d}{du}\Big[H(x,u,\tilde{v}(x,u))\Big] \\
&= -H_u(x,u,\tilde{v}(x,u)) - H_v(x,u,\tilde{v}(x,u))\tilde{v}_u(x,u) \\
&= v_x(x,\alpha(x,u)) - u_x(x,\alpha(x,u))\tilde{v}_u(x,u) \\
&= \tilde{v}_x(x,u) + \tilde{v}_u(x,u)u'(x,\alpha(x,u)) - u'(x,\alpha(x,u))\tilde{v}_u(x,u) \\
&= \tilde{v}_x(x,u)\,.
\end{aligned}$$

Wegen $(*)$ folgt hieraus durch Integration

$$\begin{aligned}
S_u(x,u) - \tilde{v}(x,u) &= S_u(a,u) - \tilde{v}(a,u) \\
&= S_0'(u) - v(a,u) \\
&= S_0'(u) - v_0(u) \\
&= 0\,.
\end{aligned}$$

Damit ist der Satz bewiesen.

5.13 Beispiel (Geometrische Optik, Eikonalgleichung). In 5.9.2 wurde gezeigt, daß die Hamiltonfunktion zum Fermatschen Prinzip des kürzesten Lichtweges durch

$$H(x,u,v) = -\sqrt{n(x,u)^2 - v^2}\,, \quad -n(x,u) < v < n(x,u)$$

gegeben ist. Die Hamilton–Jacobi Gleichung lautet in diesem Fall

$$S_x(x,u) - \sqrt{n(x,u)^2 - S_u(x,u)^2} = 0\,.$$

Aus Symmetriegründen ersetze ich die Bezeichnung u durch y. Dann kann die Hamilton–Jacobi Gleichung auch in der Form

$$S_x(x,y)^2 + S_y(x,y)^2 = n(x,y)^2$$

oder

$$|\nabla S(x,y)|^2 = n(x,y)^2$$

geschrieben werden. Diese Gleichung nennt man in der geometrischen Optik auch **Eikonalgleichung**.

Sei $(u(x,\alpha), v(x,\alpha))$ die Lösung von

$$\begin{aligned} u'(x,\alpha) &= H_v(x,u(x,\alpha),v(x,\alpha))\,, & u(a,\alpha) &= \alpha \\ v'(x,\alpha) &= -\,H_u(x,u(x,\alpha),v(x,\alpha))\,, & v(a,\alpha) &= v_0(\alpha) \end{aligned}$$

mit $v_0(y) = S_0'(y) \in W(a,y) = (-n(a,y), n(a,y))$. In 5.9.2 wurde gezeigt, daß $v(x,\alpha)$ die zweite Komponente eines Vektors V mit $|V| = n(x,y)$ ist. Wegen

$$S_y(x,y) = \tilde{v}(x,y) = v(x,\alpha(x,y))\,, \quad |\nabla S(x,y)| = n(x,y)$$

folgt

$$\nabla S(x,y) = V\,.$$

Es wurde auch gezeigt, daß der Tangentenvektor $(1, u'(x,\alpha))$ an die Kurve $x \mapsto (x, u(x,\alpha))$ die Richtung von V, also von $\nabla S(x, u(x,\alpha))$ hat. Die Kurven $x \mapsto (x, u(x,\alpha))$ sind also die Kurven des steilsten Anstiegs der Funktion $S(x,y)$, sie stehen senkrecht auf den Niveaulinien $S(x,y) = \text{const}$ von S.

Schreibt man zum Beispiel die Anfangsbedingung $S_0(y) = 0$ vor (also $v_0(y) = S_0'(y) = 0$), dann ist die Gerade $\{a\} \times \mathbb{R}$ selbst eine Niveaulinie von S. Die Kurven $x \mapsto (x, u(x,\alpha))$ beginnen also im Punkt (a,α) senkrecht zu dieser Geraden:

Zeichnung - Niveaulinien

In diesem Beispiel ist die Gerade $\{a\} \times \mathbb{R}$ eine "leuchtende Linie" . Die Kurven $x \mapsto (x, u(x, \alpha))$ sind die von dieser Linie ausgehenden Lichtstrahlen. Entlang der Lichtstrahlen, also jeweils in Richtung von ∇S , breitet sich das Licht mit der Geschwindigkeit $\frac{1}{n(x,y)} = \frac{1}{|\nabla S(x,y)|}$ aus. Daher ist die Niveaulinie $S(x, y) = t$ eine Linie, die alle Punkte auf den Lichtstrahlen verbindet, die vom Licht zum Zeitpunkt t erreicht werden. Wenn (x_1, y_1) und (x_2, y_2) zwei Punkte auf demselben Lichtstrahl sind, dann gibt die Differenz

$$S(x_2, y_2) - S(x_1, y_1)$$

die Zeit an, die das Licht entlang des Lichtstrahles vom Punkt (x_1, y_1) zum Punkt (x_2, y_2) benötigt.

6. Schwache Topologie

6.1 Definition (Konvexe Mengen). Sei V ein normierter Raum und $M \subseteq V$. Die Menge

$$\mathrm{conv\,M} = \{\sum_{\mathrm{i}\in\mathrm{N}} \lambda_{\mathrm{i}}\mathrm{x}_{\mathrm{i}} \mid \mathrm{N \text{ endliche Teilmenge von } \mathbb{N}\,,\ x_i \in M\,,\ \lambda_i \geq 0\,,\ } \sum_{\mathrm{i}\in\mathrm{N}} \lambda_{\mathrm{i}} = 1\,\}$$

heißt konvexe Hülle von M.

Ist M konvex, dann auch $\overline{M}$. Ist M offen, dann auch $\mathrm{conv\,M}$, wegen

$$\mathrm{conv\,M} = \bigcup_{\substack{\mathrm{I}\subseteq[0,1] \\ \mathrm{I\,endlich} \\ \sum_{\lambda\in\mathrm{I}}\lambda=1}} (\sum_{\lambda\in\mathrm{I}} \lambda\mathrm{M})\,.$$

6.2 Satz. Sei H ein Hilbertraum, und sei $M \subseteq H$ abgeschlossen und konvex, $M \neq \emptyset$. Dann existiert für jedes $x \in H$ ein eindeutiges $y \in M$ mit

$$\|x-y\| = \inf_{\eta\in M} \|x-\eta\|\,.$$

Beweis. Sei $\eta_k \in M$ mit $\lim_{k\to\infty} \|\eta_k - x\| = \inf_{\eta\in M} \|\eta - x\| = d$.
Die Parallelogrammgleichung

$$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2$$

ergibt mit $u = x - \eta_k$, $v = x - \eta_l$

$$(*) \qquad \|\eta_k - \eta_l\|^2 = 2\|x-\eta_k\|^2 + 2\|x-\eta_l\|^2 - 4\|x - \frac{1}{2}(\eta_k+\eta_l)\|^2\,.$$

Da M konvex ist, ist $\frac{1}{2}(\eta_k+\eta_l) \in M$, somit

$$d^2 \leq \|x - \frac{1}{2}(\eta_k+\eta_l)\|^2\,.$$

Zu jedem $\varepsilon > 0$ gibt es l_0 mit $\|\eta_l - x\| < d + \varepsilon$ für alle $l \geq l_0$. Aus $(*)$ folgt also für $0 < \varepsilon \leq 1$ und $k, k \geq l_0$

$$\begin{aligned} \|\eta_k - \eta_l\|^2 &\leq 2\|x-\eta_k\|^2 + 2\|x-\eta_l\|^2 - 4d^2 \\ &\leq 2(d+\varepsilon)^2 + 2(d+\varepsilon)^2 - 4d^2 \\ &= 8d\varepsilon + 4\varepsilon^2 \leq (8d+4)\varepsilon\,. \end{aligned}$$

Also ist $\{\eta_l\}_{l=1}^{\infty}$ eine Cauchyfolge und hat den Grenzwert y. Da M abgeschlossen ist, ist $y \in M$ und für dieses y gilt

$$\|x-y\| = \|x - \lim_{k\to\infty} \eta_k\| = \lim_{k\to\infty} \|x-\eta_k\| = d\,.$$

Zum Beweis des Satzes genügt es also zu zeigen, daß y eindeutig ist. Sei y' ein zweites Element aus M mit $\|x-y'\| = d$. Dies ergibt

$$\begin{aligned}\|y-y'\| &= 2\|x-y\|^2 + 2\|x-y'\|^2 - 4\|x - \frac{1}{2}(y+y')\|^2 \\ &\le 4d^2 - 4d^2 = 0\,,\end{aligned}$$

also $y = y'$, also ist y eindeutig bestimmt.

6.3 Folgerung. Sei S ein abgeschlossener Unterraum des Hilbertraumes $H, S \neq \emptyset$, sei $x \in H$ und sei $\delta = \inf\{\|x-y\| \mid y \in S\}$. Dann gibt es genau ein $y_0 \in S$ mit $\|x-y_0\| = \delta$.

Beweis: Ergibt sich aus dem vorangehenden Satz mit $M = S$.

6.4 Definition (Projektion auf konvexe Mengen). Sei H ein Hilbertraum und $M \subseteq H$ abgeschlossen und konvex. Für $x \in H$ bezeichne $\mathrm{Proj_M\,x}$ das nach dem letzten Satz eindeutig bestimmte Element aus M mit

$$\|\mathrm{Proj_M\,x - x}\| = \inf_{\eta\in \mathrm{M}} \|\mathrm{x} - \eta\|\,.$$

$\mathrm{Proj_M} : \mathrm{H} \to \mathrm{M}$ heißt die Projektion von H nach M.

6.5 Lemma. Sei $M \subseteq H$ eine abgeschlossene und konvexe Menge. Dann gilt $y = \mathrm{Proj_M\,x}$ genau dann, wenn $y \in M$ ist und

$$\mathrm{Re\,(y, \eta - y) \ge Re\,(x, \eta - y)}$$

für alle $\eta \in M$.

Beweis: Sei $x \in H$, $y = \mathrm{Proj_M\,x}$. Da M konvex ist, folgt

$$(1-t)y + t\eta = y + t(\eta - y) \in M$$

für alle $0 \le t \le 1$, folglich nimmt

$$\varphi(t) = \|x-y-t(\eta-y)\|^2 = \|x-y\|^2 - 2t\,\mathrm{Re\,(x-y, \eta-y) + t^2\|\eta - y\|^2}$$

das Minimum an für $t = 0$. Also gilt

$$0 \leq \varphi'(0) = -2\,\mathrm{Re}\,(\mathrm{x}-\mathrm{y}, \eta-\mathrm{y})\ ,$$

somit

$$\mathrm{Re}\,(\mathrm{y}, \eta-\mathrm{y}) \geq \mathrm{Re}\,(\mathrm{x}, \eta-\mathrm{y})\ .$$

Umgekehrt sei für $y \in M$ und für alle $\eta \in M$

$$\mathrm{Re}\,(\mathrm{y}, \eta-\mathrm{y}) \geq \mathrm{Re}\,(\mathrm{x}, \eta-\mathrm{y})\ .$$

Dies ergibt

$$0 \leq \mathrm{Re}\,(\mathrm{y}-\mathrm{x}\,,\,(\eta-\mathrm{x})+(\mathrm{x}-\mathrm{y})) = -\|\mathrm{x}-\mathrm{y}\|^2 + \mathrm{Re}\,(\mathrm{y}-\mathrm{x}, \eta-\mathrm{x})\ ,$$

also

$$\|x-y\|^2 \leq \mathrm{Re}\,(\mathrm{y}-\mathrm{x}, \eta-\mathrm{x}) \leq \|\mathrm{x}-\mathrm{y}\|\,\|\eta-\mathrm{x}\|\ ,$$

somit

$$\|x-y\| \leq \|\eta-x\|$$

für alle $\eta \in M$. Also ist $y = \mathrm{Proj}_{\mathrm{M}}\,\mathrm{x}$.

6.6 Folgerung. Sei H ein Hilbertraum und S ein abgeschlossener Unterraum von $H, S \neq \emptyset$. Dann gilt $y = \mathrm{Proj}_{\mathrm{S}}\,\mathrm{x}$, genau dann, wenn $y \in S$ mit

$$(x-y, z) = 0$$

für alle $z \in S$. (In diesem Fall ist $\mathrm{Proj}_{\mathrm{S}}$ also die orthogonale Projektion auf S .)

Beweis. Sei $(x-y, z) = 0$ für alle $z \in H$. Dann folgt für alle $\eta \in S$

$$\mathrm{Re}\,(\mathrm{y}-\mathrm{x}, \eta-\mathrm{x}) = 0 \geq 0\ ,$$

wegen $\eta - y \in S$, also $y = \mathrm{Proj}_{\mathrm{S}}\,\mathrm{x}$, nach Lemma 6.5. Sei umgekehrt $y = \mathrm{Proj}_{\mathrm{S}}\,\mathrm{x}$. Aus Lemma 6.5 folgt dann

$$\mathrm{Re}\,(\mathrm{y}-\mathrm{x}, \mathrm{z}) \geq 0$$

für alle $z \in S$, also

$$\mathrm{Re}\,(\lambda(\mathrm{y}-\mathrm{x}, \mathrm{z})) = \mathrm{Re}\,(\mathrm{y}-\mathrm{x}, \overline{\lambda}\mathrm{z}) \geq 0$$

für alle $\lambda \in \mathbb{K}$, somit $(y - x, z) = 0$ für $z \in S$.

6.7 Definition (Reell lineare Abbildungen). Sei V ein Vektorraum über $\mathbb{K}$. Eine Abbildung $f : V \to \mathbb{R}$ heißt reell linear, wenn für alle $x, x_1, x_2 \in V$ und alle $\lambda \in \mathbb{R}$ gilt

$$\begin{aligned} f(x_1 + x_2) &= f(x_1) + f(x_2) \\ f(\lambda x) &= \lambda f(x) . \end{aligned}$$

6.7.1 Bemerkung. Sei V ein normierter Raum über $\mathbb{K}$. Für eine reell lineare Abbildung f gilt genau wie für eine lineare Abbildung, daß sie stetig ist, genau dann, wenn sie beschränkt ist. D.h. die reell lineare Abbildung $f : V \to \mathbb{R}$ ist stetig, genau dann, wenn eine Konstante $C \geq 0$ existiert mit

$$|f(x)| \leq C \, \|x\|$$

für alle $x \in V$. Denn V ist auch ein normierter Vektorraum über $\mathbb{R}$, und für diesen normierten Vektorraum sind die reell linearen Abbildungen gerade die linearen Abbildungen.

6.8 Satz (Trennungssatz). Sei H ein Hilbertraum über $\mathbb{K}$, sei $M \subseteq H$ eine abgeschlossene, nichtleere, konvexe Teilmenge von H, und sei $x \in H \setminus M$. Dann gibt es eine stetige, reell lineare Abbildung $f : H \to \mathbb{R}$ mit

$$f(x) > \sup_{z \in M} f(z) .$$

Beweis: Sei $y = \mathrm{Proj_M\, x}$. Dann ist $f = H \to \mathbb{R}$ mit

$$f(z) = \mathrm{Re\,(x - y, z)}$$

eine reell lineare Abbildung mit

$$|f(z)| \leq |\mathrm{Re\,(x - y, z)}| \leq \mathrm{\|x - y\| \, \|z\|} ,$$

also ist f stetig. Außerdem gilt nach Lemma 6.5 für alle $\eta \in M$

$$\begin{aligned} f(x) &= \mathrm{Re\,(x - y, x) = (x - y, x - y) + Re\,(x - y, y)} \\ &= \|x - y\|^2 + \mathrm{Re\,[(y - x, \eta - y) + (x - y, \eta)]} \\ &\geq \|x - y\|^2 + \mathrm{Re\,(x - y, \eta) = \|x - y\|^2 + f(\eta)} , \end{aligned}$$

und somit

$$\sup_{\eta \in M} f(\eta) \leq f(x) - \|x-y\|^2 < f(x) ,$$

wegen $x \in H \backslash M$, $y \in M$, also $x \neq y$.

6.9 Folgerung. Sei H ein Hilbertraum. Dann ist jede nichtleere abgeschlossene, konvexe Teilmenge M von H der Durchschnitt aller abgeschlossenen Halbräume, die diese Menge enthalten.

(Wenn $f : H \to \mathbb{R}$ eine stetige, reell lineare Abbildung ist und $\alpha \in \mathbb{R}$, dann heißen die abgeschlossenen Mengen

$$\{x \in H \mid f(x) \geq \alpha\}$$

und

$$\{x \in H \mid f(x) \leq \alpha\}$$

abgeschlossene Halbräume.)

Beweis: Sei $x \in H \backslash M$. Nach Satz 6.8 gibt es $\alpha \in \mathbb{R}$ und eine stetige, reell lineare Abbildung $f : H \to \mathbb{R}$ mit $\sup_{z \in M} f(z) \leq \alpha < f(x)$, also ist x nicht Element des Durchschnitts aller abgeschlossenen Teilräume, die M enthalten.

6.10 Schwache Topologie. Sei V ein normierter Raum und V' der Dualraum, d.h. der Raum aller stetigen linearen Abbildungen $f : V \to \mathbb{K}$.

Für $f \in V'$ sei

$$U_f = \{x \in V \mid |f(x)| < 1\} .$$

Die schwache Topologie auf V wird folgendermaßen erklärt. Eine Menge $U \subseteq V$ ist nach Definition eine Umgebung von 0 (Nullumgebung), genau dann, wenn es endlich viele stetige lineare Abbildungen $f_1, \ldots, f_n \in V'$ gibt mit

$$\bigcap_{i=1}^{n} U_{f_i} \subseteq U .$$

Für beliebiges $x \in V$ ist $W \subseteq V$ eine Umgebung von x , genau dann, wenn es eine Nullumgebung U gibt mit

$$W = x + U = \{x + y \mid y \in U\} .$$

Das so definierte System von Umgebungen erfüllt die Axiome, die von einem System von Umgebungen verlangt werden, und definiert damit eine Topologie auf V. Die so definierte Topologie auf V hat außerdem noch folgende Eigenschaft: Versieht man V mit der schwachen Topologie, und versieht man $\mathbb{K}$ mit der üblichen Topologie, dann sind die Vektorraumaddition und Multiplikation

$$\begin{array}{cccccc} (x,y) & \to & x+y & : & V \times V & \to & V \\ (\lambda,x) & \to & \lambda x & : & \mathbb{K} \times V & \to & V \end{array}$$

stetige Abbildungen. Also ist die schwache Topologie auf V eine Vektorraumtopologie auf V, und V ist ein topologischer Vektorraum mit der schwachen Topologie.

Die schwache Topologie auf V ist die gröbste Vektorraumtopologie auf V, in der noch alle $f \in V'$ stetig sind. Umgekehrt ist jede lineare Abbildung $g : V \to \mathbb{K}$, die in der schwachen Topologie stetig ist, auch in der Normtopologie stetig, weil die Normtopologie feiner ist als die schwache Topologie, also stimmt der Dualraum V' zu V mit der Normtopologie überein mit dem Dualraum zu V mit der schwachen Topologie.

Nach dem Satz von Hahn-Banach gibt es zu jedem $x \in V$ mit $x \neq 0$ ein $f \in V'$ mit $|f(x)| > 1$, also ist $x \notin U_f$, folglich ist die schwache Topologie eine separierte Topologie, d.h. Grenzwerte sind eindeutig.

6.11 Satz. Sei V ein normierter Raum. Eine Folge $\{x_n\}_{n=1}^{\infty} \subseteq V$ konvergiert schwach gegen $x \in V$ (d.h. in der schwachen Topologie von V), genau dann, wenn für alle $f \in V'$ gilt

$$\lim_{n\to\infty} f(x_n) = f(x) .$$

Beweis. Wenn $\{x_n\}_{n=1}^{\infty}$ schwach gegen $x \in V$ konvergiert, gilt $\lim_{n\to\infty} f(x_n) = f(x)$, da jedes $f \in V'$ stetig ist in der schwachen Toplogie. Umgekehrt gelte $\lim_{n\to\infty} f(x_n) = f(x)$ für alle $f \in V'$. Sei $x+U$ eine Umgebung von x, wobei U eine Nullumgebung ist. Dann gibt es $f_1, \dots, f_m \in V'$ mit

$$\bigcap_{i=1}^{m} U_{f_i} = \{x \in V \mid \underset{i=1,\dots,m}{\forall} |f_i(x)| < 1\} \subseteq U .$$

Wegen $\lim_{n\to\infty} f_i(x_n) = f_i(x)$ für $i = 1, \dots, m$ gibt es k_i mit

$$|f_i(x_n - x)| = |f_i(x_n) - f_i(x)| < 1$$

für alle $n \geq k_i$, also auch für alle $n \geq k_0 = \max\{k_1, \dots, k_m\}$, somit

$$x_n - x \in \bigcap_{i=1}^{m} U_{f_i} \subseteq U$$

für alle $n \geq k_0$, also

$$x_n = x + (x_n - x) \in x + U \ , \quad n \geq u_0 \ .$$

Da $x+U$ eine beliebige Umgebung von x ist, bedeutet dies, daß $\lim_{n\to\infty} x_n = x$ gilt in der schwachen Topologie von V .

Wenn $\{x_n\}_{n=1}^{\infty}$ in der schwachen Topologie gegen x konvergiert, schreibt man auch $x_n \rightharpoonup x$. In einem Hilbertraum gilt $x_n \rightharpoonup x$, genau dann, wenn

$$\lim_{n\to\infty} (x_n, z) = (x, z)$$

gilt für alle $z \in H$, weil in einem Hilbertraum jede stetige lineare Abbildung $f \in H'$ dargestellt werden kann in der Form

$$f(x) = (x, z)$$

mit einem geeigneten $z = z(f) \in H$.

6.12 Lemma (Abschätzung der Norm des schwachen Grenzwertes).
Sei H ein Hilbertraum und $x_n \rightharpoonup x$.
(i) Dann gilt

$$\|x\| \leq \liminf_{n\to\infty} \|x_n\| \ .$$

(ii) Aus $\limsup\limits_{n\to\infty} \|x_n\| \leq \|x\|$ folgt $x_n \to x$.

Beweis: (i)

$$\|x_n - x\|^2 = \|x_n\|^2 - 2(x, x_n) + \|x\|^2$$

folgt wegen $(x_n, x) \to (x, x) = \|x\|^2$, daß

$$\begin{aligned} 0 \ &\leq \ \liminf_{n\to\infty} \|x_n - x\|^2 = \liminf_{n\to\infty} \|x_n\|^2 - 2 \lim_{n\to\infty} (x_n, x) + \|x\|^2 \\ &= \ \liminf_{n\to\infty} \|x_n\|^2 - \|x\|^2 \ , \end{aligned}$$

also $\|x\|^2 \leq \liminf_{n\to\infty} \|x_n\|^2$.

(ii) Es gilt

$$\begin{aligned}\limsup_{n\to\infty} \|x - x_n\|^2 &= \limsup_{n\to\infty} \|x_n\|^2 - 2\lim_{n\to\infty}(x_n, x) + \|x\|^2 \\ &= \limsup_{n\to\infty} \|x_n\|^2 - \|x\|^2 \leq 0\,,\end{aligned}$$

folglich

$$0 \leq \lim_{n\to\infty} \|x - x_n\|^2 \leq \limsup_{n\to\infty} \|x - x_n\|^2 \leq 0\,,$$

d.h. $x_n \to x$.

6.13 Satz (Abschluß konvexer Mengen in der schwachen und in der Normtopologie). Sei H ein Hilbertraum, und sei M eine konvexe Teilmenge. Dann stimmen der Abschluß von M in der Normtopologie und in der schwachen Topologie überein.

Beweis. Sei $\overline{M}^{\|\cdot\|}$ der Abschluß von M in der Normtopologie, und sei $\overline{M}^W$ der Abschluß von M in der schwachen Topologie. Da die Normtopologie feiner ist als die schwache Topologie, ist auch $\overline{M}^W$ in der Normtopologie abgeschlossen, also

$$\overline{M}^{\|\cdot\|} \subseteq \overline{M}^W .$$

Nach Folgerung 6.9 ist jede in der Normtopologie abgeschlossene konvexe Menge der Durchschnitt von (in der Normtopologie) abgeschlossenen Halbräumen.

Also ist $\overline{M}^{\|\cdot\|}$ der Durchschnitt von abgeschlossenen Halbräumen, da mit M auch $\overline{M}^{\|\cdot\|}$ konvex ist.

Es wird nun gezeigt, daß diese Halbräume auch in der schwachen Topologie abgeschlossen sind. Wenn dies gezeigt ist, folgt, daß $\overline{M}^{\|\cdot\|}$ der Durchschnitt von in der schwachen Topologie abgeschlossenen Mengen ist, also ist $\overline{M}^{\|\cdot\|}$ selbst schwach abgeschlossen, somit folgt

$$\overline{M}^W \subseteq \overline{M}^{\|\cdot\|}\,,$$

zusammen also $\overline{M}^W = \overline{M}^{\|\cdot\|}$.

Jeder abgeschlossene Halbraum ist nach Definition von der Form

$$\{x \in H \mid f(x) \leq \alpha\}$$

mit einer geeigneten stetigen, reell linearen Abbildung $f : H \to \mathbb{R}$ und mit geeignetem $\alpha \in \mathbb{R}$.

Ein solches f ist aber auch in der schwachen Topologie stetig, also ist der Halbraum auch in der schwachen Topologie abgeschlossen als Urbild einer abgeschlossenen Menge unter einer stetigen Abbildung.

Um zu zeigen, daß f schwach stetig ist, sei $\mathbb{K} = \mathbb{C}$, weil sonst nichts zu zeigen ist. Dann ist die Abbildung $g : H \to \mathbb{C}$,

$$g(x) = f(x) - if(ix)$$

eine stetige, (komplex) lineare Abbildung. Denn für $\lambda = \lambda_1 + i\lambda_2 \in \mathbb{C}$ mit $\lambda_1, \lambda_2 \in \mathbb{R}$ gilt

$$\begin{aligned} g((\lambda_1 + i\lambda_2)x) &= f((\lambda_1 + i\lambda_2)x) - if(i(\lambda_1 + i\lambda_2)x) \\ &= \lambda_1 f(x) + \lambda_2 f(ix) - \lambda_1 i f(ix) + \lambda_2 i f(x) \\ &= (\lambda_1 + i\lambda_2) f(x) - (\lambda_1 + i\lambda_2) i f(ix) \\ &= (\lambda_1 + i\lambda_2) g(x)\,. \end{aligned}$$

Somit ist $g \in H'$, also ist g auch in der schwachen Topologie stetig, also auch $f = \operatorname{Re} g$, als Hintereinanderausführung von g und der Projektion auf die reelle Achse.

6.14 Definition (Beschränkte Mengen). Sei V ein topologischer Vektorraum. Eine Menge $M \subseteq V$ heißt beschränkt, wenn es zu jeder Nullumgebung U eine Zahl $\lambda_0 > 0$ gibt mit $M \subseteq \lambda U$ für alle $\lambda \in \mathbb{K}$ mit $|\lambda| \geq \lambda_0$.

Als Beispiel betrachte man folgenden Spezialfall: Sei V normierter Raum: Eine Menge $M \subseteq V$ ist beschränkt, genau dann, wenn

$$\sup_{x \in M} \|x\| < \infty\,.$$

M ist beschränkt in der schwachen Topologie, genau dann, wenn für jedes $f \in V'$ gilt

$$\sup_{x \in M} |f(x)| < \infty\,.$$

6.15 Satz (Äquivalenz von schwacher und Normbeschränktheit). Sei V ein normierter Raum und $M \subseteq V$. Es gilt: M ist schwach beschränkt, genau dann, wenn M beschänkt ist.

Zum Beweis braucht man folgenden

6.15.1 Satz von Banach-Steinhaus. Sei X ein Banachraum, Y ein normierter Raum und F eine Menge von stetigen linearen Abbildungen von X nach Y mit

$$\sup_{T \in F} \|Tx\|_Y < \infty$$

für alle $x \in X$. Dann gilt

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

Hierbei ist die Norm von T wie üblich definiert:

$$\|T\| = \sup_{\|x\| \leq 1} \|Tx\|_Y .$$

Beweis von Satz 6.15: Wenn M beschränkt ist, folgt für alle $f \in V'$

$$\sup_{x \in M} |f(x)| \leq \sup_{x \in M} \|f\| \, \|x\| \leq \|f\| \sup_{x \in M} \|x\| < \infty ,$$

also ist M schwach beschränkt.

Sei umgekehrt M schwach beschränkt. V' ist ein Banachraum, und jedes $x \in V$ definiert eine lineare stetige Abbildung $J[x] \in V''$ durch

$$(J[x])(f) = f(x) \text{ mit } \|J[x]\|_{V''} = \|x\|_V .$$

Wenn M schwach beschränkt ist, bedeutet dies für alle $f \in V'$, daß

$$\sup_{x \in M} |(J[x])(f)| = \sup_{x \in M} |f(x)| < \infty .$$

Die Menge $\{J[x] \mid x \in M\}$ von stetigen linearen Abbildungen auf dem Banachraum V' ist also punktweise beschränkt. Aus dem Satz von Banach-Steinhaus folgt somit, daß

$$\sup_{x \in M} \|x\|_V = \sup_{x \in M} \|J[x]\|_{V''} < \infty$$

gilt, also ist M normbeschränkte Teilmenge von V.

6.16 Folgerung. Sei H ein Hilbertraum mit $\{x_n\}_{n=1}^{\infty}, \{y_n\}_{n=1}^{\infty} \subseteq H$. Es gelte

$$x_n \to x \in H\,, \quad y_n \rightharpoonup y \in H\,.$$

Dann gilt $(x_n, y_n) \to (x, y)$.

Beweis: Es gilt $(y_n, z) \to (y, z)$ für alle $z \in H$, also gilt für alle $z \in H$

$$\sup_{n\in\mathbb{N}} |(y_n, z)| < \infty\,,$$

folglich ist $\{y_n\}_{n=1}^{\infty}$ schwach beschränkt, und somit auch beschränkt:

$$\|y_n\| \leq C$$

für alle $n \in \mathbb{N}$. Also ergibt sich

$$\begin{aligned} |(x,y) - (x_n, y_n)| &= |(x,y) - (x_n - x, y_n) - (x, y_n)| \\ &\leq |(x,y) - (x, y_n) - (x_n - x, y_n)| \\ &\leq |(x,y) - (x, y_n)| + \|x_n - x\| C \to 0 \end{aligned}$$

für $n \to \infty$.

6.17 Satz (schwache Folgenkompaktheit der Einheitskugel). Sei H ein Hilbertraum. Dann ist die abgeschlossene Einheitskugel $\overline{B_1(0)} = \{x \in H \mid \|x\| \leq 1\}$ schwach folgenkompakt.

Beweis: Sei $\{x_n\}_{n=1}^{\infty} \subseteq \overline{B_1(0)}$. Es muß gezeigt werden, daß es eine Teilfolge $\{x_{n_k}\}_{k=1}^{\infty}$ und $x \in \overline{B_1(0)}$ gibt mit

$$(y, x_{n_k}) \to (y, x)$$

für alle $y \in H$. Hierzu sei

$$\hat{H} = \overline{\mathrm{span}\,\{\mathrm{x_n}\}_{\mathrm{n=1}}^{\infty}}\,.$$

Dies ist ein abgeschlossener Unterraum von H, also selbst ein Hilbertraum. Außerdem ist $\hat{H}$ separabel, das heißt, es gibt eine abzählbare dichte Teilmenge $\{y_l\}_{l=1}^{\infty}$ von $\hat{H}$.

Es genügt nun zu zeigen, daß es eine Teilfolge $\{x_{n_k}\}_{k=1}^{\infty}$ und $x \in \hat{H} \cap \overline{B_1(0)}$ gibt mit

$$(*) \qquad (z, x_{n_k}) \to (z, x)$$

für alle $z \in \hat{H}$. Denn sei $H^{\perp}$ der Orthogonalraum von $\hat{H}$. Dann kann jedes $y \in H$ zerlegt werden in

$$y = z + z^{\perp}$$

mit $z \in \hat{H}$ und $z^{\perp} \in H^{\perp}$. Aus $(*)$ folgt dann

$$\begin{aligned} &(y, x_{n_k}) = (z, x_{n_k}) + (z^{\perp}, x_{n_k}) = (z, x_{n_k}) \\ \longrightarrow \quad &(z, x) = (z, x) + (z^{\perp}, x) = (y, x)\,, \end{aligned}$$

wegen x_{n_k}, $x \in \hat{H}$.

Zum Beweis von $(*)$ wähle mit dem Diagonalverfahren eine Teilfolge $\{x_{n_k}\}_{k=1}^{\infty}$ so aus, daß die Folge

$$\{(y_l, x_{n_k})\}_{k=1}^{\infty}$$

in $\mathbb{K}$ konvergiert für alle l. Dann konvergiert auch die Folge

$$\{(y, x_{n_k})\}_{k=1}^{\infty}$$

für jedes $y \in Y = \mathrm{span}\,\{\mathrm{y_l}\}_{\mathrm{l=1}}^{\infty} \subseteq \hat{\mathrm{H}}$, wegen

$$y = \alpha_1 y_{l_1} + \ldots + \alpha_m y_{l_m}$$

für jedes $y \in Y$ mit geeigneten m und $\alpha_1, \ldots, \alpha_m \in \mathbb{K}$. Definiere eine Abbildung $f : Y \to \mathbb{K}$ durch

$$f(y) = \lim_{k \to \infty} (y, x_{n_k})\,.$$

Es gilt für $y, z \in Y$

$$\begin{aligned} f(y + z) &= \lim_{k \to \infty} (y + z, x_{n_k}) = \lim_{k \to \infty} (y, x_{n_k}) + \lim_{k \to \infty} (z, x_{n_k}) \\ &= f(y) + f(z)\,, \end{aligned}$$

und ebenso $f(\lambda y) = \lambda f(y)$, also ist f linear. Weiterhin gilt

$$\begin{aligned} |f(y)| &= |\lim_{k\to\infty}(y, x_{n_k})| \le \lim_{k\to\infty} |(y, x_{n_k})| \\ &\le \limsup_{k\to\infty} \|y\|\,\|x_{n_k}\| \le \|y\|\ , \end{aligned}$$

also, für $y, z \in Y$

$$|f(y) - f(z)| = |f(y-z)| \le \|y - z\|\ ,$$

d.h. f ist gleichmäßig stetig auf Y , und kann somit zu einer stetigen Abbildung g auf $\overline{Y} = \hat{H}$ fortgesetzt werden. g ist automatisch auch linear, also $g \in \hat{H}'$ mit $\|g\| \le 1$, und folglich existiert nach dem Rieszschen Darstellungssatz ein eindeutiges $x \in \hat{H}$ mit $f(z) = (z, x)$ für alle $z \in \hat{H}$ und mit $\|x\| = \|g\| \le 1$, also $x \in \overline{B_1(0)} \cap \hat{H}$. Sei nun $z \in \hat{H}$ beliebig und $\varepsilon > 0$. Wähle $y \in Y$ mit $\|z - y\| < \varepsilon/3$. Es folgt

$$\begin{aligned} |(z,x) - (z, x_{n_k})| &\le |(z,x) - (y,x)| + |(y,x) - (y, x_{n_k})| + |(y, x_{n_k}) - (z, x_{n_k})| \\ &\le \|z-y\|\,\|x\| + |f(y) - (y, x_{n_k})| + \|y - z\|\,\|x_{n_k}\| \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 \;=\; \varepsilon \end{aligned}$$

für $k \ge k_0$, $k_0 = k_0(\varepsilon)$ geeignet. Also folgt

$$\lim_{k\to\infty}(z, x_{n_k}) = (z, x)$$

für alle $z \in \hat{H}$. Damit ist $(*)$ bewiesen, und es folgt $x_{n_k} \rightharpoonup x$.

7. Konvexe Funktionale

In diesem Abschnitt sei H ein Hilbertraum.

7.1 Definition (Unterhalbstetige Funktionen). Sei $M \subseteq H, M \neq \emptyset$ und sei $f : M \to (-\infty, \infty]$. f heißt von unten halbstetig im Punkt $x \in M$, wenn für jede Folge $\{x_n\}_{n=1}^{\infty} \subseteq M$ mit $x_n \to x$ gilt

$$f(x) \leq \liminf_{n\to\infty} f(x_n)\,.$$

f heißt schwach von unten halbstetig im Punkt $x \in M$, wenn für jede Folge $\{x_n\}_{n=1}^{\infty} \subseteq M$ mit $x_n \rightharpoonup x$ gilt

$$f(x) \leq \liminf_{n\to\infty} f(x_n)\,.$$

f heißt (schwach) von unten halbstetig, wenn f in jedem Punkt von M (schwach) von unten halbstetig ist.

7.2 Lemma. (i) Seien $f_1, f_2 : M \to \mathbb{R}$ konvex. Dann ist $f_1 + f_2$ konvex. Ist f_1 strikt konvex und f_2 konvex, dann ist $f_1 + f_2$ strikt konvex.
(ii) Eine reell linear Abbildung $H \to \mathbb{R}$ ist konvex.

Beweis: klar.

7.3 Satz (Äquivalenz von Unterhalbstetigkeit und schwacher Unterhalbstetigkeit bei konvexen Funktionalen). Sei $M \subseteq H$ eine konvexe Menge und sei $f : M \to (-\infty, \infty]$ konvex. Dann gilt: f ist schwach von unten halbstetig genau dann wenn f von unten halbstetig ist.

Beweis: Wenn f schwach von unten halbstetig ist, ist f auch von unten halbstetig. Denn aus $x_n \to x$ folgt $x_n \rightharpoonup x$ und somit

$$f(x) \leq \liminf_{n\to\infty} f(x_n)\,.$$

Um die Umkehrung zu beweisen beachte man, daß f schwach von unten halbstetig ist falls die Menge

$$\{x \in M \mid f(x) \leq \varepsilon\}$$

für jedes $\varepsilon \in \mathbb{R}$ schwach abgeschlossen ist in der auf M durch die schwache Topologie induzierten Topologie. Denn sei $\{x_n\}_{n=1}^{\infty} \subseteq M$ mit $x_n \rightharpoonup x_0 \in M$ und sei $\{x_{n_k}\}_{k=1}^{\infty}$ eine Teilfolge mit

$$f(x_{n_k}) \to \liminf_{n\to\infty} f(x_n) = a\,.$$

Für jedes $\delta > 0$ gibt es dann ein m mit

$$f(x_{n_k}) \leq a + \delta$$

also

$$x_{n_k} \in \{x \in M \mid f(x) \leq a + \delta\}$$

für alle $k \geq m$. Wenn diese Menge schwach abgeschlossen ist, gehört folglich auch der schwache Grenzwert x_0 von $\{x_{n_k}\}_{k=1}^{\infty}$ zu dieser Menge, also

$$f(x_0) \leq a + \delta\,,$$

und somit

$$f(x_0) \leq a = \liminf_{n\to\infty} f(x_n)\,,$$

weil $\delta > 0$ beliebig war. Somit ist f schwach von unten halbstetig.

Um den Beweis von Satz 7.3 fertig zu machen, sei also f von unten halbstetig. Es genügt zu zeigen, daß die Menge $\{x \in M \mid f(x) \leq \varepsilon\}$ für jedes $\varepsilon \in \mathbb{R}$ schwach abgeschlossen ist. Da M und f konvex sind, ist auch diese Menge konvex wegen

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) \leq \varepsilon\,.$$

Konvexe Mengen sind nach Satz 6.13 schwach abgeschlossen, genau dann wenn sie abgeschlossen sind. Da f von unten halbstetig ist, ist $A_\varepsilon = \{x \in M \mid f(x) \leq \varepsilon\}$ aber abgeschlossen, also schwach abgeschlossen. Somit ist f schwach von unten halbstetig.
Zum Beweis sei $x_0 \in \overline{A_\varepsilon}$ und $\{x_n\}_{n=1}^{\infty} \subseteq A_\varepsilon$ mit $x_n \to x_0$. Wegen der Halbstetigkeit von f folgt dann

$$f(x_0) \leq \liminf_{n\to\infty} f(x_n) \leq \varepsilon\,,$$

also $x_0 \in A_\varepsilon$ und somit $A_\varepsilon = \overline{A_\varepsilon}$.

7.4 Beispiel (Konvexität der Norm). Sei H ein Hilbertraum mit Norm $\|\cdot\|$. Dann ist $x \mapsto \|x\|^2 : H \to \mathbb{R}$ stetig und strikt konvex. Denn für $0 < t < 1$ und $x \neq y$ gilt

$$\begin{aligned} \frac{d^2}{dt^2}\|tx + (1-t)y\|^2 &= \frac{d^2}{dt^2}\Big[t^2\|x\|^2 + 2t(1-t)\mathrm{Re}\,(\mathrm{x},\mathrm{y}) + (1-\mathrm{t})^2\|\mathrm{y}\|^2\Big] \\ &= 2\|x\|^2 - 4\mathrm{Re}\,(\mathrm{x},\mathrm{y}) + 2\|\mathrm{y}\|^2 = 2\|\mathrm{x}-\mathrm{y}\|^2 > 0\,. \end{aligned}$$

Somit ist die Abbildung strikt konvex. Da diese Abbildung stetig ist, ist sie auch von unten halbstetig, und somit schwach von unten halbstetig, d.h. für $x_n \rightharpoonup x$ gilt

$$\|x\|^2 \leq \liminf_{n\to\infty} \|x_n\|^2\,.$$

Dies wurde schon in Lemma 6.12 (i) gezeigt.

7.5 Definition (Koerzitivität). Eine Funktion $f : H \to (-\infty, \infty]$ heißt koerzitiv, wenn Konstanten $c_1 > 0, c_2 \geq 0$ existieren mit

$$f(x) \geq c_1\|x\|^2 - c_2\,.$$

7.6 Lemma. Sei H ein Hilbertraum, sei $f_1 : H \to (-\infty, \infty]$ koerzitiv. Sei $f_2 : H \to (-\infty, \infty]$ und sei $g : H \to \mathbb{R}$ eine stetige, reell lineare Abbildung mit

$$f_2(x) \geq g(x) - C$$

für alle $x \in H$. Dann ist $f_1 + f_2$ koerzitiv.

Beweis: Es gibt $c_1 > 0$ und $c_2 \in \mathbb{R}$ mit

$$\begin{aligned} f_1(x) + f_2(x) &\geq c_1\|x\|^2 - c_2 + g(x) - C \\ &= c_1\|x\|^2 + \|x\|g\Big(\frac{x}{\|x\|}\Big) - C - c_2 \\ &\geq c_1\|x\|^2 - \|x\| \sup_{\|y\|\leq 1} |g(y)| - C - c_2 \\ &\geq \frac{c_1}{2}\|x\|^2 + C_1 - C - c_2\,, \end{aligned}$$

wobei

$$C_1 = \min_{x\in H}\Big(\frac{c_1}{2}\|x\|^2 - \|x\| \sup_{\|y\|\le 1} |g(y)|\Big) = -\frac{1}{2c_1}\Big(\sup_{\|y\|\le 1} |g(y)|\Big)^2 .$$

Folglich ist $f_1 + f_2$ koerzitiv.

7.7 Satz (Existenz eines Minimums bei konvexen Funktionalen). Sei $f : H \to (-\infty, \infty]$ eine konvexe, von unten halbstetige, koerzitive Funktion. Dann gibt es ein $x_0 \in H$ mit

$$f(x_0) = \min_{x\in H} f(x) .$$

Wenn f strikt konvex ist, ist das Minimum eindeutig.

Beweis: Wenn $f \equiv \infty$ gilt, ist die Behauptung richtig. Also sei

$$\inf_{x\in H} f(x) = a < \infty .$$

Wegen der Koerzitivität gilt für alle $x \in H$ mit $\|x\|^2 \ge r = \frac{a+1+c_2}{c_1}$, daß

$$(*) \qquad f(x) \ge c_1\|x\|^2 - c_2 \ge a + 1 + c_2 - c_2 = a + 1 .$$

Wähle nun eine Folge $\{x_n\}_{n=1}^\infty \subseteq H$ mit

$$\lim_{n\to\infty} f(x_n) = a = \inf_{x\in H} f(x) .$$

Wenn $(*)$ existiert eine Zahl n_0 mit

$$\|x_n\| \le r ,$$

also

$$x_n \in \overline{B_r(0)} = \{x \in H \mid \|x\| \le r\}$$

für alle $n \ge n_0$. Nach Satz 6.17 ist $\overline{B_r(0)}$ schwach folgenkompakt, also gibt es $x_0 \in \overline{B_r(0)}$ und eine Teilfolge $\{x_{n_k}\}_{k=1}^\infty$ mit

$$x_{n_k} \rightharpoonup x_0 .$$

Da f konvex und von unten halbstetig ist, ist f nach Satz 7.3 auch schwach von unten halbstetig, also gilt

$$\begin{aligned} f(x_0) &\leq \liminf_{k\to\infty} f(x_{n_k}) = \lim_{k\to\infty} f(x_{n_k}) \\ &= \lim_{n\to\infty} f(x_n) = \inf_{x\in H} f(x)\,, \end{aligned}$$

und somit

$$f(x_0) = \min_{x\in H} f(x)\,.$$

Um zu zeigen, daß das Minimum eindeutig ist, soll angenommen werden, daß $x_0, x_0' \in H$ existierten mit $x_0 \neq x_0'$ und mit

$$f(x_0) = f(x_0') = \min_{x\in H} f(x)\,.$$

Dann folgte für alle $t \in (0,1)$

$$f(tx_0 + (1-t)x_0') < tf(x_0) + (1-t)f(x_0') = \min_{x\in H} f(x)\,.$$

Dies ist ein Widerspruch, also kann das Minimum nur in einem Punkt angenommen werden.

7.8 Lemma (Beschränktheit konvexer, unterhalbstetiger Funktionale). Sei $f : H \to (-\infty, \infty]$ eine konvexe, von unten halbstetige Funktion. Dann existiert ein reell lineares, stetiges Funktional $g : H \to \mathbb{R}$ und eine Konstante $c_0 \in \mathbb{R}$ mit

$$f(x) > g(x) + c_0$$

für alle $x \in H$.

Beweis: Falls $f \equiv \infty$ ist, ist die Behauptung klar. Also genügt es den Fall zu betrachten, wo ein $x_0 \in H$ existiert mit $f(x_0) < \infty$. Beachte zunächst, daß der Raum $H \times \mathbb{K}$ ein Hilbertraum ist mit dem Skalarprodukt

$$((x,\lambda)\,,\,(y,\mu))_{H\times\mathbb{K}} = (x,y)_H + \lambda\overline{\mu}$$

und der Norm

$$\|(x,\lambda)\|_{H\times\mathbb{K}} = \sqrt{\|x\|_H^2 + |\lambda|^2}\,.$$

Sei nun
$$K(f) = \{(x,\lambda) \in H \times \mathbb{R} \mid f(x) \leq \lambda\} \subseteq H \times \mathbb{K}\,.$$
Dies ist eine konvexe, abgeschlossene, nichtleere Menge. Denn seien (x,λ) und $(y,\mu) \in K(f)$. Dann folgt für $0 < t < 1$

$$\begin{aligned} f(tx + (1-t)y) &\leq tf(x) + (1-t)f(y) \\ &\leq t\lambda + (1-t)\mu\,, \end{aligned}$$

also ist

$$t(x,\lambda) + (1-t)(y,\mu) = (tx + (1-t)y, t\lambda + (1-t)\mu) \in K(f)\,,$$

und somit ist $K(f)$ konvex.
Um zu sehen, daß $K(f)$ abgeschlossen ist, sei $(x,\lambda) \in \overline{K(f)}$ und $\{(x_m,\lambda_m)\}_{m=1}^{\infty} \subseteq K(f)$ eine Folge mit $\lim_{m\to\infty}(x_m,\lambda_m) = (x,\lambda)$. Da f von unten halbstetig ist, folgt

$$\begin{aligned} f(x) &\leq \liminf_{m\to\infty} f(x_m) \leq \liminf_{m\to\infty} \lambda_m \\ &= \lim_{m\to\infty} \lambda_m = \lambda\,, \end{aligned}$$

also $(x,\lambda) \in K(f)$; folglich ist $K(f)$ abgeschlossen.
Sei nun $\mu \in \mathbb{R}$ mit $\mu < f(x_0)$. Dann ist $(x_0,\mu) \notin K(f)$, also gibt es nach Satz 6.8 eine stetige, reel lineare Abbildung $G : H \times \mathbb{K} \to \mathbb{R}$ mit

$$(*) \qquad G((x_0,\mu)) > \sup_{(x,\lambda)\in K(f)} G((x,\lambda))\,.$$

Definiere die Abbildung $h : H \to \mathbb{R}$ durch

$$h(x) = G((x,0))\,.$$

Dann ist h reell linear und stetig. Auch die Abbildung $\lambda \mapsto G((0,\lambda)) : \mathbb{K} \to \mathbb{R}$ ist reell linear und stetig, also gibt es ein $\omega \in \mathbb{K}$ mit

$$G((0,\lambda)) = \mathrm{Re}\,(\lambda,\omega)_{\mathbb{K}} = \mathrm{Re}\,(\lambda\overline{\omega})$$

für alle $\lambda \in \mathbb{K}$. Zum Beweis beachte man, daß für jede stetige reell lineare Abbildung u gilt
$$u = \mathrm{Re}\,\hat{\mathrm{u}}$$

mit einer stetigen linearen Abbildung $\hat{u}$. Dies wurde im Beweis von Satz 6.13 gezeigt. Also gilt für alle $(x, \lambda) \in H \times \mathbb{K}$

$$G((x,\lambda)) = G((x,0)) + G((0,\lambda)) = h(x) + \mathrm{Re}\,(\lambda\overline{\omega})\,,$$

und wegen $\mu \in \mathbb{R}$ kann somit $(*)$ auch in der Form

$$G((x_0,\mu)) = h(x_0) + \mu(\mathrm{Re}\,\overline{\omega}) > \sup_{\mathrm{x}\in\mathrm{H}}\Big[\sup_{\lambda\geq \mathrm{f(x)}}(\mathrm{h(x)} + \lambda(\mathrm{Re}\,\overline{\omega}))\Big]$$

geschrieben werden. Insbesondere folgt hieraus

$$G((x_0,\mu)) = h(x_0) + (\mathrm{Re}\,\overline{\omega})\mu > \mathrm{h(x_0)} + (\mathrm{Re}\,\overline{\omega})\sup_{\lambda\geq \mathrm{f(x_0)}}\lambda\,,$$

was wegen $f(x_0) < \infty$ nur sein kann wenn $(\mathrm{Re}\,\overline{\omega}) < 0$. Außerdem folgt

$$G((x_0,\mu)) > h(x) + f(x)(\mathrm{Re}\,\overline{\omega})$$

für alle $x \in H$, d.h.

$$f(x) > -\frac{1}{\mathrm{Re}\,\overline{\omega}}h(x) + \frac{1}{\mathrm{Re}\,\overline{\omega}}G((x_0,\mu)) = g(x) + c_0$$

mit $g(x) = -\frac{1}{\mathrm{Re}\,\overline{\omega}}h(x)$ und

$$c_0 = \frac{1}{\mathrm{Re}\,\overline{\omega}}G((x_0,\mu))\,.$$

7.9 Folgerung. Sei H ein Hilbertraum und $f : H \to (-\infty, \infty]$ eine konvexe, von unten halbstetige Funktion. Dann ist die Funktion

$$x \mapsto f(x) + \frac{\alpha}{2}\|x\|^2$$

für jedes $\alpha > 0$ strikt konvex, von unten halbstetig und koerzitiv. Also gibt es ein eindeutiges $x_0 \in H$ mit

$$f(x_0) + \frac{\alpha}{2}\|x_0\|^2 = \min_{x\in H}\big(f(x) + \frac{\alpha}{2}\|x\|^2\big)\,.$$

Beweis: $f(x) + \frac{\alpha}{2}\|x\|^2$ ist von unten halbstetig als Summe einer von unten halbstetigen und einer stetigen Funktion. Im Beispiel 7.4 wurde gezeigt, daß die Abbildung $x \mapsto \|x\|^2$ strikt konvex ist. Also ist $\frac{\alpha}{2}\|x\|^2$ strikt konvex, folglich nach Lemma 7.2 auch $f(x) + \frac{\alpha}{2}\|x\|^2$. Nach Lemma 7.8 gibt es eine Konstante $c_0 \in \mathbb{R}$ und ein reell lineares, stetiges Funktional $g : H \to \mathbb{R}$ mit $f(x) > g(x) + c_0$ für alle $x \in H$. Da $x \mapsto \frac{\alpha}{2}\|x\|^2$ koerzitiv ist, ist also auch $f(x) + \frac{\alpha}{2}\|x\|^2$ koerzitiv nach Lemma 7.6. Der Rest der Behauptung ergibt sich nun aus Satz 7.7.

Für den Rest dieses Kapitels sei H immer ein reeller Hilbertraum. Dies ist keine Einschränkung der Allgemeinheit, weil jeder Hilbertraum mit dem Skalarprodukt $\mathrm{Re}\,(\cdot,\cdot)$ zu einem reellen Hilbertraum wird.

7.10 Definition (Subdifferential). Seien $M \subseteq H$ und $f : M \to (-\infty, \infty]$. Das Subdifferential ∂f von f ist eine Abbildung $\partial f : M \to \mathcal{P}(H)(=$ Potenzmenge von $H)$, die folgendermaßen definiert ist

$$y \in \partial f(x) \Longleftrightarrow \forall \xi \in M : f(\xi) \geq (y, \xi - x) + f(x)\,.$$

7.11 Beispiele für Subdifferentiale.
a) Sei $H = \mathbb{R}$ und sei $f : \mathbb{R} \to \mathbb{R}$ definiert durch

$$f(x) = \begin{cases} 0, & x \leq 0 \\ 1, & 0 < x \leq 1 \\ \infty, & x > 1 \end{cases}$$

Dann gilt

$$\partial f(x) = \begin{cases} \{0\}\,, & x < 0 \\ \{y \in \mathbb{R} \mid 0 \leq y \leq 1\}\,, & x = 0 \\ \emptyset, & 0 < x < 1 \\ \{y \in \mathbb{R} \mid 1 \leq y\}, & x = 1 \\ \emptyset, & x > 1\,. \end{cases}$$

b) Sei H ein reeller Hilbertraum, sei $M \subseteq H$ eine offene und konvexe Menge und sei $f : M \to (-\infty, \infty]$ ein konvexes Funktional. Sei $x \in M$. Dann gilt: Existiert grad $f(x)$, dann ist

$$\partial f(x) = \{\text{grad } f(x)\} .$$

Beweis: Es gilt für alle $z \in M$ und alle $0 \leq t \leq 1$

$$tf(z) + (1-t)f(x) \geq f((1-t)x + tz) = f(x + t(z-x)) .$$

Wenn grad $f(x)$ existiert, folgt hieraus

$$\begin{aligned} f(z) - f(x) &\geq \lim_{t \to 0+} \frac{1}{t}\Big[f(x + t(z-x)) - f(x)\Big] \\ &= (\text{grad } f(x), z - x) , \end{aligned}$$

also grad $f(x) \in \partial f(x)$. Angenommen, es sei $y \in \partial f(x)$ und $y \neq$ grad $f(x)$. Dann existiert $h \in H$ mit

$$(\text{grad } f(x), h) < (y, h) .$$

$\|h\|$ kann so klein gewählt werden, daß $x+th \in M$ ist für alle t mit $0 \leq t \leq 1$, weil M offen ist. Also folgt

$$\begin{gathered} f(th + x) - f(x) = t\Big((\text{grad } f(x), h) + o(1)\Big) \\ < t(y, h) = (y, th + x - x) \leq f(th + x) - f(x) \end{gathered}$$

für genügend kleines $t > 0$. Dies ist ein Widerspruch, also folgt
$y =$ grad $f(x)$, also

$$\partial f(x) = \{\text{grad } f(x)\} .$$

c) In Beispiel 7.4 wurde gezeigt, daß $f(x) = \|x\|^2$ konvex ist. In einem reellen Hilbertraum gilt

$$\begin{aligned} (\text{grad } f(x), h) &= \lim_{t \to 0} \frac{1}{t}[f(x + th) - f(x)] \\ &= \lim_{t \to 0} \frac{1}{t}(\|x + th\|^2 - \|x\|^2) \\ &= \lim_{t \to 0} (2(x, h) + t\|h\|^2) \;=\; 2(x, h) , \end{aligned}$$

also grad $f(x) = 2x$, und somit

$$\partial f(x) = \{2x\},$$

also, wenn x mit der Menge $\{x\}$ identifiziert wird,

$$\partial f = 2I$$

(I = Identität).
d) Sei $g : H \to \mathbb{R}$ eine stetige lineare Abbildung. Nach dem Rieszschen Darstellungssatz gibt es dann ein eindeutiges $y \in H$ mit

$$g(x) = (y, x)$$

für alle $x \in H$, und es gilt

$$\begin{aligned}(\text{grad } g(x), h) &= \lim_{t \to 0} \frac{1}{t}(g(x + th) - g(x)) \\ &= \lim_{t \to 0} \frac{1}{t}(y, th) = (y, h),\end{aligned}$$

also $\partial g(x) = \{y\}$.

7.12 Definition (Wertebereich des Subdifferentials). Sei $f : H \to (-\infty, \infty]$ eine konvexe Funktion und $\alpha \geq 0$. Die Menge

$$R(\alpha I + \partial f) = \bigcup_{x \in H} [\alpha x + \partial f(x)]$$

heißt Wertebereich der Abbildung $\alpha I + \partial f : H \to \mathcal{P}(H)$, die durch

$$(\alpha I + \partial f)(x) = \alpha x + \partial f(x) = \{\alpha x + y \mid y \in \partial f(x)\}$$

definiert ist.

In den nächsten Abschnitten wird gezeigt, daß es zu vielen elliptischen Differentialoperatoren der Form

$$\sum_{i=1}^{n} \frac{\partial}{\partial x_i} a_i\Big(\nabla u(x)\Big)$$

in Sobolevräumen eine konvexe Funktion f gibt mit

$$\partial f(u) = \left\{ -\sum_{i=1}^{n} \frac{\partial}{\partial x_i} a_i\Big(\nabla u(x)\Big) \right\}.$$

Die Gleichung

$$-\sum_{i=1}^{n} \frac{\partial}{\partial x_i} a_i\Big(\nabla u(x)\Big) + \alpha u(x) = g(x)$$

kann also für jede beliebige gegebene Funktion g aus dem Hilbertraum H (meistens ist $H = L^2(\Omega)$) gelöst werden, genau dann wenn $R(\alpha I + \partial f) = H$ gilt. Um die Frage zu untersuchen, wann dies gilt, benötigen wir folgendes

7.13 Lemma. Sei $f : H \to (-\infty, \infty]$ eine konvexe Funktion und $g : H \to \mathbb{R}$ eine lineare stetige Funktion mit

$$g(x) = (y, x)$$

für alle $x \in H$. Dann nimmt $f + g$ das Minimum an an der Stelle x_0 , genau dann wenn $-y \in \partial f(x_0)$. Insbesondere nimmt f das Minimum an an der Stelle x_0 , genau dann wenn $0 \in \partial f(x_0)$.

Beweis: $-y \in \partial f(x_0)$ ist äquivalent zu

$$f(x) \geq (-y, x - x_0) + f(x_0) \text{ für alle } x \in H\,,$$

und dies ist äquivalent zu

$$f(x) + g(x) = f(x) + (y, x) \geq f(x_0) + (y, x_0) = f(x_0) + g(x_0)$$

für alle $x \in H$, und dies ist schließlich äquivalent zu

$$f(x_0) + g(x_0) = \min_{x \in H} \left(f(x) + g(x) \right).$$

7.14 Satz (Surjektivität und Injektivität des Subdifferentials). Sei $f : H \to (-\infty, \infty]$ eine konvexe, von unten halbstetige, koerzitive Funktion. Dann gilt

$$R(\partial f) = H\,.$$

Wenn f strikt konvex ist, dann ist ∂f injektiv, d.h. zu jedem $y \in H$ gibt es genau ein $x \in H$ mit

$$y \in \partial f(x) .$$

Beweis: Sei $y \in H$. Um zu zeigen, daß $y \in R(\partial f)$ ist, betrachte die stetige lineare Abbildung $g : H \to \mathbb{R}$ mit

$$g(x) = (-y, x) .$$

Die Funktion $f + g$ ist von unten halbstetig als Summe einer stetigen und einer von unten halbstetigen Funktion. Nach Lemma 7.2 ist $f + g$ konvex beziehungsweise strikt konvex, und nach Lemma 7.6 ist $f + g$ koerzitiv. Nach Satz 7.7 gibt es also ein $x_0 \in H$ mit

$$f(x_0) + g(x_0) = \min_{x \in H} (f(x) + g(x)) ,$$

und nach Lemma 7.13 ist dies äquivalent zu

$$y \in \partial f(x_0) ,$$

also ist $R(\partial f) = H$. Wenn $f + g$ strikt konvex ist, nimmt nach Satz 7.7 $f + g$ das Minimum an in genau einem Punkt $x_0 \in H$, also gilt $y \in \partial f(x_0)$ nur genau für dieses x_0 , also ist ∂f injektiv.

7.15 Folgerung Sei $f : H \to (-\infty, \infty]$ eine konvexe, von unten halbstetige Funktion. Dann gilt für alle $\alpha > 0$, daß

$$R(\alpha I + \partial f) = H ,$$

und $\alpha I + \partial f$ ist injektiv.

Beweis: Sei $g(x) = f(x) + \frac{\alpha}{2} \|x\|^2$. Nach Beispiel 7.11 c) ist

$$\partial(\|x\|^2) \ = \{2x\} ,$$

also folgt für alle $y \in H$ und alle $z \in \partial f(x)$

$$\begin{aligned} g(y) &= f(y) + \frac{\alpha}{2}\|y\|^2 \geq (z, y-x) + f(x) + \frac{\alpha}{2}[(2x, y-x) + \|x\|^2] \\ &= (z + \alpha x, y - x) + f(x) + \frac{\alpha}{2}\|x\|^2 = (z + \alpha x, y - x) + g(x)\,, \end{aligned}$$

also $z + \alpha x \in \partial g(x)$, und somit $\alpha x + \partial f(x) \subseteq \partial g(x)$. Es gilt auch $\partial g(x) \subseteq \alpha x + \partial f(x)$, also $\partial g(x) = \alpha x + \partial f(x)$. Denn angenommen, es gäbe ein $z \in \partial g(x)$ mit $z - \alpha x \notin \partial f(x)$. Dann existierte ein $h \in H, h \neq 0$, und ein $\varepsilon > 0$ mit $f(x+h) \leq (z - \alpha x, h) + f(x) - \varepsilon$. Aus der Konvexität von f folgte dann für alle $0 < t < 1$

$$\begin{aligned} f(x + th) &= f(t(x+h) + (1-t)x) \leq tf(x+h) + (1-t)f(x) \\ &\leq t(z - \alpha x, h) + tf(x) - t\varepsilon + (1-t)f(x) \\ &= t(z - \alpha x, h) - t\varepsilon + f(x)\,, \end{aligned}$$

also

$$\begin{aligned} g(x + th) &= f(x + th) + \frac{\alpha}{2}\|x + th\|^2 \\ &\leq t(z - \alpha x, h) + f(x) - t\varepsilon + \frac{\alpha}{2}(\|x\|^2 + 2t(x, h) + t^2\|h\|^2) \\ &= t(z, h) + g(x) - t\varepsilon + t^2\frac{\alpha}{2}\|h\|^2 \\ &= (z, th) + g(x) - t(\varepsilon - t\frac{\alpha}{2}\|h\|^2) < (z, th) + g(x) \end{aligned}$$

für hinreichend kleines $t > 0$. Dies ist ein Widerspruch zu $z \in \partial g(x)$, also ist $\partial g(x) = \alpha x + \partial f(x)$, und somit

$$\partial g = \alpha I + \partial f\,.$$

Nach Folgerung 7.9 ist $g = f + \frac{\alpha}{2}\|x\|^2$ strikt konvex, von unten halbstetig und koerzitiv, also ergibt Satz 7.14 daß $R(\partial g) = R(\alpha I + \partial f) = H$, und daß $\alpha I + \partial f$ injektiv ist.

8. Direkte Methoden.

In diesem Abschnitt werde ich die Ergebnisse der beiden vorangehenden Abschnitte benützen, um die Existenz der Lösung von Variationsproblemen zu beweisen.
Die dabei angewandten Beweismethoden heißen direkte Methoden der Variationsrechnung. Ich studiere das folgende Variationsproblem im $\mathbb{R}^n$: Sei $\Omega \subseteq \mathbb{R}^n$, sei b eine nichtnegative Zahl oder sei $b = +\infty$, und sei $F : \mathbb{R}^n \to \mathbb{R}$. Gesucht ist eine Funktion $u : \overline{\Omega} \to \mathbb{R}$, die für alle Punkte x des Randes $\partial\Omega$ von Ω die Bedingung

$$0 \leq u(x) \leq b$$

erfüllt, und für die das Integral

$$\int_\Omega F(\nabla u(x))dx$$

unter allen derartigen Funktionen einen Minimalwert annimmt. Diese Randbedingung heißt "Hindernisrandbedingung"und ist allgemeiner als die früher betrachtete "Dirichletsche Randbedingung". Für $b = 0$ erhält man jedoch die homogene Dirichletsche Randbedingung

$$u(x) = 0\,,\ x \in \partial\Omega\,.$$

Der Einfachheit halber betrachte ich nur Integranden, die nicht explizit von x und $u(x)$ abhängen. Es ist aber möglich, die folgenden Beweise auf Integranden der Form $F(x, u(x), \nabla u(x))$ zu verallgemeinern.

Bevor ich den allgemeinen Existenzsatz formulieren kann, benötige ich einige Definitionen:
Wie im letzten Abschnitt werde ich in diesem Abschnitt nur reellwertige Funktionen betrachten. Sei

$$\tilde{M}(b) = \{u \in C_1(\overline{\Omega}) \cap H_1(\Omega) \mid \forall x \in \partial\Omega : 0 \leq u(x) \leq b\}\,,$$

und

$$M(b) = \overline{\tilde{M}(b)}^{H_1(\Omega)}\,,$$

wobei $C_1(\overline{\Omega})$ die Menge aller Funktionen $u \in C_1(\Omega)$ ist, die selbst und deren erste Ableitungen stetig auf $\overline{\Omega}$ fortgesetzt werden können. Die Menge $M(b)$

ist eine abgeschlossene und konvexe Teilmenge des Hilbertraumes $H_1(\Omega)$ und es gilt

$$\overset{\circ}{C}_\infty(\Omega) \subseteq \tilde{M}(b) \subseteq M(b)\,.$$

Zum Beweis genügt es zu zeigen, daß $\tilde{M}(b)$ konvex ist, weil der Abschluß einer konvexen Menge wieder konvex ist. Für $u, v \in \tilde{M}(b)$ und $0 \le t \le 1$ gilt aber $tu + (1-t)v \in C_1(\overline{\Omega}) \cap H_1(\Omega)$. Für $x \in \partial\Omega$ gilt außerdem

$$0 \le tu(x) + (1-t)v(x) \le tb + (1-t)b = b\,,$$

also $tu + (1-t)v \in \tilde{M}(b)$, und somit ist $\tilde{M}(b)$ konvex.

8.1 Definition (strikte Koerzitivität). Sei $f : \mathbb{R}^n \to (-\infty, \infty]$. f heißt strikt koerzitiv, wenn eine Konstante $c > 0$ existiert mit

$$f(x) \ge c\,|x|^2$$

für alle $x \in \mathbb{R}^n$.

In Zukunft werde ich immer voraussetzen, daß $F : \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar, konvex und strikt koerzitiv ist. Insbesondere gilt dann $F(\xi) \ge 0$ für alle $\xi \in \mathbb{R}^n$. Sei

$$M(F, b) = \{u \in M(b) \mid \int_\Omega F(\nabla u(x))dx < \infty\}\,.$$

Es gilt $M(F, b) \subseteq M(b) \subseteq H_1(\Omega)$, und wenn Ω beschränkt ist oder wenn $F(0) = 0$ gilt, folgt $\overset{\circ}{C}_\infty(\Omega) \subseteq M(F, b)$. Wenn jedoch Ω unbeschränkt ist und $F(0) \ne 0$ gilt, kann $M(F, b) = \emptyset$ sein. Ich will diesen Fall ausschliessen und setze daher immer $M(F, b) \ne \emptyset$ voraus.

Die Funktionale $V : H_1(\Omega) \to (-\infty, \infty]$ und $\tilde{V} : L^2(\Omega) \to (-\infty, \infty]$ seien definiert durch

$$V(u) = \begin{cases} \displaystyle\int_\Omega F(\nabla u(x))dx\,, & u \in M(F, b) \\ +\infty\,, & u \in H_1(\Omega)\backslash M(F, b) \end{cases}$$

$$\tilde{V}(u) = \begin{cases} \displaystyle\int_\Omega F(\nabla u(x))dx\,, & u \in M(F,b) \\ +\infty\,, & u \in L^2(\Omega)\backslash M(F,b)\,. \end{cases}$$

8.2 Satz (Existenz eines Minimums für Variationsfunktionale). Sei $F : \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar, konvex, strikt koerzitiv und sei $M(F,b) \neq \emptyset$.
(i) Sei $\lambda > 0$. Dann gibt es zu jedem $f \in L^2(\Omega)$ ein eindeutiges $u \in M(F,b)$ mit

$$\tilde{V}(u) + \frac{\lambda}{2}\|u\|_\Omega^2 - (f,u)_\Omega = \min_{v\in L^2(\Omega)} \Big[\tilde{V}(v) + \frac{\lambda}{2}\|v\|_\Omega^2 - (f,v)_\Omega\Big]\,.$$

(ii) Sei $\Omega \subseteq \mathbb{R}^n$ offen und beschränkt. Dann gibt es zu jedem $f \in L^2(\Omega)$ ein $u \in M(F,b)$ mit

$$\tilde{V}(u) - (f,u)_\Omega = \min_{v\in L^2(\Omega)} \Big[\tilde{V}(v) - (f,v)_\Omega\Big]\,.$$

Ich beweise diesen Satz in mehreren Schritten. Zunächst benötige ich folgendes grundlegende Resultat:

8.3 Satz (Konvexität und Unterhalbstetigkeit von Variationsfunktionalen auf $\boldsymbol{H_1(\Omega)}$). Sei $F : \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar, konvex und strikt koerzitiv. Dann ist $M(F,b)$ konvex, V ist konvex, von unten halbstetig, und es existiert eine Konstante $c > 0$ mit

$$V(u) \geq c\,|u|_{1,\Omega}^2$$

für alle $u \in H_1(\Omega)$.

Beweis: Seien $v_1, v_2 \in M(F,b)$ und $t \in (0,1)$. Dann folgt

$$\begin{aligned}
&\int_\Omega F(\nabla(tv_1(x) + (1-t)v_2(x)))dx \\
&\quad = \int_\Omega F(t\nabla v_1(x) + (1-t)\nabla v_2(x))dx \leq \int_\Omega tF(\nabla v_1(x)) + (1-t)F(\nabla v_2(x))dx \\
&\quad = tV(v_1) + (1-t)V(v_2) < \infty\,.
\end{aligned}$$

Hieraus folgt zunächst, daß $tv_1 + (1-t)v_2 \in M(F,b)$ ist, also ist $M(F,b)$ konvex. Dann aber folgt auch

$$
\begin{aligned}
V(tv_1 + (1-t)v_2) &= \int_\Omega F(\nabla(tv_1(x) + (1-t)v_2(x)))dx \\
&\le tV_1(v_1) + (1-t)V(v_2)\,.
\end{aligned}
$$

Diese Ungleichung ist auch richtig, wenn wenigstens eine der Funktionen v_1, v_2 nicht in $M(F,b)$ enthalten ist, weil dann die rechte Seite den Wert $+\infty$ hat, also ist V konvex.

Wenn F strikt koerzitiv ist, gibt es eine Konstante $c > 0$ mit $F(\xi) \ge c\,|\xi|^2$. Also folgt

$$
V(u) = \int_\Omega F(\nabla u(x))dx \ge c \int_\Omega |\nabla u(x)|^2 dx = c\,|u|_{1,\Omega}^2\,.
$$

Also bleibt zu zeigen, daß V von unten halbstetig ist. Hierzu muß bewiesen werden, daß für $\{u_m\}_{m=1}^\infty \subseteq H_1(\Omega)$ und $u \in H_1(\Omega)$ mit $\|u - u_m\|_{1,\Omega} \to 0$ gilt

$$
V(u) \le \liminf_{m\to\infty} V(u_m)\,.
$$

Es genügt eine Folge $\{u_m\}_{m=1}^\infty$ mit $\liminf\limits_{m\to\infty} V(u_m) = \alpha < \infty$ zu betrachten, weil sonst nichts zu zeigen ist. Wähle eine Teilfolge $\{u'_m\}_{m=1}^\infty$ aus mit

$$
\lim_{m\to\infty} V(u'_m) = \alpha\,.
$$

Nach Definition von V folgt hieraus $u'_m \in M(F,b) \subseteq M(b)$ für alle genügend großen m. Da $\{u'_m\}_{m=1}^\infty$ in der Norm von $H_1(\Omega)$ gegen u konvergiert, und da $M(b)$ in $H_1(\Omega)$ abgeschlossen ist, folgt hieraus

$$
u \in M(b)\,.
$$

Da $\int_\Omega |\frac{\partial}{\partial x_i}u(x) - \frac{\partial}{\partial x_i}u'_m(x)|^2 dx \to 0$ gilt für $m \to \infty$ und für $i = 1, \ldots, n$ gibt es nach einem Satz aus der Lebesgueschen Integrationstheorie eine Teilfolge $\{u''_m\}_{m=1}^\infty$ von $\{u'_m\}_{m=1}^\infty$ mit

$$
\lim_{m\to\infty} \nabla u''_m(x) = \nabla u(x)
$$

für fast alle $x \in \Omega$. Nach Voraussetzung ist $F : \mathbb{R}^n \to \mathbb{R}$ differenzierbar, also insbesondere stetig. Daher folgt

$$
\lim_{m\to\infty} F(\nabla u''_m(x)) = F(\nabla u(x))
$$

für fast alle $x \in \Omega$. Nach Voraussetzung gilt außerdem $F(\nabla u''_m(x)) \geq 0$ und

$$\lim_{m\to\infty} \int_\Omega F(\nabla u''_m(x))dx = \lim_{m\to\infty} V(u''_m) = \lim_{m\to\infty} V(u_m) = \alpha \,.$$

Somit sind die Voraussetzungen des Lemmas von Fatou erfüllt, und es folgt

$$\begin{aligned}
\int_\Omega F(\nabla u(x))dx &= \int_\Omega \lim_{m\to\infty} F(\nabla u''_m(x))dx \\
&= \int_\Omega \liminf_{m\to\infty} F(\nabla u''_m(x))dx \leq \liminf_{m\to\infty} \int_\Omega F(\nabla u''_m(x))dx \\
&= \lim_{m\to\infty} \int_\Omega F(\nabla u''_m(x))dx = \alpha = \liminf_{m\to\infty} V(u_m) < \infty \,.
\end{aligned}$$

Dies bedeutet zunächst, daß $u \in M(F,b)$ ist. Dann aber folgt

$$V(u) = \int_\Omega F(\nabla u(x))dx \,,$$

und somit

$$V(u) \leq \liminf_{m\to\infty} V(u_m) \,,$$

folglich ist V von unten halbstetig.

Als nächstes brauche ich folgendes Resultat:

8.4 Lemma. Sei $M \subseteq H_1(\Omega)$ eine konvexe und bezüglich der $H_1(\Omega)$–Norm abgeschlossene Teilmenge, und sei

$$W : L^2(\Omega) \to (-\infty, \infty] \quad \text{mit } W\big|_{L^2(\Omega)\setminus M} = +\infty \,.$$

Außerdem sei $W\big|_M$ konvex, bezüglich der $H_1(\Omega)$–Norm von unten halbstetig, und es gebe Konstanten $c_1 > 0$, $c_2 \geq 0$ mit

$$W(u) \geq c_1 |u|^2_{1,\Omega} - c_2$$

für alle $u \in M$. Dann ist W konvex und bezüglich der L^2–Norm von unten halbstetig. Ist $W\big|_M$ auch koerzitiv bezüglich der $H_1(\Omega)$–Norm, dann ist W zusätzlich auch koerzitiv bezüglich der L^2–Norm.

Beweis: Um zu zeigen, daß W konvex ist, muß für alle $u, v \in L^2(\Omega)$ und für alle $0 < t < 1$ gezeigt werden, daß

$$(*) \qquad W(tu + (1-t)v) \leq tW(u) + (1-t)W(v)$$

gilt. Wenn $u, v \in M$ sind, ist dies nach Voraussetzung klar. Wenn wenigstens eine der Funktionen u, v nicht zu M gehört, dann ist der Wert der rechten Seite dieser Ungleichung $+\infty$, also ist sie erfüllt. Somit ist W konvex. Wenn $W|_M$ koerzitiv ist, existieren Konstanten $c_1 > 0$, $c_2 \geq 0$ mit

$$W(u) \geq c_1 \|u\|_{1,\Omega}^2 - c_2 \geq c_1 \|u\|_\Omega^2 - c_2$$

für alle $u \in M$. Für alle anderen u ist $W(u) = +\infty$, also ist W koerzitiv.

Es bleibt zu zeigen, daß W von unten halbstetig ist. Hierzu muß bewiesen werden, daß für $u \in L^2(\Omega)$ und für $\{u_m\}_{m=1}^\infty \subseteq L^2(\Omega)$ mit $\|u - u_m\|_\Omega \to 0$ gilt

$$W(u) \leq \liminf_{m\to\infty} W(u_m)\,.$$

Falls $\liminf_{m\to\infty} W(u_m) = +\infty$ gilt, ist dies richtig. Sei also $\liminf_{m\to\infty} W(u_m) = a < \infty$. Wähle eine Teilfolge $\{u'_m\}_{m=1}^\infty$ von $\{u_m\}_{m=1}^\infty$ aus mit $\lim_{m\to\infty} W(u'_m) = \alpha$. Für alle genügend großen m ist dann $W(u'_m) < \infty$, also $u'_m \in M$; ohne Beschränkung der Allgemeinheit kann also angenommen werden, daß $\{u'_m\}_{m=1}^\infty \subseteq M \subseteq \overset{\circ}{H}_1(\Omega)$ ist. Nach Voraussetzung gilt für alle $m \in \mathbb{N}$

$$|u'_m|_{1,\Omega}^2 \leq \frac{1}{c_1}(W(u'_m) + c_2) \leq r^2$$

mit einer geeigneten Konstanten $r > 0$, also ist

$$\{\frac{\partial}{\partial x_i} u'_m\}_{m=1}^\infty \subseteq \overline{B_r(0)} = \big\{v \in L^2(\Omega) \mid \|v\|_\Omega \leq r\big\}\ , \quad i = 1, \ldots, n\,.$$

Da $\overline{B_r(0)}$ nach Satz 6.17 schwach folgenkompakt ist, gibt es eine Teilfolge $\{u''_m\}_{m=1}^\infty$ von $\{u'_m\}_{m=1}^\infty$ und Funktionen $w_1, \ldots, w_n \in \overline{B_r(0)}$ mit

$$\frac{\partial}{\partial x_i} u''_m \underset{L^2(\Omega)}{\rightharpoonup} w_i\,,$$

für alle $i = 1, \dots, n$. Für alle $\varphi \in \overset{\circ}{C}_\infty(\Omega)$ folgt nun

$$(w_i, \varphi)_\Omega = \lim_{m\to\infty} (\frac{\partial}{\partial x_i} u''_m, \varphi)_\Omega = - \lim_{m\to\infty} (u''_m, \frac{\partial}{\partial x_i} \varphi)_\Omega = -(u, \frac{\partial}{\partial x_i} \varphi)_\Omega \,,$$

und dies bedeutet, daß $u \in H_1(\Omega)$ ist mit

$$\frac{\partial}{\partial x_i} u = w_i \,, \quad i = 1, \dots, n \,.$$

Es gilt somit $\frac{\partial}{\partial x_i} u''_m \underset{L^2(\Omega)}{\rightharpoonup} \frac{\partial}{\partial x_i} u$, und daraus folgt für alle $v \in H_1(\Omega)$

$$\begin{aligned} \lim_{m\to\infty} (u''_m, v)_{1,\Omega} &= \lim_{m\to\infty} [(u''_m, v)_\Omega + (\nabla u''_m, \nabla v)_\Omega] \\ &= (u, v)_\Omega + (\nabla u, \nabla v)_\Omega = (u, v)_{1,\Omega} \,, \end{aligned}$$

also

$$u''_m \underset{H_1(\Omega)}{\rightharpoonup} u \,.$$

Nach Voraussetzung ist M konvex und bezüglich der $H_1(\Omega)$–Norm abgeschlossen. Nach Satz 6.13 ist M also auch bezüglich der schwachen Toplogie von $H_1(\Omega)$ abgeschlossen, folglich gilt $u \in M$.

Wieder nach Voraussetzung ist $W|_M$ konvex und bezüglich der $H_1(\Omega)$–Norm von unten halbstetig. Nach Satz 7.3 ist somit $W|_M$ auch schwach von unten halbstetig bezüglich der schwachen Topologie von $H_1(\Omega)$, also folgt

$$\begin{aligned} W(u) &\leq \liminf_{m\to\infty} W(u''_m) = \lim_{m\to\infty} W(u'_m) \\ &= \liminf_{m\to\infty} W(u_m) \,; \end{aligned}$$

dies bedeutet, daß W von unten halbstetig ist.

8.5 Folgerung (Konvexität und Unterhalbstetigkeit von Variationsfunktionalen auf $L^2(\Omega)$). Sei $F : \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar, konvex und strikt koerzitiv. Dann ist $\tilde{V} : L^2(\Omega) \to (-\infty, \infty]$ konvex und von unten halbstetig.

Beweis: Satz 8.3 zeigt, daß die Voraussetzungen von Lemma 8.4 erfüllt sind. Also folgt die Behauptung aus diesem Lemma.

Als letztes Zwischenergebnis benötige ich

8.6 Lemma (Eine Version der Poincaréschen Ungleichung). Sei $\Omega \subseteq \mathbb{R}^n$ eine offene und beschränkte Menge mit

$$d = \sup_{x,y\in\Omega} |x-y| \,,$$

und sei $0 \le b < \infty$. Dann gilt für $u \in M$

$$\|u\|_\Omega^2 \le d^2 |u|_{1,\Omega}^2 + 2b^2 \operatorname{meas}\Omega \,,$$

wobei $\operatorname{meas}\Omega = \int_\Omega \mathrm{dx}$ sei.

Beweis: Sei zunächst $u \in \tilde{M}(b) = \{u \in C_1(\overline{\Omega}) \cap H_1(\Omega) \mid \forall x \in \partial\Omega : 0 \le u(x) \le b\}$, sei $x = (x_1, \dots, x_n) \in \Omega$, $x' = (x_2, \dots, x_n)$ und sei $L(x') = \{z \in \mathbb{R}^n \mid (z_2, \dots, z_n) = x'\}$. Schließlich sei $y \in (L(x') \cap \partial\Omega)$ der Randpunkt von Ω mit $y_1 < x_1$ und mit der Eigenschaft, daß alle Punkte auf der Geraden $L(x')$ zwischen y und x zu Ω gehören. Dann folgt aus dem Hauptsatz der Differential– und Integralrechung und aus der Cauchy-Schwarzschen Ungleichung

$$u(x) = \int_{y_1}^{x_1} \frac{\partial}{\partial x_1} u(z) dz_1 + u(y) \,,$$

also

$$\begin{aligned} |u(x)|^2 &\le 2 \,| \int_{y_1}^{x_1} u(z) dz_1|^2 + 2\,|u(y)|^2 \\ &\le 2(x_1 - y_1) \int_{y_1}^{x_1} |\frac{\partial}{\partial x_1} u(z)|^2 dz_1 + 2\,|u(y)|^2 \,, \end{aligned}$$

folglich

$$\begin{aligned} \int_{y_1}^{x_1} |u(z)|^2 dz_1 &\le (x_1 - y_1)^2 \int_{y_1}^{x_1} |\frac{\partial}{\partial x_1} u(z)|^2 dz_1 + 2 \int_{y_1}^{x_1} |u(y)|^2 dz_1 \\ &\le d^2 \int_{y_1}^{x_1} |\nabla u(z)|^2 dz_1 + 2 \int_{y_1}^{x_1} b^2 \, dz \end{aligned}$$

und somit

$$\int_{L(x')\cap\Omega} |u(z)|^2 dz \le d^2 \int_{L(x')\cap\Omega} |\nabla u(z)|^2\, dz + 2b^2 \int_{L(x')\cap\Omega} dz\,.$$

Die behauptete Ungleichung ergibt sich aus dieser Ungleichung für $u \in \tilde{M}(b)$ durch Integration bezüglich x'. Um diese Ungleichung für $u \in M(b)$ zu beweisen, wähle man eine Folge $\{u_m\}_{m=1}^\infty \subseteq M_1$ mit $\|u - u_m\|_{1,\Omega} \to 0$. Dann folgt

$$\begin{aligned} \|u\|_\Omega^2 &= \lim_{m\to\infty} \|u_m\|_\Omega^2 \le \lim_{m\to\infty} (d^2 |u_m|_{1,\Omega}^2 + 2b^2 \operatorname{meas}\Omega) \\ &= d^2 |u|_{1,\Omega}^2 + 2b^2 \operatorname{meas}\Omega\,. \end{aligned}$$

Beweis von Satz 8.2. (i) Nach Folgerung 8.5 und Folgerung 7.9 ist für $\lambda > 0$

$$v \mapsto \tilde{V}(v) + \frac{\lambda}{2}\, \|v\|_\Omega^2 : L^2(\Omega) \to L^2(\Omega)$$

strikt konvex, von unten halbstetig und koerzitiv. Nach Lemma 7.6 gilt dies auch für

$$v \mapsto \tilde{V}(v) + \frac{\lambda}{2}\, \|v\|_\Omega^2 - (f, v)_\Omega\,.$$

Also besitzt $\tilde{V}(v) + \frac{\lambda}{2}\, \|v\|_\Omega^2 - (f, v)_\Omega$ nach Satz 7.7 ein eindeutiges Minimum $u \in L^2(\Omega)$. Wegen $M(F, b) \ne \emptyset$ ist $\tilde{V}(u) + \frac{\lambda}{2}\, \|u\|_\Omega^2 - (f, u)_\Omega < \infty$, und somit $u \in M(F, b)$.

(ii) Der Beweis folgt aus Lemma 8.6. Denn nach Satz 8.3 folgt aus
$u \in M(b) \subseteq H_1(\Omega)$

$$\|u\|_\Omega^2 \le d^2\, |u|_{1,\Omega}^2 + 2b^2 \operatorname{meas}\Omega \le \frac{d^2}{c}\, V(u) + 2b^2 \operatorname{meas}\Omega$$

mit $d = \sup_{x,y\in\Omega} |x - y|$. Wegen $\tilde{V}(u) = \infty$ für $u \in L^2(\Omega)\backslash M(b)$ folgt hieraus

$$\tilde{V}(u) \ge \frac{c}{d^2}\, \|u\|_\Omega^2 - \frac{2cb^2}{d^2} \operatorname{meas}\,(\Omega)$$

für alle $u \in L^2(\Omega)$. Also ist $\tilde{V}$ koerzitiv, und folglich ist $\tilde{V}(v) - (f, v)$ koerzitiv, konvex und von unten halbstetig. Nach Satz 7.7 existiert also ein Minimum

u von $\tilde{V}(v) - (f, v)$. Wie oben folgt, daß $u \in M(F, b)$.

Im nächsten Schritt wird das Subdifferential von $\tilde{V}$ bestimmt.

8.7 Satz (Subdifferential und Variationsungleichung). Sei $F : \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar, konvex, strikt koerzitiv und sei $M(F, b) \neq \emptyset$. Genau dann ist $w \in \partial\tilde{V}(u)$, wenn $u \in M(F, b)$, wenn

$$\int_\Omega |a(\nabla u(x)) \cdot (\nabla v(x) - \nabla u(x))| dx < \infty$$

und wenn

$$\int_\Omega a(\nabla u(x)) \cdot (\nabla v(x) - \nabla u(x)) dx \geq (w, v - u)_\Omega$$

für alle $v \in M(F, b)$. Hierbei ist $a(\xi) = \mathrm{grad}\, \mathrm{F}(\xi)$.

Beweis: Beachte zunächst, daß für $u, v \in M(F, b)$, $0 \leq t \leq 1$ und für alle $x \in \Omega$ die Funktion

$$\frac{1}{t} \left[F(t\nabla v(x) + (1-t)\nabla u(x)) - F(\nabla u(x)) \right]$$

monoton fallend ist für $t \searrow 0$. Denn sei

$$H(t) = F(t\nabla v(x) + (1-t)\nabla u(x)) \, ,$$

$0 < t_1 \leq t_2 \leq 1$ und $\alpha = \frac{t_1}{t_2}$. Dann ist H konvex, und somit folgt

$$\begin{aligned} \frac{1}{t_1}(H(t_1) - H(0)) &= \frac{1}{t_2} \frac{1}{\alpha} \left(H(\alpha t_2 + (1-\alpha)0) - H(0) \right) \\ &\leq \frac{1}{t_2} \frac{1}{\alpha} \left(\alpha H(t_2) + (1-\alpha) H(0) - H(0) \right) \\ &= \frac{1}{t_2} \left(H(t_2) - H(0) \right) . \end{aligned}$$

Hieraus folgt die Behauptung. Da $tv + (1-t)u \in M(F, b)$ ist, ist also

$$x \to \frac{1}{t} \left[F(t\nabla v(x) + (1-t)\nabla u(x)) - F(\nabla u(x)) \right]$$

für $t \searrow 0$ eine monoton fallende Folge integrierbarer Funktionen.

Sei nun $w \in \partial\tilde{V}(u)$. Dann folgt für alle $v \in L^2(\Omega)$, daß

$$\tilde{V}(v) \geq (w, v-u) + \tilde{V}(u)$$

gilt. Da $M(F,b)$ nicht leer ist, gibt es $v \in M(F,b)$ mit $\tilde{V}(v) < \infty$. Dies kann nur sein, wenn $\tilde{V}(u) < \infty$ ist, und dies impliziert $u \in M(F,b)$. Für $v \in M(F,b)$ folgt also aus dem Satz von Beppo Levi

$$\begin{aligned}
(w, v-u)_\Omega &\leq \lim_{t\searrow 0} \frac{1}{t}\,[\tilde{V}(tv + (1-t)u) - \tilde{V}(u)] \\
&= \lim_{t\searrow 0} \int_\Omega \frac{1}{t}\,[F(t\nabla v(x) + (1-t)\nabla u(x)) - F(\nabla u(x))]\,dx \\
&= \int_\Omega \lim_{t\searrow 0} \frac{1}{t}\,[F(t\nabla v(x) + (1-t)\nabla u(x)) - F(\nabla u(x))]\,dx \\
&= \int_\Omega \frac{d}{dt}\,F(t\nabla v(x) + (1-t)\nabla u(x))\,|_{t=0}\ dx \\
&= \int_\Omega \mathrm{grad\,F(\nabla u(x)) \cdot (\nabla v(x) - \nabla u(x))dx} = \int_\Omega \mathrm{a(\nabla u) \cdot (\nabla v - \nabla u)dx}\,,
\end{aligned}$$

und der Satz von Beppo Levi liefert auch, daß das letzte Integral existiert, also daß

$$(*) \qquad \int_\Omega |a(\nabla u) \cdot (\nabla v - \nabla u)|\,dx < \infty\,.$$

Wenn umgekehrt $u \in M(F,b)$ ist, das Integral $(*)$ endlich ist, und

$$\int_\Omega a(\nabla u) \cdot (\nabla v - \nabla u)dx \geq (w, v-u)$$

gilt für alle $v \in M(F,b)$, dann folgt

$$\begin{aligned}
(w, v-u)_\Omega &\leq \int_\Omega a(\nabla u) \cdot (\nabla v - \nabla u)dx \\
&= \int_\Omega \lim_{t\searrow 0} \frac{1}{t}\,[F(t\nabla v + (1-t)\nabla u) - F(\nabla u)]\,dx \\
&\leq \int_\Omega (F(\nabla v) - F(\nabla u))dx = \tilde{V}(v) - \tilde{V}(u)\,,
\end{aligned}$$

also

$$\tilde{V}(v) \geq (w, v-u)_\Omega + \tilde{V}(u)\,.$$

Diese Ungleichung gilt auch für alle $v \in L^2(\Omega)\backslash M(F,b)$, weil dann $\tilde{V}(v) = +\infty$ ist, also ist $w \in \partial\tilde{V}(u)$.

8.8 Folgerung (Variationsungleichung zum Variationsfunktional. Schwache Form der Eulergleichung). Sei $\Omega \subseteq \mathbb{R}^n$ eine offene Menge. $F : \mathbb{R}^n \to \mathbb{R}$ sei stetig differenzierbar, konvex, strikt koerzitiv und sei $M(F,b) \neq 0$. Sei $f \in L^2(\Omega)$ und sei $\lambda \geq 0$. Dann sind die folgenden Aussagen (i) und (ii) äquivalent:

(i) $u \in L^2(\Omega)$ und

$$\tilde{V}(u) + \frac{\lambda}{2}\,\|u\|_\Omega^2 - (f,u)_\Omega = \min_{v \in L^2(\Omega} [\tilde{V}(v) + \frac{\lambda}{2}\,\|v\|_\Omega^2 - (f,v)_\Omega\,]\,;$$

(ii) $u \in M^*$,

$$\int_\Omega |a(\nabla u(x)) \cdot (\nabla v(x) - \nabla u(x))|\,dx < \infty$$

und

$$(**)\ \int_\Omega [a(\nabla u(x)) \cdot (\nabla v(x) - \nabla u(x)) + (\lambda u(x) - f(x))\,(v(x) - u(x))\,]\,dx \geq 0$$

für alle $v \in M^*$. Hierbei sei $a(\xi) = \operatorname{grad} \mathrm{f}(\xi)$.

Bemerkung. Die Ungleichung $(**)$ heißt Variationsungleichung. Dies ist eine Verallgemeinerung der schwachen Eulergleichung aus 4.3.

Beweis: Im Beweis von Folgerung 7.15 wurde gezeigt, daß

$$\partial(\tilde{V}(u) + \frac{\lambda}{2}\,\|u\|_\Omega^2) = \lambda u + \partial\tilde{V}(u)$$

ist. Da $\tilde{V}(u) + \frac{\lambda}{2}\,\|u\|_\Omega^2$ konvex ist, ist (i) nach Lemma 7.13 äquivalent zu

$$f \in \partial(\tilde{V}(u) + \frac{\lambda}{2}\,\|u\|_\Omega^2) = \lambda u + \partial\tilde{V}(u)\,,$$

also äquivalent zu
$$f - \lambda u \in \partial \tilde{V}(u)\,,$$
und dies ist nach Satz 8.7 äquivalent zu (ii).

8.9 Das Subdifferential. Es soll nun der Zusammenhang zwischen der Variationsungleichung und der partiellen Differentialgleichung untersucht werden. Nach Satz 8.7 ist $w \in \partial \tilde{V}(u)$ genau dann, wenn $u \in M(F,b)$ und

$$(*) \qquad \int_\Omega a(\nabla u(x)) \cdot (\nabla v(x) - \nabla u(x))dx \geq (w, v-u)_\Omega$$

gilt für alle $v \in M(F,b)$. Hieraus folgt, daß
$$w + (\,-u + M(F,b))^\perp \subseteq \partial \tilde{V}(u)$$
gilt, wobei w ein beliebiges Element aus $\partial \tilde{V}(u)$ ist. Also kann
$$\partial \tilde{V}(u) = \{-\mathrm{div\ a(\nabla u)}\}$$
höchstens dann gelten, wenn $(-u + M(F,b))^\perp = \{0\}$ ist. Um schwache Ableitungen definieren zu können, möchte man insbesondere haben, daß $\overset{\circ}{C}_\infty(\Omega) \subseteq -u + M(F,b)$ gilt, weil man dann in $(*)$ $v = u \pm \varphi$ setzen kann für jedes $\varphi \in \overset{\circ}{C}_\infty(\Omega)$. Falls $F(\xi)$ jedoch sehr schnell wächst für $|\xi| \to \infty$, kann $M(F,b)$ eine kleine Menge sein, und es ist nicht klar, ob $\overset{\circ}{C}_\infty(\Omega)$ in $-u + M(F,b)$ enthalten ist. Aus diesem Grunde und um weitere technische Schwierigkeiten zu vermeinden, werde ich im folgenden voraussetzen, daß $F : \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar ist mit
$$F(0) = 0$$
und mit
$$|a(\xi)| = |\,\mathrm{grad\,F}(\xi)| \leq \bar{c}\,|\xi|$$
für eine geeignete Konstante $\bar{c} > 0$. Es folgt dann aus dem Mittelwertsatz (oder aus der Konvexität von F) und aus $F(0) = 0$, daß
$$\begin{aligned} |F(\xi)| &\leq |\,\mathrm{grad\,F}(\xi^*) \cdot \xi + \mathrm{F}(0)| \\ &\leq \bar{c}\,|\xi|^2 + |F(0)| = \bar{c}\,|\xi|^2 \end{aligned}$$

und somit
$$\int_\Omega F(\nabla u(x))dx \leq \int_\Omega \overline{c}\,|\nabla u(x)|^2 dx = \overline{c}\,|u|^2_{1,\Omega}$$
gilt. Dies ergibt $M(F,b) = M(b) = \overline{\tilde{M}(b)}^{H_1(\Omega)}$. Nach Definition von $\tilde{M}(b)$ ist zu allen $u \in \tilde{M}(b)$ und $\varphi \in \overset{\circ}{C}_\infty(\Omega)$ auch $u+\varphi \in \tilde{M}(b)$ und folglich ist $u+\varphi \in M(b)$ für alle $u \in M(b)$. Es folgt
$$\overset{\circ}{C}_\infty(\Omega) \subseteq -u + M(F,b) = -u + M(b)\,,$$
also $(-u + M(F,b))^\perp = \{0\}$. Weiterhin folgt für alle $u \in M(b)$
$$\int_\Omega |a(\nabla u(x))|^2 dx \leq \overline{c}^2 \int_\Omega |\nabla u(x)|^2 dx = \overline{c}^2 |u|^2_{1,\Omega} < \infty\,,$$
also $a(\nabla u(x)) \in L^2(\Omega, \mathbb{R}^n)$. Bevor $\partial\tilde{V}(u)$ nun genau bestimmt werden kann, benötigt man folgende Definition.

8.10 Definition (schwache Divergenz). Sei $\Omega \subseteq \mathbb{R}^n$ eine offene Menge und $g \in L^2(\Omega, \mathbb{R}^n)$. Wenn eine Funktion $f \in L^2(\Omega, \mathbb{R})$ existiert mit
$$(g, \nabla\varphi)_\Omega = -(f, \varphi)_\Omega$$
für alle $\varphi \in \overset{\circ}{C}_\infty(\Omega)$, dann wird f mit $\mathrm{div}\,\mathrm{g}$ bezeichnet, und man sagt $\mathrm{div}\,\mathrm{g}$ (die Divergenz von g) existiere im verallgemeinerten Sinn.

8.11 Definition (schwache Definition von Differentialoperatoren mit Randbedingung). Sei $\Omega \subseteq \mathbb{R}^n$ eine offene Menge. Der im allgemeinen nichtlineare Operator $A : D(A) \subseteq L^2(\Omega) \to L^2(\Omega)$ mit Definitionsbereich $D(A)$ sei folgendermaßen definiert: Es sei $D(A)$ die Menge aller $u \in M(b)$, für die $\mathrm{div}\,\mathrm{a}(\nabla\mathrm{u}) \in \mathrm{L}^2(\Omega)$ existiert, und für die
$$(a(\nabla u),\, \nabla v - \nabla u)_\Omega \geq -(\mathrm{div}\,\mathrm{a}(\nabla\mathrm{u}),\, \mathrm{v} - \mathrm{u})_\Omega$$
gilt für alle $v \in M(b)$. Für $u \in D(A)$ sei
$$Au = -\mathrm{div}\,\mathrm{a}(\nabla\mathrm{u})\,.$$
Hierbei wird vorausgesetzt, daß a die Voraussetzungen aus 8.9 erfüllt.

Zur Motivation dieser Definition beachte man, daß div g eindeutig bestimmt ist (Beweis wie in Lemma 2.11), und daß für $g \in C_1(\Omega)$ die verallgemeinerte Divergenz von g mit der klassischen Divergenz $\sum_{i=1}^n \frac{\partial}{\partial x_i} g(x)$ übereinstimmt. Wenn $a \in C_1(\mathbb{R}^n)$ und $u \in C_2(\overline{\Omega})$ ist, ist also $\mathrm{div\, a(\nabla u(x))}$ die klassische Divergenz. Der Gaußsche Satz liefert dann für Gebiete mit glattem Rand $\partial\Omega$ und für $v \in \tilde{M}(b)$

$$
(*) \qquad \begin{aligned}
&(a(\nabla u),\, \nabla v - \nabla u)_\Omega + (\mathrm{div\, a(\nabla u),\, v - u})_\Omega \\
&= \int_\Omega [a(\nabla u(x)) \cdot (\nabla v(x)) - \nabla u(x)) + \mathrm{div\, a(\nabla u(x))\,(v(x) - u(x))}\,]\,\mathrm{dx} \\
&= \int_\Omega \mathrm{div\,[a(\nabla u(x))\,(v(x) - u(x))\,]\,dx} \\
&= \int_{\partial\Omega} n(x) \cdot a(\nabla u(x))\,(v(x) - u(x))\,dS_x\,,
\end{aligned}
$$

wobei $n(x)$ die äußere Einheitsnormale an $\partial\Omega$ im Punkte $x \in \partial\Omega$ ist. Wenn u zu $D(A)$ gehört, muß die rechte Seite dieser Gleichung größer oder gleich Null sein, und für $b > 0$ ist dies genau dann der Fall, wenn

$$
(**) \qquad n(x) \cdot a(\nabla u(x)) \begin{cases} \geq 0\,, & u(x) = 0 \\ = 0\,, & 0 < u(x) < b \\ \leq 0\,, & u(x) = b \end{cases}
$$

gilt. Denn daß die rechte Seite von $(*)$ nicht negativ ist, wenn $(**)$ erfüllt ist, ist klar. Die Umkehrung erhält man ähnlich wie in Abschnitt 1.1, indem man $v = u + \psi \in M(b)$ setzt mit Funktionen $\psi \in C_\infty(\overline{\Omega})$, die am Rande $\partial\Omega$ geeignete Werte annehmen. Solche Funktionen ψ können wie im dritten Kapitel durch Faltung mit geeigneten $\overset{\circ}{C}_\infty$–Funktionen konstruiert werden. Die Einzelheiten bleiben dem Leser überlassen. Dies zeigt, daß die Forderung

$$
(a(\nabla u),\, \nabla v - \nabla u)_\Omega \geq -(\mathrm{div\, a(\nabla u),\, v - u})_\Omega
$$

eine Verallgemeinerung der Randbedingung $(**)$ ist auf Gebiete mit nichtglattem Rand und auf Funktionen, für die $\nabla u(x)$ am Rande nicht unbedingt zu existieren braucht.

8.12 Satz (Dirichletsches Prinzip. Äquivalenz zwischen Randwert-

problem, Variationsungleichung und Variationsproblem). Sei $\Omega \subseteq \mathbb{R}^n$ eine offene Menge, sie $F : \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar, konvex und strikt koerzitiv. Außerdem gelte $F(0) = 0$, und es gebe eine Konstante $\overline{c} > 0$ mit

$$|a(\xi)| = |\mathrm{grad\,F}(\xi)| \leq \overline{\mathrm{c}}\, |\xi|$$

für alle $\xi \in \mathbb{R}^n$. Dann gilt

$$\partial\tilde{V}(u) = \begin{cases} \{Au\} & , \quad u \in D(A)\,. \\ \emptyset & , \quad u \in L^2(\Omega)\backslash D(A)\,. \end{cases}$$

Für $f \in L^2(\Omega)$ und $\lambda \geq 0$ sind folglich äquivalent:

(i) $u \in D(A)$ und $Au + \lambda u = f$

(ii) $u \in M(b) = M(F, b)$ und

$$(a(\nabla u),\, \nabla v - \nabla u) + \lambda(u, v-u)_\Omega \geq (f, v-u)_\Omega$$

für alle $v \in M(b)$.

(iii) $u \in L^2(\Omega)$ und

$$\tilde{V}(u) + \frac{\lambda}{2}\, \|u\|_\Omega^2 - (f, u)_\Omega = \min_{v \in L^2(\Omega)} [\tilde{V}(v) + \frac{\lambda}{2}\, \|v\|_\Omega^2 - (f, v)_\Omega]\,.$$

Wenn $\lambda > 0$ ist, gibt es ein eindeutiges u, das diese drei äquivalenten Eigenschaften hat. Wenn $\lambda = 0$ ist, gibt es für $b < \infty$ bei beschränktem Ω wenigstens ein u, das diese drei Eigenschaften hat.

8.12.1 Bemerkung: Dies bedeutet natürlich, daß das Randwertproblem

$$-\mathrm{div\;a}(\nabla \mathrm{u(x)}) + \lambda \mathrm{u(x)} = \mathrm{f(x)} \quad , \quad x \in \Omega$$

$$0 \leq u(x) \leq b \qquad\qquad , \quad x \in \partial\Omega$$

$$n(x) \cdot a(\nabla u(x)) \left\{ \begin{array}{ll} \geq 0\,, & u(x) = 0 \\ = 0\,, & 0 < u(x) < b \\ \leq 0\,, & u(x) = b \end{array} \right\} x \in \partial\Omega$$

unter den angegebenen Bedingungen verallgemeinerte Lösungen besitzt.

Beweis. Sei $u \in L^2(\Omega)$ und $\partial\tilde{V}(u) \neq \emptyset$. Dann gibt es ein $w \in L^2(\Omega)$ mit $w \in \partial\tilde{V}(u)$, und nach Satz 8.7 ist dies äquivalent zu $u \in M(F,b) = M(b)$ und

$$(*) \qquad (a(\nabla u),\, \nabla v - \nabla u)_\Omega \geq (w, v-u)_\Omega$$

für alle $v \in M(b)$. Nach 8.9 ist unter den angegebenen Voraussetzungen $v = u \pm \varphi \in M(b)$ für alle $\varphi \in \overset{\circ}{C}_\infty(\Omega)$. Für diese v ergibt sich nun

$$\pm(a(\nabla u),\, \nabla\varphi)_\Omega \geq \pm(w,\varphi)_\Omega\,,$$

folglich

$$(a(\nabla u),\, \nabla\varphi)_\Omega = (w,\varphi)_\Omega$$

für alle $\varphi \in \overset{\circ}{C}_\infty(\Omega)$. Da $a(\nabla u) \in L^2(\Omega)$ ist, bedeutet dies nach Definition 8.10, daß $\mathrm{div}\,(\mathrm{a}(\nabla \mathrm{u}))$ existiert, und daß $w = -\mathrm{div}\,(\mathrm{a}(\nabla\mathrm{u}))$. Dies impliziert $\partial\tilde{V}(u) = \{-\mathrm{div}\,(\mathrm{a}(\nabla\mathrm{u}))\}$, und $(*)$ nimmt die Form

$$(a(\nabla u),\, \nabla v - \nabla u)_\Omega \geq -(\mathrm{div}\,\mathrm{a}(\nabla\mathrm{u}),\, \mathrm{v}-\mathrm{u})_\Omega$$

an. Nach Definition 8.11 ist daher $u \in D(A)$ und $\partial\tilde{V}(u) = \{Au\}$. Sei umgekehrt $u \in D(A)$. Nach Definition 8.11 bedeutet dies $u \in M(b)$ und

$$(a(\nabla u),\, \nabla v - \nabla u)_\Omega \geq -(\mathrm{div}\,\mathrm{a}(\nabla\mathrm{u}),\, \mathrm{v}-\mathrm{u})_\Omega$$

für alle $v \in M(b)$. Satz 8.7 zeigt nun, daß

$$-\mathrm{div}\,\mathrm{a}(\nabla\mathrm{u}) \in \partial\tilde{\mathrm{V}}(\mathrm{u})\,,$$

also ist $\partial\tilde{V}(u) \neq \emptyset$. Zusammen folgt

$$\partial\tilde{V}(u) = \begin{cases} \{Au\} & , \quad u \in D(A) \\ \emptyset & , \quad u \in L^2(\Omega)\backslash D(A)\,. \end{cases}$$

Aus diesem Ergebnis folgt, daß $u \in D(A)$ und $Au + \lambda u = f$ äquivalent ist zu $f - \lambda u \in \partial\tilde{V}(u)$. Nach Satz 8.7 ist dies aber äquivalent zu (ii). Also ist (i) äquivalent zu (ii). Die Äquivalenz von (ii) und (iii) wurde schon in Folgerung 8.8 bewiesen. Die Existenz von u ergibt sich aus Satz 8.2.

9. Existenztheorie für das Hindernisproblem bei nichtlinearer Randbedingung.

9.1 Das Potential beim Hindernisproblem. Wir werden nun die Lösbarkeit des in Abschnitt 1.2 betrachteten Hindernisproblems für die Membran studieren. Sei $\Omega \subseteq \mathbb{R}^n$ eine offene beschränkte Menge mit Lipschitzrind $\partial\Omega$ und sei $\psi \in C_\infty(\overline{\Omega}, \mathbb{R})$. Weiterhin sei $j : \mathbb{R} \to \mathbb{R}$ eine stetige differenzierbare, konvexe, strikt koerzitive Funktion mit $j(0) = 0$. Aus diesen Voraussetzungen folgt $j(\xi) \geq 0$ für alle $\xi \in \mathbb{R}$. Schließlich sei

$$M = \{u \in H_1(\Omega) | u(x) \geq \psi(x) \text{ fast überall in } \Omega\}\,.$$

M ist abgeschlossen in $H_1(\Omega)$, und nach Satz 8.4 existieren die Randwerte $u_{|\partial\Omega} = Bu$ für alle $u \in M$. Sei

$$M^* = \left\{ u \in M | \int_{\partial\Omega} j(Bu(x)) dS_x < \infty \right\}\,.$$

Wir definieren nun das Funktional $\tilde{V} : L^2(\Omega) \to L^2(\Omega)$ durch

$$\tilde{V}(u) = \begin{cases} \int_\Omega \frac{1}{2}|\nabla u(x)|^2 dx + \int_{\partial\Omega} j(Bu(x)) dS_x < \infty\,, & u \in M^* \\ +\infty\,, & u \in L^2(\Omega) \backslash M^*\,. \end{cases}$$

$\tilde{V}$ ist das Potential, das zum in 1.2 betrachteten Hindernisproblem gehört. Die Aufgabe ist, zu zeigen, daß dieses Potential sein Minimum annimmt.

9.2 Satz. Sei j stetig differenzierbar, konvex, strikt koerzitiv und erfülle $j(0) = 0$. Dann ist M^* konvex, $\tilde{V}$ ist konvex und von unten halbstetig und $\tilde{V}_{|M^*}$ ist strikt konvex.

Beweis: Wie in Beispiel 5.5 sieht man, daß $\omega \mapsto \frac{1}{2}|\omega|^2_{1\Omega} : H_1(\Omega) \to \mathbb{R}$ strikt konvex ist. Also folgt für $u, v \in M^*$ mit $u \neq v$ und für $t \in (0,1)$

$$\begin{aligned} & \tfrac{1}{2}|tv + (1-t)u|^2_{1,\Omega} + \textstyle\int_{\partial\Omega} j(tv + (1-t)u) dS \\ (*) \quad < \; & \tfrac{1}{2}t|v^2_{1,\Omega} + \tfrac{1}{2}(1-t)|u|^2_{1,\Omega} + \textstyle\int_{\partial\Omega} tj(v(x)) + (1-t)j(u(x)) dS_x \\ = \; & t\tilde{V}(u) + (1-t)\tilde{V}(v) < \infty\,, \end{aligned}$$

weil j konkex ist. Außerdem gilt

$$tv(x) + (1-t)u(x) \geq t\psi(x) + (1-t)\psi(x) = \psi(x)$$

für fast alle $x \in \Omega$, also ist $tv + (1-t)u \in M^*$, und somit ist diese Menge konvex. Aus $(*)$ folgt nun aber

$$\tilde{V}(tv + (1-t)u) < t\tilde{V}(u) + (1-t)\tilde{V}(v)$$

für alle $u, v \in M^*$ mit $u \neq v$ und für alle $t \in (0,1)$, also ist $\tilde{V}_{|M^*}$ strikt konvex. Wenn mindestens eine der Funktionen $u, v \in L^2(\Omega)\backslash M^*$ ist, dann gilt $\tilde{V}(tv + (1-t)u) \leq t\tilde{V}(u) + (1-t)\tilde{V}(v)$, weil dann die rechte Seite den Wert $+\infty$ hat. Also ist $\tilde{V}$ konvex.
Sei nun $V_1 : H_1(\Omega) \to (-\infty, \infty]$ definiert durch

$$V_1(u) = \begin{cases} \int_{\partial\Omega} j(u(x))dS_x\,, & u \in M^* \\ +\infty\,, & u \in H_1(\Omega)\backslash M^*\,. \end{cases}$$

Außerdem setzen wir $V = \tilde{V}_{|H_1(\Omega)}$. Dann gilt

$$V(v) = \frac{1}{2}|v|_{1,\Omega}^2 + V_1(v)$$

für alle $v \in H_1(\Omega)$. Wir zeigen nun, daß V_1 von unten halbstetig ist. Da $v \mapsto \frac{1}{2}|v|_{1,\Omega}^2 : H_1(\Omega) \to \mathbb{R}$ stetig ist, folgt dann, daß auch $V : H_1(\Omega) \to (-\infty, \infty]$ von unten halbstetig ist.
Sei $\{u_m\}_{m=1}^{\infty} \subseteq H_1(\Omega)$ und $u \in H_1(\Omega)$ mit $\|u - u_m\|_{1,\Omega} \to 0$ für $m \to \infty$.
Zu zeigen ist, daß

$$V_1(u) \leq \liminf_{m\to\infty} V_1(u_m)$$

gilt. Ohne Beschränkung der Allgemeinheit sei

$$\liminf_{m\to\infty} V_1(u_m) < \infty\,,$$

weil sonst nichts zu beweisen ist. Wähle eine Teilfolge $\{u'_m\}_{m=1}^{\infty}$ von $\{u_m\}_{m=1}^{\infty}$ mit

$$\lim_{m\to\infty} V_1(u'_m) = \liminf_{m\to\infty} V_1(u_m)\,.$$

Es folgt, daß $V_1(u'_m) < \infty$ ist für alle genügend großen m, also ist $u'_m \in M^* \subseteq M$ für alle genügend großen m. Da M abgeschlossen ist in $H_1(\Omega)$ und

da $\{u'_m\}_{m=1}^{\infty}$ in $H_1(\Omega)$ gegen u konvergiert, ergibt sich $u \in M$.
Nach Satz 8.4 ist die Abbildung $u \mapsto u_{|\partial\Omega} : H_1(\Omega) \to L^2(\partial\Omega)$ stetig, also gilt

$$\int_{\partial\Omega} |Bu(x) - Bu'_m(x)|^2 dS_x = \|u_{|\partial\Omega} - u'_{m|\partial\Omega}\|_{\partial\Omega}^2 \to 0$$

für $m \to \infty$. Also gibt es nach einem Satz aus der Lebesgueschen Integrationstheorie und nach Definition des Randintegrals in 8.3 eine Teilfolge $\{u''_m\}_{m=1}^{\infty}$ von $\{u'_m\}_{m=1}^{\infty}$ mit

$$\lim_{m\to\infty} [Bu''_m](x) = [Bu](x)$$

für fast alle $x \in \partial\Omega$. Nach Vorasussetzung ist j stetig differenzierbar. Also folgt

$$\lim_{m\to\infty} j(Bu''_m(x)) = j(Bu(x))$$

für fast alle $x \in \partial\Omega$. Wegen $j(\xi) \geq 0$ folgt aus dem Lemma von Fatou

$$\begin{aligned}
\int_{\partial\Omega} j([Bu](x))dS_x &= \int_{\partial\Omega} \lim_{m\to\infty} j([Bu''_m](x))dS_x \\
&= \int_{\partial\Omega} \liminf_{m\to\infty} j([Bu''_m](x))dS_x \\
&\leq \liminf_{m\to\infty} \int_{\partial\Omega} j([Bu''_m](x))dS_x \\
&= \liminf_{m\to\infty} V_1(u''_m) \\
&= \lim_{m\to\infty} V_1(u''_m) \\
&= \lim_{m\to\infty} V_1(u'_m) \\
&= \liminf_{m\to\infty} V_1(u_m) \\
&< \infty,
\end{aligned}$$

also ist $u \in M^*$ und somit

$$V_1(u) = \int_{\partial\Omega} j([Bu](x))dS_x \leq \liminf_{m\to\infty} V_1(u_m).$$

Dies bedeutet, daß V_1 von unten halbstetig ist, und somit auch $V : H_1(\Omega) \to (-\infty, \infty]$.

Um zu zeigen, daß $\tilde{V} : L^2(\Omega) \to (-\infty, \infty]$ von unten halbstetig ist, beachte daß nach Definition $j(\xi) \geq 0$, also $V_1(v) \geq 0$ und somit

$$V(v) = \frac{1}{2}|v^2_{1,\Omega} + V_1(v) \geq \frac{1}{2}|v|^2_{1,\Omega}$$

gilt für alle $v \in H_1(\Omega)$. Also erfüllt $\tilde{V}$ die Voraussetzung von Lemma 6.9, und dies bedeutet, daß $\tilde{V}$ nach Lemma 6.9 von unten halbstetig ist. Satz 9.2 ist bewiesen.

9.3 Satz. Sei $\Omega \subseteq \mathbb{R}^n$ eine offene und beschränkte Menge mit Lipschitzrand. Sei $j : \mathbb{R} \to \mathbb{R}$ eine stetig differenzierbare, konvexe und strikt koerzitive Funktion mit $j(0) = 0$.
Genau dann ist $\omega \in \partial\tilde{V}(u)$, wenn $u \in M^*$ ist und wenn für alle $v \in M^*$

$$\int_{\partial\Omega} |j'(Bu(x))(Bv(x) - Bu(x))| dS_x < \infty$$

sowie

$$\begin{aligned}\int_\Omega \nabla u(x) \cdot (\nabla v(x) &- \nabla u(x)) dx \\ &+ \int_{\partial\Omega} j'(Bu(x))(Bv(x) \\ &- Bu(x)) dS_x \\ &\geq (\omega, v-u)_\Omega\end{aligned}$$

gilt.

Beweis: Wie im Beweis zu Satz 7.5 folgt, daß für $u, v \in M^*, t \in (0,1]$ und $x \in \partial\Omega$

$$\frac{1}{t}\Big[j(tBv(x) + (1-t)Bu(x)) - j(Bu(x))\Big]$$

monoton fallend ist für $t \searrow 0$. Also ist

$$x \mapsto \frac{1}{t}\Big[j(tBv(x) + (1-t)Bu(x)) - j(Bu(x))\Big]$$

eine für $t \searrow 0$ monoton fallende Folge über $\partial\Omega$ integrierbarer Funktionen.
Sei nun $u \in L^2(\Omega)$ und $\omega \in \partial\tilde{V}(u)$. Dies bedeutet, daß

$$
(*) \qquad \begin{aligned} \tilde V(tv+(1-t)u) &\geq (\omega, tv+(1-t)u-u)_\Omega + \tilde V(u) \\ &= t(\omega, v-u)_\Omega + \tilde V(u) \end{aligned}
$$

gilt für alle $v \in L^2(\Omega)$ und $t \in [0,1]$. Da $\psi \in C_\infty(\overline{\Omega})$ ist und da Ω beschränkt ist, ist $\sup_{x\in\overline{\Omega}} |\psi(x)| < \infty$, also ist $\psi \in M^*$ und $\tilde V(\psi) < \infty$ nach Definition von M^* und von $\tilde V$. Setzt man also $t = 1$ und $v = \psi$ in $(*)$, so folgt, daß auch $\tilde V(u) < \infty$ sein muß, also daß $u \in M^*$ gelten muß.
Aus $(*)$ und aus dem Satz von Beppo Levi folgt nun für alle $v \in M^*$

$$
\begin{aligned}
(\omega, v-u)_\Omega &\leq \lim_{t\searrow 0} \frac{1}{t}\Big[\tilde V(tv+(1-t)u) - \tilde V(u)\Big] \\
&= \lim_{t\searrow 0} \Big[\int_\Omega \frac{1}{2t}(|t\nabla v + (1-t)\nabla u|^2 - |\nabla u|^2)dx \\
&+ \int_{\partial\Omega} \frac{1}{t}\Big(j(tBv+(1-t)Bu) - j(Bu)\Big)dS\Big] \\
&= \int_\Omega \lim_{t\to 0} \Big(\frac{t}{2}|\nabla v|^2 + (1-t)\nabla v\cdot\nabla u + \frac{1}{2}(t-2)|\nabla u|^2\Big)dx \\
&+ \int_{\partial\Omega} \lim_{t\searrow 0} \frac{1}{t}\Big[j(tBv+(1-t)Bu) - j(Bu)\Big]dS \\
&= \int_\Omega \nabla u\cdot(\nabla v - \nabla u)dx + \int_{\partial\Omega} \frac{d}{dt} j(tBv+(1-t)Bu)_{|t=0} dS \\
&= \int_\Omega \nabla u\cdot(\nabla v - \nabla u)dx + \int_{\partial\Omega} j'(Bu(x))(Bv(x) - Bu(x))dS_x\,.
\end{aligned}
$$

Außerdem ergibt der Satz von Beppo Levi, daß

$$
\int_{\partial\Omega} |j'(Bu(x))(Bv(x) - Bu(x))| dS_x
$$

existiert. Wenn umgekehrt $u \in M^*$ ist, dieses Integral für alle $v \in M^*$ exitiert und die im Satz behauptete Variationsungleichung erfüllt ist, dann folgt für alle $v \in M^*$

$$
\begin{aligned}
(\omega, v-u)_\Omega &\leq \int_\Omega \nabla u\cdot(\nabla v - \nabla u)dx + \int_{\partial\Omega} j'(Bu)(Bv-Bu)dS \\
&= \int_\Omega \lim_{t\searrow 0} \frac{1}{2t}[|t\nabla v - (1-t)\nabla u|^2 - |\nabla u|^2]dx \\
&+ \int_{\partial\Omega} \lim_{t\searrow 0} \frac{1}{t}\Big[j(tBv+(1-t)Bu) - j(Bu)\Big]dS
\end{aligned}
$$

$$\begin{aligned} &\leq \int_\Omega \Big(\frac{1}{2}|\nabla v|^2 - \frac{1}{2}|\nabla u|^2\Big)dx + \int_{\partial\Omega} (j(Bv) - j(Bu))dS \\ &= \tilde{V}(v) - \tilde{V}(u)\,, \end{aligned}$$

also

$$\tilde{V}(v) \geq (\omega, v-u)_\Omega + \tilde{V}(u)\,,$$

weil auch

$$t \to \frac{1}{2t}[|t\nabla v(x) - (1-t)\nabla u(x)|^2 - |\nabla u(x)|^2]$$

wegen der Konvexität von

$$t \mapsto \left|t\nabla v(x) + (1-t)\nabla u(x)\right|^2$$

monoton fallend ist für $t \searrow 0$. Diese Ungleichung gilt auch für alle $v \in L^2(\Omega)\backslash M^*$ weil dann $\tilde{V}(v) = \infty$ ist, also ist $\omega \in \partial\tilde{V}(u)$. Damit ist der Satz bewiesen. Es bleibt zu zeigen, daß $\tilde{V}$ koerzitiv ist.

9.4 Lemma. Sei $\Omega \subseteq \mathbb{R}^n$ eine offene und beschränkte Menge mit Lipschitzrand. Sei $j : \mathbb{R} \to \mathbb{R}$ stetig differenzierbar, konvex und strikt koerzitiv mit $j(0) = 0$. Dann ist $\tilde{V}$ koerzitiv.

Zum Beweis brauchen wir das folgende Ergebnis:

9.5 Lemma. Sei $\Omega \subseteq \mathbb{R}^n$ eine offene und beschränkte Menge mit Lipschitzrand. Dann existiert eine Konstante $C > 0$ mit

$$\|u\|_\Omega^2 \leq C(\|u\|_{2,0,\partial\Omega}^2 + |u|_{1,\Omega}^2)$$

für alle $u \in H_1(\Omega)$.

Beweis von Lemma 9.4. Da j strikt koerzitiv ist, existiert eine Konstante $c > 0$ mit $j(\xi) \geq c\xi^2$. Also folgt für $u \in M^*$ aus Lemma 9.5

$$\begin{aligned} \tilde{V}(u) &= \frac{1}{2}|u|_{1,\Omega}^2 + \int_{\partial\Omega} j(Bu(x))dS_x \\ &\geq \frac{1}{2}|u|_{1,\Omega}^2 + c\int_{\partial\Omega} |Bu(x)|^2 dS_x \end{aligned}$$

$$\begin{aligned} &= \frac{1}{2}|u|^2_{1,\Omega} + c\|u\|^2_{2,0,\partial\Omega} \\ &\geq \frac{1}{C}\min\Big(\frac{1}{2}, c\Big)\|u\|^2_{\Omega}\,. \end{aligned}$$

Wegen $\tilde{V}(u) = \infty$ für $u \in L^2(\Omega)\backslash M^*$ ist also $\tilde{V}$ koerzitiv.

Beweis von Lemma 9.5. (Durch Widerspruch.) Angenommen, die Behauptung sei falsch. Dann existierte eine Folge $\{u_m\}_{m=1}^\infty \subseteq H_1(\Omega)$ mit $\|u_m\|_{0,\Omega} = 1$ und mit

$$\|u_m\|^2_{2,0,\partial\Omega} + |u_m|^2_{1,\Omega} \to 0$$

für $m \to \infty$. Somit ist $\{u_m\}_{m=1}^\infty$ eine beschränkte Folge im Hilbertraum $H_1(\Omega)$ und hat damit nach Satz 8.11 eine in $L^2(\Omega)$ konvergente Teilfolge $\{u'_m\}_{m=1}^\infty$. Wegen $|u'_m|_{1,\Omega} \to 0$ ist diese Teilfolge sogar in $H_1(\Omega)$konvergent mit Grenzwert $u \in H_1(\Omega)$. Wegen

$$\begin{aligned} \big|\|u\|_{0,\Omega} - \|u'_m\|_{0,\Omega}\big| &\leq \|u - u'_m\|_{0,\Omega} \leq \|u - u'_m\|_{1,\Omega} \to 0\,, \\ \big||u|_{1,\Omega} - |u'_m|_{1,\Omega}\big| &\leq |u - u'_m|_{1,\Omega} \leq \|u - u'_m|_{1,\Omega} \to 0 \end{aligned}$$

und

$$\big|\|u\|_{2,0,\partial\Omega} - \|u'_m\|_{2,0,\partial\Omega}\big| \leq \|u - u'_m\|_{2,0\partial\Omega} \leq C\|u - u'_m\|_{1,\Omega} \to 0$$

für $m \to \infty$ folgt

$$(*) \qquad \|u\|_{0,\Omega} = 1, |u|_{1,\Omega} = 0 \text{ und } Bu = 0\,.$$

Nach Satz 8.10 ist dann $u \in \overset{\circ}{H}_1(\Omega)$. Die Poincarésche Ungleichung (Satz 3.14) liefert folglich

$$\|u\|_{0,\Omega} \leq C|u|_{1,\Omega} = 0\,,$$

im Widerspruch zu $(*)$, also muß die ursprüngliche Annahme falsch gewesen sein.

9.6 Folgerung. Sei $\Omega \subseteq \mathbb{R}^n$ eine offene und beschränkte Menge mit Lipschitzrand. Sei $j : \mathbb{R} \to \mathbb{R}$ eine stetig differenzierbare, konvexe und strikt koerzitive Funktion mit $j(0) = 0$. Sei $f \in L^2(\Omega)$. Dann sind äuqivalent:

a.) $u \in M^*$, für alle $v \in M^*$ gilt

$$\int_{\partial\Omega} |j'(Bu(x))(Bv(x) - Bu(x))| dS < \infty\,,$$

und es ist

$$\begin{aligned} \textstyle\int_\Omega \Big[\nabla u(x) \cdot (\nabla v(x) &- \nabla u(x)) \\ &- f(x)(v(x) - u(x))\Big] dz \\ &+ \textstyle\int_{\partial\Omega} j'(Bu(x))(Bv(x) \\ &- Bu(x)) dS \\ &\geq 0 \end{aligned}$$

für alle $v \in M^*$.

b.) $u \in M$ und

$$\begin{aligned} & \textstyle\int_\Omega \left[\frac{1}{2}|\nabla u(x)|^2 - f(x)u(x)\right] dx + \int_{\partial\Omega} j(Bu(x)\Big) dS \\ = \;& \min_{v\in M} \Big\{ \textstyle\int_\Omega \left[\frac{1}{2}|\nabla v(x)|^2 - f(x)v(x)\right] dx + \int_{\partial\Omega} j(Bv(x)) dS \Big\}\,. \end{aligned}$$

Schließlich gilt, daß es zu jedem $f \in L^2(\Omega)$ ein eindeutiges u gibt, das diese beiden äquivalenten Eigenschaften hat.

Beweis: Nach Satz 9.3 ist a.) äquivaltent zu $f \in \partial\tilde{V}(u)$, und nach Lemma 5.14 ist dies äquivalent zu

$$(*) \qquad \tilde{V}(u) - (f, u)_\Omega = \min_{v\in L^2(\Omega)} \left[\tilde{V}(v) - (f, v)_\Omega\right].$$

Nach Definition ist aber $\tilde{V}(v) = \infty$ für $v \in L^2(\Omega)\backslash M^* \supseteq L^2\backslash M$, also ist $(*)$ äquivalent zu $u \in M$ und

$$\tilde{V}(u) - (f, u)_\Omega = \min_{v\in M} \left[\tilde{V}(v) - (f, v)\right].$$

Dies ist äquivalent zu b.) .
Nach Satz 9.2 und Lemma 9.4 ist $\tilde{V}$ konvex, von unten halbstetig und koerzitiv. Also nimmt nach Lemma 5.7 und Satz 5.8 $\tilde{V}(v) - (f,v)_\Omega$ das Minimum an in wenigstens einem $u \in L^2(\Omega)$. Dieses u hat also die Eigenschaft b.) und damit auch a.) . Seien u und u' zwei Elemente mit diesen Eigenschaften. Dann sind $u, u' \in M^*$, also sind u und u' Minima von $[\tilde{V}(v) - (f,v)_\Omega]_{v \in M^*}$. Nach Satz 9.2 und Lemma 5.3 ist dieses Funktional aber strikt konvex, also sind Minima nach Satz 5.8 eindeutig und somit gilt $u = u'$.

9.7 Bemerkung. Das eindeutig bestimmte u aus Folgerung 9.6 ist die eindeutige Lösung des Hindernisproblems aus Abschnitt 1.2 . Analog zum Vorgehen in Kapitel 7 und wie in 1.2 skizziert kann ein Randwertproblem zur elliptischen partiellen Differentialgleichung

$$-\Delta u = f$$

gefunden werden, für das u Lösung ist. In 1.2 wurde schon gezeigt, daß u nur in den Bereichen von Ω Lösung dieser Gleichung ist, wo es nicht auf dem Hindernis aufliegt. Diese Bereiche hängen von der Lösung ab und sind daher zunächst nicht bekannt, müssen vielmehr gleichzeitig mit der Lösung mitbestimmt werden. Solche Probleme heißen freie Randwertprobleme. Mit solchen freien Randwertproblemen sind interessante Fragen verbunden. Zum Beispiel kann man untersuchen, ob der Rand der Menge, in der u auf dem Hindernis aufliegt, glatt oder sogar analytisch ist.

Lehrbücher

[1] D. Agmon: Lectures on Elliptic Boundary Value Problems. Van Nostrand, New York, 1965.

[2] H.W. Alt: Lineare Funktionalanalysis. Berlin, Springer, 1985.

[3] H. Brezis: Operateurs maximaux monotones et semi–groupes de contractions dans les espaces de Hilbert. North–Holland, Amsterdam, 1973.

[4] B. Dacorogna: Direct methods in the calculus of variations. Springer, 1989.

[5] B. Dacorogna: Introduction au calcul des variations. Presse Polytechniques et Universitaires Romandes, Lausanne, 1992.

[6] K. Deimling: Nonlinear Functional Analysis. Springer, Berlin, 1985.

[7] G. Duvaut; J.L. Lions: Inequalities in mechanics and physics. Springer, Berlin, 1976.

[8] L. C. Evans: Weak convergence methods for nonlinear partial differential equations. Regional Conference Series in Mathematics 74. Amer. Math. Soc., 1990.

[9] S. Fučik; J. Nečas; V. Souček: Einführung in die Variationsrechnung. Teubner, Leipzig, 1977.

[10] D. Gilbarg; N.S. Trudinger: Elliptic partial differential equations of second order. Springer, Berlin, 1977.

[11] L. Hörmander: The analysis of linear partial differential operators. Vol. 1–4. Springer, Berlin, 1985.

[12] D. Kinderlehrer; G. Stampacchia: An introduction to variational inequalities and their applications. Academic Press, New York, 1980.

[13] E. Klingbeil: Variationsrechnung. BI–Wissenschaftsverlag, Mannheim, 1988.

[14] N. Koshlyakov; M. Smirnov; E. Gliner: Differential equations of mathematical physics. North-Holland, Amsterdam, 1964.

[15] O.A. Ladyzhenskaja: The boundary value problems of mathematical physics. American Math. Society, 1977.

[16] A. Langenbach: Monotone Potentialoperatoren in Theorie und Anwendung. Springer, Berlin, 1977.

[17] R. Leis: Vorlesungen über partielle Differentialgleichungen zweiter Ordnung. Bibliographisches Institut, Mannheim, 1967.

[18] R. Leis: Initial boundary value problems in mathematical physics. Teubner, Stuttgart, 1986.

[19] A. Sommerfeld: Partielle Differentialgleichungen der Physik. 5. Auflage, Akademische Verlagsgesellschaft, Leipzig, 1962.

[20] M. Struwe: Variational methods: Applications to nonlinear partial differential equations and Hamiltonian systems. Springer, 1990.

[21] M Vainberg: Variational method und method of monotone operators in the theory of nonlinear equations. Wiley, Chichester, 1972.

[22] W. Velte: Direkte Methoden der Variationsrechnung. Teubner, Stuttgart, 1976.

[23] E. Zeidler: Vorlesungen über nichtlineare Funktionalanalysis I–III. Teubner, Leipzig, 1978. Erweiterte Fassung: Nonlinear functional analysis and its applications I–IV, Springer, 1985/86