Variationsrechnung

Andreas Klaiber

Zusammenfassung

Dies ist das Skript zu einer zweistündigen Einführung in die Variationsrechnung im Sommersemester 2016. Die Vorlesung besteht aus drei Teilen. Im ersten Teil wird die klassische Methode der Variationsrechnung in einer Raumdimension behandelt, wobei der Inhalt den Quellen [1, 2, 7] entstammt. Im zweiten Teil wird die direkte Methode der Variationsrechnung gemäß [5, Kapitel 8] behandelt. Im kurzen dritten Teil werden Ausblicke angedeutet (Probleme mit Nebenbedingung, Sattelpunktmethoden, Anwendungen).

Vorlesungsplan

1	Einleitung und Inhalt der Vorlesung
2	1D-Probleme: Euler-Lagrange-Gleichung
3	1D-Probleme: Konvexe Integranden, zweite Variation
4	1D-Probleme: Legendre-Bedingung, Jacobi-Theorie 1
5	1D-Probleme: Jacobi-Theorie 2
6	1D-Probleme: Probleme mit Nebenbedingungen; Einführung (E)
7	(E) Unterhalbstetigkeit/Konvexität
8	(E) Existenz/Eindeutigkeit, Schwache Lsg. der ELGl
9	(E) Systeme, Begriffe und Sätze
10	(E) Polykonvexität I
11	(E) Polykonvexität II; Null-Lagrangians und Brouwer
12	(E) Nebenbedingungen, Mountain-Pass-Theorem
13	(E) Anwendung auf travelling waves

Inhaltsverzeichnis

1 Einleitung und Beispiele

1.1 Das Grundproblem der Variationsrechnung

Situation: In allgemeiner Form geht es um folgendes mathematisches Problem. Für $n \in \mathbb{N}$ und $U \subset \mathbb{R}^n$ offen sei die *Lagrange-Funktion*

$$L : U \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}, (x_1, \ldots, x_n, z, p_1, \ldots, p_n) \mapsto L(x, z, p)$$

stetig. Einer Funktion $w : U \to \mathbb{R}$ ordnen wir die reelle Zahl

$$I[w] := \int_U L(x, w(x), \nabla w(x)) \, \mathrm{d}x$$

zu. Also ist I eine Abbildung vom Raum der Funktionen nach $\mathbb{R}$; eine solche Abbildung nennt man *Funktional*.

Aufgabe: Finde eine Funktion u derart, dass $I[u]$ ein Minimum von I liefert. Hinzu kommen Nebenbedingungen, z. B. vorgegebene Randwerte $u|_{\partial U} = g$ oder vorgegebene Regularität $u \in C^1(\bar{U})$.

Alternativ können wir das beschriebene *Variationsproblem* folgendermaßen formulieren: Ist $\mathcal{A}$ eine Menge *zulässiger Funktionen* $w : U \to \mathbb{R}$, dann ist eine Funktion $u \in \mathcal{A}$ derart gesucht, dass

$$I[u] = \min_{w \in \mathcal{A}} I[w]$$

gilt.

Hauptschwierigkeit: Wir bezeichnen mit $J := \{I[w] : w \in \mathcal{A}\} \subset \mathbb{R}$ die Bildmenge von I. Das Variationsproblem ist lösbar, wenn J ein Minimum besitzt. Hierfür muss zunächst J nach unten beschränkt sein, d. h., L bzw. I muss so beschaffen sein, dass $\inf\{I[w] : w \in \mathcal{A}\}$ existiert. Aber das garantiert natürlich noch nicht, dass das Minimum tatsächlich existiert!

Das **Ziel** der Variationsrechnung besteht nun genau darin, für bestimmte Funktionale oder bestimmte Klassen von Funktionalen, die Existenz des Minimums zu zeigen.

1.2 Die klassische Methode der Variationsrechnung

(a) Beispiel: Ebene Geodäten

Wir suchen diejenige Kurve $(x, u(x))$ (als Graph einer Funktion), die $(0, 0)$ und (b, B) verbindet und kürzeste Länge hat.

Gesucht: Minimum von

$$I[u] := \int_0^b \sqrt{1 + u'(x)^2} \, \mathrm{d}x \quad \text{auf} \quad \mathcal{A} := \{u \in C^1([0, b]) : u(0) = 0, u(b) = B\}.$$

(b) Beispiel: Die Brachistochrone

Wir suchen diejenige Kurve $(x, u(x))$ (als Graph einer Funktion), entlang derer ein Massenpunkt in kürzester Zeit von $(0, 0)$ nach (b, B) gleitet. (Wir suchen also eine

„schnellste Rutsche".) Wir benötigen also zunächst eine Formel für die Laufzeit: Über $s = \varphi(t)$ wissen wir $\dot{s} = \dot{\varphi}(t) = v(t) = \sqrt{-2gy(t)}$. Aus der Energieerhaltung und wegen $y(0) = 0, v(0) = 0$ folgt

$$\frac{m}{2}v(t)^2 + mgy(t) = \text{const} = 0.$$

Daraus folgt:

$$T = \varphi^{-1}(L) - \varphi^{-1}(0) = \int_0^L (\varphi^{-1})'(s)\,\mathrm{d}s = \int_0^L \frac{\mathrm{d}s}{\sqrt{-2gy(\varphi^{-1}(s))}} = \int_0^b \sqrt{\frac{1 + (u'(x))^2}{-2gu(x)}}\,\mathrm{d}x,$$

mit der Variablentransformation $s \rightsquigarrow x$ im letzten Schritt.
Gesucht: Minimum von

$$I[u] := \int_0^b \sqrt{\frac{1 + u'(x)^2}{-u(x)}}\,\mathrm{d}x \quad \text{auf} \quad \mathcal{A} := \{u \in C^1((0,b]) \cap C^0([0,b]) : u(0) = 0, u(b) = B, I[u] < \infty\}.$$

(c) *Beispiel: Analytische Mechanik (Pendel)

Bezeichnet $r(t) = (x(t), y(t), z(t)), t \in [t_0, t_1]$, die Bahnkurve eines Massepunktes m, so ist

$$T[r] = \frac{m}{2}\,|\dot{r}(t)|^2$$

die kinetische Energie zur Zeit t. Die auf m wirkende Kraft f besitze eine Potential $V = V(t, r)$, s.d. $f = -\nabla V$. Dann heißt $L := T - V$ Lagrange-Funktion und das Hamiltonsche Prinzip besagt: Die Bahnkurve $r(t)$ ist so beschaffen, dass das Wirkungsintegral

$$I[r] := \int_{t_0}^{t_1} L(t, r(t), \dot{r}(t))\,\mathrm{d}t$$

stationär wird, also bei r ein Extremum oder Sattelpunkt besitzt.

Das gilt entsprechend auch für mehrere Massepunkte. Hierbei müssen nicht kartesische Koordinaten benutzt werden, sondern es können beliebige dem Problem angepasste, sogenannte „verallgemeinerte" Koordinaten verwendet werden (deren Anzahl mit der Zahl der Freiheitsgrade übereinstimmt). Zum Beispiel beim mathematischen Pendel:

$$T = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) = \frac{1}{2}ml^2\dot{\varphi}^2, V = -mgl\cos(\varphi) \Longrightarrow L = \frac{ml^2}{2}\dot{\varphi}^2 + mgl\cos(\varphi).$$

$$I[\varphi] := \int_{t_0}^{t_1} \frac{ml^2}{2}\dot{\varphi}^2 + mgl\cos(\varphi)\,\mathrm{d}t$$

(d) Idee zur Gewinnung von Minimierern

Die betrachteten Probleme führen auf Funktionale der Form

$$I[u] = \int_a^b L(x, u(x), u'(x))\,\mathrm{d}x,$$

wobei $L = L(x, z, p)$ gegeben ist, die über einer gewissen Menge $\mathcal{A}$ zulässiger Funktionen (hier: von $\mathbb{R}$ nach $\mathbb{R}$) einen minimalen (oder allgemeiner: stationären) Wert annehmen sollen, z.B.

$$\mathcal{A} = \{u \in C^1([a, b]) : u(a) = y_0, u(b) = y_1\}.$$

Folgende Idee geht auf Euler und Lagrange zurück: Besitzt I bei einem $u \in C^2([a, b])$ ein Minimum, dann hat die skalare Funktion

$$i(\varepsilon) := I[u + \varepsilon\varphi]$$

für jede „Störung“ $\varphi \in C_0^\infty([a, b])$ ein Minimum. Ist i differenzierbar, so folgt

$$i'(0) = 0 \quad \text{und} \quad i''(0) \geq 0.$$

Wir nehmen nur die erste Bedingung her und untersuchen genauer, was sie nach sich zieht:

$$\begin{aligned} 0 = i'(0) &= \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} \int_a^b L(x, (u + \varepsilon\varphi)(x), (u + \varepsilon\varphi)'(x))\,\mathrm{d}x \\ &= \int_a^b L_z(x, u, u')\varphi + L_p(x, u, u')\varphi'\,\mathrm{d}x \\ &= \int_a^b \left\{ L_z(x, u, u') - \frac{\mathrm{d}}{\mathrm{d}x} L_p(x, u, u') \right\} \varphi\,\mathrm{d}x \end{aligned}$$

An dieser Stelle können wir das sogenannte „Fundamentallemma der Variationsrechnung“ anwenden: *Erfüllt $f \in C([a, b])$ die Gleichung $\int_a^b f\varphi\,\mathrm{d}x = 0$ für alle $\varphi \in C_0^\infty([a, b])$, dann gilt $f \equiv 0$.* Daraus folgt die (zu I gehörige) *Euler-Lagrange-Gleichung*:

$$L_z(x, u(x), u'(x)) - \frac{\mathrm{d}}{\mathrm{d}x} L_p(x, u(x), u'(x)) = 0.$$

Das ist hier eine gewöhnliche Dgl. 2. Ordnung (bzw. ein System).

Der (bezüglich φ lineare) Ausdruck

$$\delta I(y)\varphi := i'(0) = \int_a^b L_z\varphi + L_p\varphi'\,\mathrm{d}x$$

heißt *erste Variation* von I bei u in Richtung φ. Somit ist das Verschwinden der ersten Variation notwendig für die Existenz eines Minimums/Extremums und die Euler-Lagrange-Gleichung ist eine notwendige Bedingung für das Verschwinden der ersten Variation.

Die ausgeführte Rechnung lässt sich zu folgendem Theorem ausbauen.

Theorem 1.1. *Sei $L \in C^2([a, b] \times \mathbb{R} \times \mathbb{R})$ und $u \in C^2([a, b])$ ein Minimierer des Funktionals*

$$I[w] := \int_a^b L(x, w(x), w'(x))\,\mathrm{d}x$$

in der Klasse $\mathcal{A} := \{w \in C^2([a,b]) : w(a) = y_0, w(b) = y_1\}$ *mit gegebenen Randwerten* $y_0, y_1 \in \mathbb{R}$. *Dann erfüllt* u *die Euler-Lagrange-Gleichung*

$$L_z(x, w(x), w'(x)) - \frac{d}{dx} L_p(x, w(x), w'(x)) = 0.$$

Bemerkung. (1) Der Ausdruck $I[w]$ ist definiert für $w \in C^1$, sogar $C^{1,sw}$ (stückweise C^1), aber im Theorem wird $w \in C^2$ gefordert. Es bleibt die Frage, wie man Minimierer finden kann, die nur C^1 sind.

(2) In den Spezialfällen $L = L(x,p)$ und $L = L(z,p)$ lässt sich die ELGl auf eine Gleichung 1. Ordnung reduzieren.

(3) Laut Thm. sind Lösungen der Euler-Lagrange-Gleichung ausgezeichnete Kandidaten für Minimierer. Ob sie das aber wirklich sind, muss mit weiteren Mitteln geprüft werden.

Zu den Beispielen: (1) $L(x,z,p) = \sqrt{1+p^2} \Longrightarrow L_z = 0, L_p = \frac{p}{\sqrt{1+p^2}}$. Also lautet die ELGl:

$$-\frac{d}{dx} \frac{u'(x)}{\sqrt{1 + (u'(x))^2}} = 0 \quad \text{mit} u(0) = 0, u(b) = B.$$

Das liefert $u' \equiv$ const, und mit der Randbedingung folgt $u(x) = \frac{B}{b} x$.

(2) $L(x,z,p) = \frac{\sqrt{1+p^2}}{\sqrt{-z}} \Longrightarrow L_z = \frac{1}{2}\sqrt{1+p^2}(-z)^{3/2}, L_p = \frac{p}{\sqrt{1+p^2}\sqrt{-z}}$. Also lautet die ELGl:

$$\frac{1}{2} \frac{\sqrt{1+(u'(x))^2}}{\sqrt{-u(x)^3}} - \left\{ \frac{u'(x)}{\sqrt{-u(x)}\sqrt{1+(u'(x))^2}} \right\} = 0 \quad \text{mit} u(0) = 0, u(b) = B.$$

Das liefert nach längerer Rechnung einen Zykloidenbogen, in parametrischer Darstellung:

$$x(\tau) = r(\tau - \sin\tau), y(\tau) = -r(1 - \cos\tau) \quad \text{mit} \tau \in [0, \tau_*],$$

wobei τ_* so beschaffen ist, dass $x(\tau_*) = b, y(\tau_*) = B$ gilt.

(3) *Lagrange-Gleichung(en)*: $L_{r_k}(t,r,\dot{r}) - \frac{d}{dt} L_{\dot{r}_k}(t,r,\dot{r}) = 0$. Ist $k > 1$, so handelt es sich um ein System.
Explizit fürs Pendel: $-mgl\sin\varphi - \frac{d}{dt} ml^2\dot{\varphi} = 0$, also $\ddot{\varphi} + \frac{g}{l}\sin\varphi = 0$.

(e) Zusammenfassung

Die klassische Methode der VR besteht in der Herleitung der ELGl und deren Lösung. Das sind die Kandidaten für Extremalen und anschließend ist deren Typ (Minimum, Maximum, Sattelpunkt) zu prüfen. Diese Vorgehensweise ist analog zur Minimierung im $\mathbb{R}^n$: Finde die Lösungen von $\nabla f(x) = 0$ und prüfe anschließend, etwa mit $H_f(x)$, den Typ.

1.3 Die direkte Methode der Variationsrechnung

(a) Euler-Lagrange-Gleichung

Wir betrachten nun allgemeiner $L : \bar{U} \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}, (x, z, p) \mapsto L(x, z, p)$, und das zugehörige Funktional

$$I[u] := \int_U L(x, u(x), \nabla u(x))\, \mathrm{d}x \quad \text{mit} \quad u|_{\partial U} = g.$$

Ist $u \in C^2(U)$ ein Minimierer, dann gilt wieder: Für $\varphi \in C_0^\infty(U)$ hat

$$i(\varepsilon) := I[u + \varepsilon\varphi] \quad \text{ein Minimum bei } \varepsilon = 0,$$

also gilt:

$$\begin{aligned}
0 = i'(0) &= \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} \int_U L(x, (u + \varepsilon\varphi)(x), \nabla(u + \varepsilon\varphi)(x))\, \mathrm{d}x \\
&= \int_U L_z(x, u(x), \nabla u)(x))\varphi(x) + \sum_{k=1}^n L_{p_k}(x, u(x), \nabla u)(x)\varphi_{x_k}(x)\, \mathrm{d}x \\
&= \int_U \varphi(x) \left\{ L_z(x, u(x), \nabla u)(x)) - \sum_{k=1}^n \left(L_{p_k}(x, u(x), \nabla u)(x)\right)_{x_k} \right\} \mathrm{d}x,
\end{aligned}$$

wobei wir verwendet haben, dass φ kompakten Träger hat (also keine Randterme) und dass u regulär genug ist, um partielle Integration zu erlauben. Hieraus erhalten wir die zu I gehörige *Euler-Lagrange-Gleichung*:

$$L_z(x, u(x), \nabla u)(x)) - \sum_{k=1}^n \left(L_{p_k}(x, u(x), \nabla u)(x)\right)_{x_k} = 0.$$

(b) Beispiel: Das Dirichlet-Integral

Betrachte

$$I[u] := \int_U \frac{1}{2} |\nabla u(x)|^2 - f(x)u(x)\, \mathrm{d}x \quad \text{mit Randwert } u|_{\partial U} = g,$$

also $L(x, z, p) = \frac{1}{2}|p|^2 - f(x)z$ und somit ELGl:

$$-\Delta u = f, \quad u|_{\partial U} = g.$$

(c) *Beispiel: Poisson-Gleichung

Betrachte

$$I[u] := \int_U \frac{1}{2} |\nabla u(x)|^2 - F(u(x))\, \mathrm{d}x \quad \text{mit Randwert } u|_{\partial U} = g,$$

also $L(x, z, p) = \frac{1}{2}|p|^2 - F(z)$ und somit ELGl (wobei $F' = f$):

$$-\Delta u = f(u), \quad u|_{\partial U} = g.$$

(d) Beispiel: Minimalflächen

Ist eine Fläche F als Graph über $U \subset \mathbb{R}^n$ gegeben, d.h.

$$F = \{(b, \ldots, x_n, u(b, \ldots, x_n)) \in \mathbb{R}^{n+1} : (b, \ldots, x_n) \in U\},$$

dann ist der (verallgemeinerte) Flächeninhalt:

$$I[u] := \int_U \sqrt{1 + |\nabla u(x)|^2} \, \mathrm{d}x.$$

Außerdem sei eine Randbedingung $u|_{\partial U} = g$ gegeben. Ein lokaler Minimierer von I heißt *Minimalfläche*. Die ELGl lautet:

$$\nabla \cdot \left(\frac{\nabla u}{\sqrt{1 + |\nabla u(x)|^2}} \right) = 0, u|_{\partial U} = g.$$

Die linke Seite ist „$n \times$ mittlere Krümmung“, also besagt die Gleichung, dass die mittlere Krümmung verschwindet.

(e) Idee zur Gewinnung von Minimierern

Es gilt weiterhin, dass die ELGl hinreichend glatte Kandidaten für Minimierer liefert. Im Gegensatz zum eindimensionalen Fall ist die ELGl jetzt aber eine partielle Dgl., deren Lösbarkeit i. A. offen ist. Deshalb hat sich folgende Idee entwickelt (Hilbert): Finde Voraussetzungen an I, sodass man *direkt* auf die Existenz eines Minimierers schließen kann. Wenn dieser auch noch genügend regulär ist, dann erfüllt er (sogar) die ELGl.

Vergleich mit der Situation für Minimierer von $f : K \subset \mathbb{R}^n \to \mathbb{R}$, wobei f stetig, K kompakt: (Die klassische Methode entspricht dem Lösen von $\nabla f = 0$, siehe oben.) In diesem Fall können wir anhand der folgenden drei Schritte einen Minimierer konstruieren.

(I) **Existenz einer Minimalfolge**
$f(K) \subset \mathbb{R}$ ist nach unten beschränkt, also existiert $m := \inf_{x \in K} f(x)$.
Wähle eine *Minimalfolge* $(x_n)_{n \in \mathbb{N}}$ mit $f(x_n) \to m$, die nach Definition des Infimums existiert.

(II) **Existenz einer konvergenten Teilfolge („Kompaktheit“)**
Wegen $(x_n) \subset K$ und K beschränkt, ist (x_n) beschränkt. Nach Bolzano-Weierstraß exisitert daher eine konvergente Teilfolge $(x_{n_k})_{k \in \mathbb{N}}$ mit $x_{n_k} \to x_*$. Da K abgeschlossen, ist $x_* \in K$.

(III) **Limes ist Minimierer**
Aus der Stetigkeit von f folgt: $f(x_*) = \lim_{k \to \infty} f(x_{n_k}) = m$. Somit haben wir einen Minimierer $x_* \in K$ konstruiert.

Dieses Programm lässt sich auf Funktionale $I : D \subset X \to \mathbb{R}$ übertragen, sofern die drei fett gedruckten Behauptungen wahr sind. Dafür sind weitreichende Modifikationen nötig, insbesondere eine geschickte Wahl des zugrundeliegenden Raums X. Wir formulieren Bedingungen, die das ermöglichen.

(I) **Existenz einer Minimalfolge**
Gilt, wenn das Funktional nach unten beschränkt ist, d. h. $I(x) \geq c$

(II) **Existenz einer (schwach) konvergenten Teilfolge**
Wenn z. B. X ein Hilbert-Raum ist und die Minimalfolge beschränkt ist, dann erhält man eine schwach(!) konvergente Teilfolge. Die Beschränktheit der Minimalfolge erhält man etwa, wenn I *koerziv* ist: Aus $\|x_n\| \to \infty$ folgt $I[x_n] \to \infty$.

(III) **Limes ist Minimierer**
Dass der schwache Limes der Teilfolge tatsächlich Minimierer ist, bekommt man aus der *schwachen Unterhalbstetigkeit*:

$$x_n \overset{n\to\infty}{\rightharpoonup} x_* \Longrightarrow \liminf_{n\to\infty} I[x_n] \geq I[x_*].$$

I hat diese Eigenschaft, wenn die Lagrange-Funktion konvex bezüglich p ist.

1.4 Inhalt der Vorlesung

In der Vorlesung werden wir die klassische Methode und die direkte Methode der Variationsrechnung ausführlich kennenlernen.

Im ersten Abschnitt beschäftigen wir uns mit eindimensionalen Problemen und ihrer Lösung anhand der klassichen Methode. Wir werden notwendige und hinreichende Bedingungen (etwa nach Legendre, Jacobi) für die Existenz von Minimierern untersuchen; dafür brauchen wir folgende Konzepte: Euler-Lagrange-Gleichung, erste und zweite Variation, Jacobi-Felder, konjugierte Punkte. Schließlich kommen wir auf Probleme mit Nebenbedingungen und auf natürliche Randbedingungen zu sprechen.

Im zweiten Abschnitt behandeln wir ([5, Kap. 8] folgend) die direkte Methode zur Lösung des Minimierungsproblems in L^p für konvexe Lagrange-Funktionen und zeigen Existenz und Eindeutigkeit von Minimierern sowohl für skalare Probleme als auch für Systeme (d. h. die Euler-Lagrange-Gleichung ist eine System von part. Dgl.). Dabei bedienen wir uns einiger Resultate aus der Funktionalanalysis und der Theorie der Sobolev-Räume. Wir betrachten auch hier die Behandlung von Nebenbedingungen (anhand typischer Beispiele).

Im kurzen dritten Abschnitt betrachten wir die Existenz von Sattelpunkten (Mountain-Pass-Theorem) sowie die Anwendung der Variationsrechnung im Kontext von travelling waves.

2 Variationsrechnung in 1D

2.1 Gegenstand des Abschnitts

Die Untersuchung von eindimensionalen Variationsproblemen ist in verschiedenen Aspekten einfacher als die allgemeine Theorie – in etwa so, wie es auch bei gewöhnlichen Dgl im Unterschied zu partiellen der Fall ist. In dieser Situation werden wir notwendige und hinreichende Bedingungen für die Existenz von Minimierern kennenlernen und auf klassische Probleme anwenden.

Im Folgenden betrachten wir – sofern nicht anders angegeben – das Funktional $I : C^1([a,b]) \to \mathbb{R}$, gegeben durch

$$I[w] := \int_a^b L(x, w(x), w'(x))\,\mathrm{d}x \quad \text{mit } a < b,$$

wobei $L : [a,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ zweimal stetig partiell differenzierbar bezüglich jedes Arguments sei. Außerdem geben wir zusätzlich die Randwerte $w(a) = A$ und $w(b) = B$ vor. Wir suchen also

$$I[u] = \min_{w \in \mathcal{A}} I[w] \qquad \text{mit } \mathcal{A} := \left\{ w \in C^1([a,b]) : w(a) = A, w(b) = B \right\}.$$

Man spricht vom *nichtparametrischen eindimensionalen Variationsproblem mit festen Endpunkten.* Wir werden im Folgenden notwendige und hinreichende Kriterien für die Existenz eines Minimierers angeben. Zunächst wiederholen wir die ELGl. Dann betrachten wir die zweite Variation und notwendige Kritierien für einen Minimierer. Danach arbeiten wir auf eine praktische hinreichende Bedingung von Jacobi hin, für die wir den Begriff der konjugierten Punkte brauchen. Schließlich betrachten wir Probleme, deren Randpunkte nicht fixiert sind, sowie Probleme mit Nebenbedingungen.

2.2 Erste Variation und Existenz von Extremalen

(a) Minimierer und Extremalen in C^1

Auf $C^1([a,b])$ betrachten wir die Norm

$$\|w\|_{C^1} := \sup_{x \in [a,b]} |w(x)| + \sup_{x \in [a,b]} \left|w'(x)\right|.$$

Dementsprechend ist die folgende Definition einer Minimumstelle naheliegend.

Definition. u heißt *(strikter) lokaler Minimierer* von I, wenn gilt:

$$\exists \varepsilon > 0 \forall w \in \mathcal{A} : \|u - w\|_{C^1} \leq \varepsilon \Longrightarrow I[u] \leq I[w] \quad (\text{bzw. } I[u] < I[w] \text{ für } w \neq u).$$

In der Einleitung und Präsenzübung wurde behandelt: Ist $u \in \mathcal{A} \cap C^2$ lokaler Minimierer, dann erfüllt u die ELGl. Im Beweis zeigt man zunächst, dass für die *erste Variation von I in Richtung φ*

$$\delta I[u]\varphi := \int_a^b L_z(x, u, u')\varphi + L_p(x, u, u')\varphi'\,\mathrm{d}x = 0 \tag{2.1}$$

gilt; durch partielle Integration und das Fundamentallemma ergibt sich die ELGl. Die linke Seite ist bereits wohldefiniert, wenn lediglich $L \in C^1$ und $u \in C^1$ gilt. Daher nennt man (2.1) auch *schwache Formulierung der ELGl.*

Wir zeigen, dass man unter der schwächeren Voraussetzung $u \in \mathcal{A}$ dennoch die (starke) ELGl erhält. Zu diesem Zweck brauchen wir eine stärkere Version des Fundamentallemmas, die genau für die Situation der schwachen ELGl geeignet ist.

Lemma 2.1 (Lemma von DuBois-Reymond)**.** *Gilt für $f, g \in C([a,b])$ die Gleichung*

$$\int_a^b f(x)\varphi(x) + g(x)\varphi'(x)\,\mathrm{d}x = 0 \quad \textit{für alle } \varphi \in C_0^1([a,b]),$$

dann ist $g \in C^1([a,b])$ und es gilt $g' \equiv f$.

Bemerkung. Die folgende Überlegung kommt nun öfter vor: Es genügt, die Gleichung für $\varphi \in C_0^\infty([a,b])$ zu fordern. Für $\tilde{\varphi} \in C_0^1([a,b])$ finden wir anhand der üblichen Approximation (durch Faltung mit Glättungskernen) eine Folge $\varphi_n \in C_0^\infty([a,b])$ mit $\|\tilde{\varphi} - \varphi_n\|_{W^{1,p}} \to 0$ Gilt die Gleichung für alle $\varphi \in C_0^\infty([a,b])$, können wir schließen, dass sie für ganz $C_0^1([a,b])$ gilt.

Beweis. **1. Schritt:** Setze

$$F(x) := \int_a^x f(\xi)\,\mathrm{d}\xi.$$

Dann ist $F \in C^1$ und $F' = f$. Somit können wir partiell integrieren und erhalten

$$\int_a^b f\varphi\,\mathrm{d}x = \int_a^b F'\varphi\,\mathrm{d}x = -\int_a^b F\varphi'\,\mathrm{d}x$$

Also gilt

$$0 = \int_a^b f\varphi + g\varphi'\,\mathrm{d}x = \int_a^b F'\varphi + g\varphi'\,\mathrm{d}x = \int_a^b (-F+g)\varphi'\,\mathrm{d}x.$$

Der Beweis ist komplett, wenn wir von hier auf $-F + g = \mathrm{const}$ schließen könnten. Das zeigen wir als nächstes.

2. Schritt: Wir zeigen: *Gilt für $h \in C([a,b])$ die Gleichung*

$$\int_a^b h(x)\varphi'(x)\,\mathrm{d}x = 0 \quad \textit{für alle } \varphi \in C_0^1([a,b]), \tag{2.2}$$

dann ist $h \equiv \mathrm{const}$. Setze

$$c := \frac{1}{b-a}\int_a^b h(\xi)\,\mathrm{d}\xi \quad \text{und} \quad \varphi(x) := \int_a^x h(\xi) - c\,\mathrm{d}\xi.$$

Dann ist $\varphi \in C^1$ und $\varphi(a) = \varphi(b) = 0$, d.h. $\varphi \in C_0^1([a,b])$, und $\varphi' = h - c$. Also gilt

$$\int_a^b (h-c)\varphi'\,\mathrm{d}x = \int_a^b h\varphi'\,\mathrm{d}x - c\int_a^b \varphi'\,\mathrm{d}x = 0$$

wegen (2.2) und Randbedingungen von φ. Andererseits gilt

$$0 = \int_a^b (h-c)\varphi' \, \mathrm{d}x = \int_a^b (h-c)^2 \, \mathrm{d}x,$$

also $h \equiv c$ aufgrund der Stetigkeit von h.

Anwendung von Schritt 2 auf $h := g - F$ aus dem ersten Schritt beendet den Beweis, denn $g' = F' = f$. □

Bemerkung. Bei der Aussage handelt es sich um einen Regularitätssatz: Es wird $g \in C^0$ vorausgesetzt und auf $g \in C^1$ geschlossen.

Wir können somit die Euler-Lagrange-Gleichung bereits unter der Voraussetzung $u \in C^1$ ableiten.

Satz 2.2. *Sei $L \in C^1([a,b] \times \mathbb{R} \times \mathbb{R})$ und $u \in \mathcal{A}$ ein Minimierer des Funktionals*

$$I[w] := \int_a^b L(x, w(x), w'(x)) \, \mathrm{d}x.$$

Dann gilt $L_p(x, u(x), u'(x)) \in C^1([a,b])$ und u erfüllt die Euler-Lagrange-Gleichung

$$L_z(x, u(x), u'(x)) = \frac{d}{dx} L_p(x, u(x), u'(x)). \tag{2.3}$$

Beweis. Analog zur Einleitung betrachte $i(\varepsilon) := I[u + \varepsilon\varphi]$, diesmal für $\varphi \in C_0^1([a,b])$. Es gilt $i'(0) = 0$ und daraus folgt:

$$\int_a^b L_z(x,u,u')\varphi + L_p(x,u,u')\varphi' \, \mathrm{d}x = 0 \quad \forall \varphi \in C_0^1([a,b]).$$

Hier können wir Lemma 2.1 anwenden und erhalten $L_p(x,u,u') \in C^1$ und

$$L_z(x,u,u') = \frac{d}{dx} L_p(x,u,u').$$

□

Bemerkung. (1) Für $u \in C^2$ kann man auf der rechten Seite der ELGl die Kettenregel anwenden; sonst nicht.

(2) Die Voraussetzungen können noch beträchtlich abgeschwächt werden, d.h. u kann weniger regulär sein, z.B. stückweise C^1, Lipschitz-stetig, absolutstetig, schwach differenzierbar. In diesen Fällen gilt die ELGl immer noch in geeignetem Sinn, z.B. stückweise für $u \in C^{1,stw}$. Das Beispiel $L(x,z,p) = (1-p^2)^2$ zeigt, dass es wirklich nicht genügt in C^2 zu arbeiten, um unter den Lösungen der ELGl Minimierer zu finden.

Definition. Wir nennen $u \in \mathcal{A}$ eine *(C^1-)Extremale bzw. C^2-Extremale*, wenn u die Euler-Lagrange-Gleichung erfüllt und in C^1 bzw. C^2 liegt.

Beispiel (Eulers Paradoxon)**.** Für $I[w] = \int_0^1 (1-(w')^2)^2 \,\mathrm{d}x$ mit $w(0) = w(1) = 0$ ist die zugehörige ELGl:

$$0 = \frac{d}{dx}\left((1-(u')^2)u'\right), \quad \text{also: } (1-(u')^2)u' = c.$$

Bezeichnen t_1, t_2, t_3 die Lösungen von $(1-t^2)t = c$ zu festem c, dann ist jede stückweise lineare Funktion mit $u'(x) \in \{t_1, t_2, t_3\}$ eine Extremale. Da $I[w] \geq 0$ und $I[w] = 0$ genau dann, wenn $|u'(x)| = 1$ f.ü., muss für einen Minimierer $c = 0$ gelten. Dann ist $t_1 = -1, t_2 = 0, t_3 = 1$. Ist $u'(x) = 0$ auf einem Teilintervall, dann ist $I > 0$, also kein Minimierer. Deshalb sind nur diejenigen Funktionen mit $|u'(x)| = 1$ f.ü. Minimierer. Insbesondere gibt es keinen C^1-Minimierer. Man erkennt hierbei, dass sogar Minimierer aus $C^{1,stw}$ die ELGl stückweise erfüllen. Funktionen, die nicht in C^1 liegen, aber die ELGl in geeignetem Sinne erfüllen, nennt man *gebrochene Extremalen*.

Beispiel (Gegenbeispiel von Weierstraß)**.** Bei den Zeitgenossen von Weierstraß wurde üblicherweise angenommen, dass nach unten beschränkte Funktionale auch ein Minimum besitzen. Weierstraß widerlegte dies mit folgendem Beispiel. Betrachte $I[w] = \int_{-1}^1 x^2(w'(x))^2 \,\mathrm{d}x$ auf $\mathcal{A} := \{w \in C^1([-1,1]) : w(-1) = -1, w(1) = 1\}$. Dann gilt $I[w] \geq 0$ für alle $w \in \mathcal{A}$ und für die Folge

$$w_n(x) := \frac{\arctan(nx)}{\arctan(n)} \in \mathcal{A}$$

gilt:

$$I[w_n] = \int_{-1}^1 \frac{n^2x^2}{\arctan(n)^2(1+n^2x^2)^2} \,\mathrm{d}x < \frac{1}{\arctan(n)^2} \int_{-1}^1 \frac{\mathrm{d}x}{(1+n^2x^2)} \,\mathrm{d}x = \frac{2}{n\arctan(n)} \to 0.$$

Somit ist $\inf_{w\in\mathcal{A}} I[w] = 0$. Es gibt jedoch keine Funktion $u \in \mathcal{A}$, die das Infimum annimmt. Eine solche Funktion müsste $xu'(x) = 0$ erfüllen, also $u'(x) = 0$ für $x \neq 0$. Aus den Randbedingungen ergibt sich $u(x) = -1$ für $x \in [-1, 0)$ und $u(x) = 1$ für $x \in (0, 1]$. Eine solche Funktion gibt es nicht in $\mathcal{A}$!

Die ELGl lautet: $x^2u' = c_1$ mit den Lösungen $u(x) = -\frac{c_1}{x} + c_2$ (wobei $c_1 \neq 0, c_2$ sich aus den Randbedingungen ergeben). Keine dieser Lösungen liegt in $\mathcal{A}$.

Unter einer einfachen Voraussetzung an L erhält man automatisch C^2-Regularität des Minimierers. Das zeigt der folgende Satz.

Satz 2.3. *Angenommen, u ist C^1-Extremale von I, $L_p \in C^1([a,b] \times \mathbb{R} \times \mathbb{R})$ und $L_{pp}(x, u(x), u'(x)) \neq 0$ für alle $x \in [a,b]$. Dann ist $u \in C^2([a,b])$.*

Beweis. Nach Voraussetzung erfüllt $u \in C^1([a,b])$ die Euler-Lagrange-Gleichung, also ist

$$f \equiv L_p(\cdot, u, u') \in C^1([a,b]).$$

Wir definieren $G : [a,b] \times \mathbb{R} \to \mathbb{R}, (x,p) \mapsto L_p(x, u(x), p) - f(x)$. Nach den Voraussetzungen an L ist G somit stetig differenzierbar und es gilt bei jedem $x_0 \in [a,b]$:

$$G(x_0, u'(x_0)) = 0, \quad G_p(x_0, u'(x_0)) = L_{pp}(x_0, u(x_0), u'(x_0)) \neq 0.$$

In dieser Situation liefert der Satz über implizite Funktionen, dass es zu kleinem $\delta > 0$ genau eine Funktion $p \in C^1((x_0 - \delta, x_0 + \delta))$ gibt mit $G(x, p(x)) = 0$ und $p(x_0) = u'(x_0)$. Andererseits gilt $G(x, u'(x)) = 0$ (ELGl), sodass $p \equiv u'$ ist und $u' \in C^1$ bedeutet $u \in C^2$. Da x_0 beliebig und $[a, b]$ kompakt ist, können wir p auf ganz $[a, b]$ fortsetzen. □

Bemerkung. Die Voraussetzung $L_{pp} \neq 0$ ist eine Elliptizitätsbedingung in einfachster Form. Wir werden sie später wieder antreffen.

(b) Konvexe Integranden

Zur Erinnerung wiederholen wir einige Fakten zur Konvexität.

(i) Eine Menge M heißt *konvex*, wenn gilt:

$$\forall x, y \in M, t \in [0, 1] : \quad tx + (1 - t)y \in M.$$

(ii) Sei M konvex. Eine Funktion $f : M \to \mathbb{R}$ heißt *konvex*, wenn gilt:

$$\forall x, y \in M, t \in [0, 1] : \quad f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y).$$

(iii) f heißt *strikt konvex*, wenn gilt:

$$\forall x, y \in M, x \neq y, t \in (0, 1) : \quad f(tx + (1 - t)y) < tf(x) + (1 - t)f(y).$$

(iv) Falls $f \in C^1(\mathbb{R}^n)$, so gilt: f ist genau dann konvex, wenn

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

(v) Falls $f \in C^2(\mathbb{R}^n)$, so gilt: f ist genau dann konvex, wenn die Hessematrix $H_f(x)$ für jedes $x \in \mathbb{R}^n$ positiv semidefinit ist, d.h., für deren Eigenwerte: $\lambda_1, \ldots, \lambda_n \geq 0$.
Alternativ: Für jedes x sind alle Hauptminoren von $H_f(x)$ größer gleich Null.
Für $n = 2$ heißt das: $f_{xx} \geq 0, f_{yy} \geq 0, f_{xx}f_{yy} - f_{xy}^2 \geq 0$.

Für den Fall konvexer Integranden können wir auch hinreichende Bedingungen an die Existenz angeben, weil sich die Konvexität auf das Funktional überträgt. Hierfür nehmen wir wieder C^2-Regularität von L an.

Satz 2.4. *Sei $L \in C^2([a, b] \times \mathbb{R} \times \mathbb{R})$.*

(i) *Sei u C^1-Extremale und sei*

$$L(x, \cdot, \cdot) : (z, p) \mapsto L(x, z, p) \quad \text{konvex für jedes } x \in [a, b],$$

dann ist u lokaler Minimierer von I.

(ii) *Ist $L(x, \cdot, \cdot)$ strikt konvex für jedes x, dann besitzt I höchstens einen lokalen Minimierer. Falls dieser existiert, ist er der globale Minimierer.*

Beweis. (i) Wenn $(\mu,\nu) \mapsto L(x,\mu,\nu)$ konvex ist, so gilt für alle $\mu,\nu \in \mathbb{R}$:

$$L(x,\mu,\nu) \geq L(x,u(x),u'(x)) + L_z(x,u(x),u'(x))(\mu - u(x)) + L_p(x,u(x),u'(x))(\nu - u'(x)).$$

Daraus folgt:

$$\begin{aligned}
I[w] &= \int_a^b L(x,w(x),w'(x))\,\mathrm{d}x \\
&\geq \int_a^b L(x,u(x),u'(x)) + L_z(x,u(x),u'(x))(w(x)-u(x)) \\
&\qquad + L_p(x,u(x),u'(x))(w'(x)-u'(x))\,\mathrm{d}x \\
&= I[u] + \int_a^b \left[L_z(x,u(x),u'(x)) - \frac{d}{dx} L_p(x,u(x),u'(x)) \right] (w(x)-u(x))\,\mathrm{d}x \\
&\qquad + L_p(b,u(b),u'(b))(w(b)-u(b)) - L_p(a,u(a),u'(a))(w(a)-u(a)) \\
&= I[u],
\end{aligned}$$

da u (2.3) erfüllt und u,w die gleichen Randwerte haben. Somit ist u ein Minimum von I.

(ii) Sei $u \in \mathcal{A}$ lokaler Minimierer von I. Dann existiert eine Umgebung W von u in C^1, sodass mit $I[u] \leq I[w]$ für alle $w \in W$ gilt. Wir zeigen nun, dass $I[u] < I[w]$ für $u \neq w$ gilt. Für $t \in (0,1)$, $w \in \mathcal{A}$ setze

$$w^t := tw + (1-t)u.$$

Aus der strikten Konvexität folgt

$$L(x,w^t,(w^t)') = L(x,tw+(1-t)u,tw'+(1-t)u') < tL(x,w,w') + (1-t)L(x,u,u'),$$

sofern $(u(x),u'(x)) \neq (w(x),w'(x))$. Nach Voraussetzung gilt $u \neq w$, daher existiert ein Teilintervall von $[a,b]$, auf dem $u(x) \neq w(x)$ gilt. Also folgt durch Integration:

$$\begin{aligned}
I[w^t] &= \int_a^b L(x,w^t,(w^t)')\,\mathrm{d}x \\
&< t\int_a^b L(x,w,w')\,\mathrm{d}x + (1-t)\int_a^b L(x,u,u')\,\mathrm{d}x \\
&= tI[w] + (1-t)I[u].
\end{aligned}$$

Angenommen, es wäre nun $I[w] \leq I[u]$, dann folgt für alle $t \in (0,1)$:

$$I[w^t] < tI[w] + (1-t)I[u] \leq tI[u] + (1-t)I[u] = I[u].$$

Aufgrund von $w^t \to u$ für $t \to 0$ existiert jedoch ein $t_0 > 0$ mit $v^t \in W$ für alle $0 < t < t_0$. Für diese t muss nach Voraussetzung $I[w^t] \geq I[u]$ gelten – im Widerspruch zur Ungleichung $I[w^t] < I[u]$. Also gilt tatsächlich $I[w] > I[u]$. □

Beispiel (Geodäte in $\mathbb{R}^2$). $L(x,z,p) = \sqrt{1+p^2}$, also $L_{pp} = (1+p^2)^{-3/2} > 0, L_{pz} = L_{zz} = 0$, somit ist $L(x,\cdot,\cdot)$ konvex (aber nicht streng konvex). Somit ist die Gerade lokaler Minimierer. Da es der einzige Kandidat in C^1 ist, handelt es sich um einen globalen Minimierer.

Bemerkung. Es ist mitunter möglich, dass man durch eine Variablentransformation zu einem konvexen Integranden gelangt. Das ist etwa bei der Brachistochrone der Fall. Hier ist $L(x,u,u') = (-u)^{-1/2}\sqrt{1+(u')^2}$. Die Transformation $v := \sqrt{-2u}$ führt auf die strikt konvexe Lagrange-Funktion

$$\tilde{L}(x,v,v') = \sqrt{2}\left(\frac{1}{v^2} + (v')^2\right)^{1/2}.$$

Siehe Übung.

2.3 Zweite Variation und Typ der Extremalen

(a) Definition der zweiten Variation und Eigenschaften

In der Einleitung haben wir als notwendige Bedingung für die Existenz eines Minimierers u von I hergeleitet, dass $i(\varepsilon) := I[u+\varepsilon\varphi]$ bei $\varepsilon = 0$ lokales Minimum bei $\varepsilon = 0$ hat. Für $i \in C^2$ heißt das $i'(0) = 0$ und $i''(0) \geq 0$. Aus der ersten Bedingung haben wir die ELGl gewonnen; nun schauen wir uns die zweite Bedingung an. Formal gilt:

$$\begin{aligned}
i''(0) &= \frac{d^2}{d\varepsilon^2}\bigg|_{\varepsilon=0} I[u+\varepsilon\varphi] \\
&= \frac{d^2}{d\varepsilon^2}\bigg|_{\varepsilon=0} \int_a^b L(x,(u+\varepsilon\varphi)(x),(u+\varepsilon\varphi)'(x))\,\mathrm{d}x \\
&= \int_a^b \frac{d^2}{d\varepsilon^2}\bigg|_{\varepsilon=0} L(x,(u+\varepsilon\varphi)(x),(u+\varepsilon\varphi)'(x))\,\mathrm{d}x \\
&= \int_a^b \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \big[L_z(x,(u+\varepsilon\varphi)(x),(u+\varepsilon\varphi)'(x))\varphi(x) + L_p(x,(u+\varepsilon\varphi)(x),(u+\varepsilon\varphi)'(x))\varphi'(x)\big]\,\mathrm{d}x \\
&= \int_a^b L_{zz}\varphi^2 + 2L_{zp}\varphi\varphi' + L_{pp}(\varphi')^2\,\mathrm{d}x.
\end{aligned}$$

Definition. Der (in φ quadratische) Ausdruck

$$\delta^2 I[u]\varphi := \int_a^b L_{zz}\varphi^2 + 2L_{zp}\varphi\varphi' + L_{pp}(\varphi')^2\,\mathrm{d}x \tag{2.4}$$

heißt *zweite Variation* von I bei u in Richtung φ.

Aus $i''(0) \geq 0$ erhalten wir somit folgende notwendige Bedingung.

Lemma 2.5. *Ist $u \in \mathcal{A}$ lokaler Minimierer, dann gilt $\delta^2 I[u]\varphi \geq 0$ für alle $\varphi \in C_0^\infty([a,b])$ (sogar $C_0^1([a,b])$).*

Bemerkung. Man beachte, dass – im Unterschied zu Funktionen $f : \mathbb{R}^n \to \mathbb{R}$ – die Bedingung $\delta^2 I[u]\varphi > 0$ für alle $\varphi \not\equiv 0$, d. h. $\delta^2 I[u]$ *positiv definit*, nicht hinreichend ist. Für ein Beispiel: Siehe Übung.

Satz 2.6. *Sei $X \subset \mathcal{A}$ konvex und $u \in X$ eine C^1-Extremale.*
(i) Wenn $\delta^2 I[v]\varphi \geq 0$ für alle $v \in X$ und $\varphi \in C_0^\infty([a,b])$ gilt, dann ist u Minimierer von I.
(ii) Wenn $\delta^2 I[v]\varphi > 0$ für alle $v \in X$ und $\varphi \in C_0^1([a,b]) \setminus \{0\}$ gilt, dann ist u strikter Minimierer von I.

Bemerkung. Der Satz besagt: (i) Jede Extremale in X liefert das absolute Minimum. (ii) Eine Extremale in X ist der eindeutige Minimierer.

Beweis. Durch Approximation gilt die Ungleichung $\delta^2 I[u]\varphi \geq 0$ für alle $\varphi \in C_0^1([a,b])$. Für beliebiges v nahe u mit den gleichen Randwerten wie u ist dann $\varphi := v-u \in C_0^1([a,b])$ und

$$\psi(t) := u + t\varphi = tv + (1-t)u \in X \quad \forall t \in [0,1].$$

Da $i(t) := I[\psi(t)]$ in C^2 ist und u Extremale ist gilt: $i'(0) = \delta I[u]\varphi = 0$ und somit

$$i(1) - i(0) = i'(0) + \int_0^1 (1-t) i''(t)\,\mathrm{d}t = \int_0^1 (1-t)\delta^2 I[\psi(t)]\varphi\,\mathrm{d}t \geq 0,$$

d. h., u ist Minimierer im Fall (i). Im Fall (ii) erhalten wir sogar > 0, d. h. u ist strikter Minimierer. □

Wenn die zweite Variation *gleichmäßig positiv definit* ist, lässt sich dieser Satz anwenden.

Satz 2.7. *Gilt für eine Extremale u die Ungleichung*

$$\delta^2 I[u]\varphi \geq 2\lambda \left\|\varphi\right\|_{W^{1,2}}^2 \equiv 2\lambda \int_a^b \varphi^2 + (\varphi')^2 \quad \forall \varphi \in C_0^\infty$$

mit einem $\lambda > 0$, dann ist u strikter Minimierer.

Beweis. Durch Approximation gilt die Ungleichung auch für alle $\varphi \in C_0^1$. Für $L \in C^2$ gilt aufgrund der Stetigkeit des Integranden

$$\delta^2 L(u,\varphi) \equiv L_{zz}(x,u,u')\varphi^2 + 2L_{zp}(x,u,u')\varphi\varphi' + L_{pp}(x,u,u')(\varphi')^2$$

von $\delta^2 I[u]\varphi$:

$$\left|\delta^2 L(u,\varphi) - \delta^2 L(v,\varphi)\right| \leq c(\varphi^2 + 2\varphi\varphi' + (\varphi')^2) \leq \lambda(\varphi^2 + (\varphi')^2),$$

sofern $\left\|u - v\right\|_{C^1}$ klein ist. Also folgt:

$$\delta^2 I[v]\varphi \geq \int_a^b -\left|\delta^2 L(u,\varphi) - \delta^2 L(v,\varphi)\right| + \delta^2 I[u]\varphi \geq \lambda \int_a^b \varphi^2 + (\varphi')^2 > 0 \quad \text{für } \varphi \not\equiv 0.$$

Somit können wir Satz 2.6 anwenden und schließen, dass u strikter Minimierer ist. □

(b) Legendre-Bedingung

Da sich Bedingungen an die zweite Variation im Allgemeinen schlecht direkt nachprüfen lassen, haben die Kriterien wenig praktische Bedeutung. Leichter zu prüfen ist die Legendre-Bedingung, die sich von hier aus ableiten lässt. Dafür wird der Integrand der zweiten Variation zunächst als Summe von Quadraten geschrieben, d. h. in der Form

$$\delta^2 I[u]\varphi = \int_a^b P(x)(\varphi'(x))^2 + Q(x)\varphi(x)^2 \,\mathrm{d}x$$

mit geeigneten Ausdrücken $P(x), Q(x)$.

Ausgehend von $2\varphi\varphi' = (\varphi^2)'$, gilt (sofern L_{zp} differenzierbar):

$$\int_a^b 2L_{zp}\varphi\varphi' \,\mathrm{d}x = \varphi^2 L_{zp}|_a^b - \int_a^b \frac{d}{dx}L_{zp}(x,u(x),u'(x))\varphi^2 \,\mathrm{d}x = -\int_a^b \frac{d}{dx}L_{zp}(x,u(x),u'(x))\varphi^2 \,\mathrm{d}x$$

unter Verwendung von $\varphi(a) = \varphi(b) = 0$. Somit lässt sich die zweite Variation auch so schreiben:

$$\delta^2 I[u]\varphi = \int_a^b \left[L_{zz} - \frac{d}{dx}L_{zp}\right]\varphi^2 + L_{pp}(\varphi')^2 \,\mathrm{d}x. \tag{2.5}$$

Das ist die gesuchte Form mit den (in Abhängigkeit von u bekannten) Funktionen

$$P(x) := L_{pp}(x,u(x),u'(x)) \quad \text{und} \quad Q(x) := L_{zz}(x,u(x),u'(x)) - \frac{d}{dx}L_{zp}(x,u(x),u'(x))$$

geschrieben. Durch geeignete Wahl von φ können wir nun weitere Informationen herausbekommen: Wir wählen φ so, dass φ^2 beschränkt bleibt, aber $(\varphi')^2$ groß wird. Für $n \in \mathbb{N}$ leistet dies (SKIZZE!)

$$\varphi_n(x) := \begin{cases} 1 + n(x-x_0) & \text{für } x \in \left[x_0 - \frac{1}{n}, x_0\right], \\ 1 - n(x-x_0) & \text{für } x \in \left[x_0, x_0 + \frac{1}{n}\right], \\ 0 & \text{sonst.} \end{cases} \tag{2.6}$$

Dann ist

$$\varphi_n'(x) = \begin{cases} n & \text{für } x \in \left[x_0 - \frac{1}{n}, x_0\right], \\ -n & \text{für } x \in \left[x_0, x_0 + \frac{1}{n}\right], \\ 0 & \text{sonst.} \end{cases}$$

Beachte, dass nur $\varphi \in C^{1,stw}$ gilt, aber durch Glättung der Ecken finden wir auch Kandidaten in C_0^∞. (Alternativ könnte man nehmen: $\varphi_n(x) = n^3(1/n^2 - |x-x_0|^2)^2$ für $|x-x_0| \leq 1/n$ und Null sonst oder $\varphi_n(x) = \sin^2(n\pi(x-x_0))$ für $|x-x_0| \leq 1/n$ und Null sonst). Für diese Funktionen gilt

$$\int_a^b \varphi_n(x)^2 \,\mathrm{d}x = \frac{1}{n} \quad \text{und} \quad \int_a^b (\varphi_n'(x))^2 \,\mathrm{d}x = 2n.$$

Damit können wir folgenden Satz beweisen.

Satz 2.8. *Ist u lokaler Minimierer, dann gilt die* Legendre-Bedingung*:*

$$L_{pp}(x, u(x), u'(x)) \geq 0 \tag{2.7}$$

Beweis. Angenommen, es gelte $L_{pp}(x_0, u(x_0), u'(x_0)) < 0$ bei $x_0 \in [a, b]$. Dann gilt (aufgrund der Stetigkeit) $P(x) = L_{pp}(x, u(x), u'(x)) \leq -c_1 < 0$ für alle $x \in [x_0 - \gamma, x_0 + \gamma]$ mit $c_1, \gamma > 0$. Auch aufgrund der Stetigkeit gilt $Q(x) \leq c_2$ für alle $x \in [x_0 - \gamma, x_0 + \gamma]$ (mit möglicherweise kleinerem γ).

Dann gilt mit φ_n wie oben:

$$0 \leq \int_a^b Q(x)\varphi_n^2 + P(x)(\varphi_n')^2 \,\mathrm{d}x \leq c_2 \int_{x_0-\gamma}^{x_0+\gamma} \varphi_n^2 \,\mathrm{d}x - c_1 \int_{x_0-\gamma}^{x_0+\gamma} (\varphi_n')^2 \,\mathrm{d}x = \frac{1}{n}c_2 - 2nc_1 < 0$$

für hinreichend großes n. Dieser Widerspruch beweist die Behauptung. □

Bemerkung. (i) Diese Bedingung ist nützlich, um Extremalen als Minimierer auszuschließen. Betrachte $I[w] := \int_{-1}^1 x\sqrt{1 + (w'(x))^2} \,\mathrm{d}x$. Anhand der ELGl könnten wir Extremalen finden, aber das ist nicht einmal nötig, um Minimierer auszuschließen: Es gilt $L(x, z, p) = x\sqrt{1 + p^2}$, also ist $P(x) \equiv L_{pp}(x, u(x), u'(x)) \equiv \frac{x}{(1+(w'(x))^2)^{3/2}} > 0$ nicht auf $[-1, 1]$ erfüllt.

(ii) Selbst die *starke Legendre-Bedingung*

$$L_{pp}(x, u(x), u'(x)) > 0 \tag{2.8}$$

ist nicht hinreichend dafür, dass u Minimierer ist: Betrachte $I[w] := \frac{1}{2}\int_0^l (w')^2 - w^2$ mit $w(0) = w(l) = 0$. Die ELGl lautet: $w'' = -w$. Es gilt $L_{pp} = 1 > 0$. Andrerseits gilt für jede Extremale u

$$\delta^2 I[u]\varphi = \int_0^l (\varphi')^2 - \varphi^2 \,\mathrm{d}x.$$

Speziell für $\varphi(x) := \sin\left(\frac{\pi}{l}x\right)$ erhalten wir

$$\delta^2 I[u]\varphi = \int_0^l \frac{\pi^2}{l^2}\cos^2\left(\frac{\pi x}{l}\right) - \sin^2\left(\frac{\pi x}{l}\right) \,\mathrm{d}x = \frac{1}{l}(\pi^2 - l^2).$$

Somit können wir aussagen: Für $l > \pi$ ist die Extremale $u \equiv 0$ kein Minimierer, für $l \leq \pi$ ist die Frage offen. Beachte, dass der Integrand indefinit ist.

Um zu einer hinreichenden Bedingung zu kommen, benötigen wir mehr als nur punktweise Informationen, nämlich Aussagen über das globale Verhalten.

2.4 Jacobi-Theorie

(a) Jacobi-Gleichung

Wir werden im Folgenden auf eine hinreichende Bedingung hinarbeiten. Dafür nehmen wir an: u ist C^2-Extremale und erfüllt die starke Legendre-Bedingung (2.8).

Vorüberlegung: Wir bringen eine weitere Umformulierung von $\delta^2 I[u]$ ins Spiel (nach Legendre und Jacobi). Sollte es möglich sein, die zweite Variation zu schreiben als

$$\delta^2 I[u]\varphi = \int_a^b L_{pp}(x, u(x), u'(x)) Y(x)^2 \, \mathrm{d}x = \int_a^b P(x) Y(x)^2 \, \mathrm{d}x$$

mit einer geeigneten Funktion Y, die zu gegebenem u von x und $\varphi(x)$ abhängt, dann ist $\delta^2 I[u]$ sicherlich positiv semidefinit. Dazu folgende Überlegung. Wir wissen bereits:

$$\delta^2 I[u]\varphi = \int_a^b Q(x)\varphi^2 + P(x)(\varphi')^2 \, \mathrm{d}x.$$

Für jede glatte Funktion w liefert die Randbedingung $\varphi(a) = \varphi(b) = 0$:

$$\int_a^b (w\varphi^2)' \, \mathrm{d}x = 0.$$

Folglich können wir im Integranden von $\delta^2 I$ einen solchen Term hinzufügen, ohne den Wert des Integrals zu ändern. Wir versuchen nun, w so zu wählen, dass tatsächlich

$$Q\varphi^2 + P(\varphi')^2 + (w\varphi^2)' = PY^2$$

gilt. Da wir die starke Legendre-Bedingung $P(x) > 0$ voraussetzen, können wir P ausklammern:

$$\text{LHS} = P(\varphi')^2 + Q\varphi^2 + w'\varphi^2 + 2w\varphi\varphi' = P\left((\varphi')^2 + \frac{2w}{P}\varphi\varphi' + \frac{Q + w'}{P}\varphi^2\right)$$

Wenn w die sogenannte *Jacobi-Gleichung*

$$\text{(JGl)} \qquad w^2 = P(Q + w')$$

erfüllt, dann gilt

$$Q\varphi^2 + P(\varphi')^2 + (w\varphi^2)' = P\left(\varphi' + \frac{w}{P}\varphi\right)^2,$$

also ist $Y := \varphi' + \frac{w}{P}\varphi$ der richtige Kandidat für die beabsichtigte Umformulierung.

Man beachte, dass $Y \equiv 0$ nur gilt, wenn $\varphi \equiv 0$ ist, denn: $Y \equiv 0$ bedeutet, φ erfüllt $\varphi' + \frac{w}{P}\varphi = 0$ mit den Randbedingungen $\varphi(a) = \varphi(b) = 0$. Offenbar ist $\varphi \equiv 0$ eine Lösung, die hier (wegen $P > 0$) eindeutig ist. Wir haben somit gezeigt:

Lemma 2.9. *Wenn $P(x) > 0$ für alle $x \in [a, b]$ gilt und eine Lösung w von* (JGl) *auf $[a, b]$ existiert, dann ist $\delta^2 I[u]$ positiv definit.*

(b) Jacobische Hilfsgleichung und Jacobi-Felder

Die Aussage von 2.9 hängt an der Existenz einer Lösung w von (JGl). Dabei handelt es sich um eine Riccati-Gleichung (nichtlinear, 1. Ordnung), die zwar lokal stets eine Lösung besitzt, aber die Existenz einer Lösung auf ganz $[a,b]$ ist unklar. (In der Tat besitzt die Gleichung keine Lösung auf ganz $[a,b]$ für $b-a>\pi$ und $-P=Q=1$, denn die lokale Lösung $w(x)=\tan(c-x)$ besitzt keine globale Fortsetzung.) Leichter zu behandeln ist die lineare Differentialgleichung 2. Ordnung, die durch die Transformation

$$w = -\frac{v'}{v}L_{pp} \equiv -\frac{v'}{v}P \tag{2.9}$$

aus (JGl) hervorgeht, die sogenannte *Jacobische Hilfsgleichung*:

$$-\left(P(x)v'(x)\right)' + Q(x)v(x) \equiv -\frac{d}{dx}(L_{pp}v') + \left(L_{zz} - \frac{d}{dx}L_{zp}\right)v = 0. \tag{JHGl}$$

Hat diese Gleichung eine Lösung v auf $[a,b]$, sodass $v(x)\neq 0$ für alle x gilt, dann erhält man vermöge der Transformation eine Lösung der Riccati-Gleichung (JGl).

Definition. Eine Lösung $v\in C^2([a,b])$ von (JHGl) mit $v\not\equiv 0$ heißt *Jacobi-Feld* entlang u.

Satz 2.10. *Gilt $P\equiv L_{pp}>0$ auf $[a,b]$ und existiert ein Jacobi-Feld v mit $v(x)>0$ auf $[a,b]$, dann ist u ein strikter Minimierer von I.*

Beweis. Sei $\varphi\in C_0^1([a,b])$ beliebig. Ausgehend von v erhalten wir die Lösung $w\equiv -\frac{v'}{v}P$ der Jacobi-Gleichung. Weiter ist

$$Y = \varphi' + \frac{w}{P}\varphi = \varphi' - \frac{v'}{v}\varphi = v\left(\frac{\varphi}{v}\right)',$$

also gilt mit $\psi := \varphi/v \in C_0^1([a,b])$:

$$\begin{aligned}\mathcal{D} := \delta^2 I[u]\varphi = \int_a^b PY^2\,\mathrm{d}x = \int_a^b Pv^2(\psi')^2\,\mathrm{d}x \geq \inf\{Pv^2\}\int_a^b(\psi')^2 \overset{\text{Poincare}}{\geq} \frac{\inf\{Pv^2\}}{(a-b)^2}\int_a^b \psi^2 \\ \geq \frac{\inf\{Pv^2\}}{(a-b)^2}\inf\{1/v^2\}\int_a^b\varphi^2 \equiv \mu_1\int_a^b\varphi^2.\end{aligned}$$

Wir zeigen, dass

$$\mathcal{D} \geq \lambda\int_a^b \varphi^2 + (\varphi')^2\,\mathrm{d}x$$

gilt, und schließen dann mit Satz 2.7.

Es gilt $\mathcal{D} = \int Q\varphi^2 + P(\varphi')^2$, also $P(\varphi')^2 = \mathcal{D} - \int Q\varphi^2$. Mit $\alpha := \sup|Q|$ und $\beta := \inf P > 0$ folgt:

$$\beta\int(\varphi')^2 \leq \mathcal{D}(\varphi) + \alpha\int\varphi^2 \leq (1+\frac{\alpha}{\mu_1})\mathcal{D}(\varphi),$$

also

$$\mathcal{D} \geq \mu_2\int(\varphi')^2 \Longrightarrow \mathcal{D} \geq \frac{\mu_1}{2}\int\varphi^2 + \frac{\mu_2}{2}\int(\varphi')^2$$

und $\lambda := \min\{\mu_1/2, \mu_2/2\}$ beendet den Beweis. □

Bemerkung. Zusammen mit Lemma 2.9 erhalten wir die Aussage: *Wenn $P(x) > 0$ und $\delta^2 I[u]\varphi > 0$, dann ist u strikter Minimierer.* Beachte, dass diese Aussage i.A. falsch ist, sobald auch nur an einer Stelle $P(x) = 0$ gilt. (Siehe Übung, Aufgabe 4)

Beispiel. Für das Funktional $I[w] = \int_a^b w'(x)^2 \, \mathrm{d}x$ ist $L = p^2$. Die JHGl lautet somit $-(2v')' = 0$ und hat die allgemeine Lösung $v(x) = c_1 x + c_2$. Offenbar gibt es ein v mit $v(x) > 0$ in $[a, b]$.

Als nächstes werden wir sehen, dass man die Frage nach einem positiven Jacobi-Feld in die Frage nach sog. konjugierten Punkten überführen kann.

(c) Konjugierte Punkte

Zur Erinnerung: Wir interessieren uns dafür, ob die Jacobi-Gleichung eine globale Lösung w auf ganz $[a, b]$ besitzt. Diese Frage haben wir zurückgeführt auf die Frage, ob die Jacobische Hilfsgleichung eine positive Lösung besitzt. In diesem Zusammenhang spielen Punkte x, bei denen v verschwindet eine ausgezeichnete Rolle, denn dort wird die Transformation (2.9) singulär.

Wir sammeln zunächst einige Fakten über die Lösungen von (JHGl). Aus der Theorie lineare Differentialgleichungen zweiter Ordnung ergibt sich sofort: Zu gegebenen Anfangsdaten $v(x_0) = v_0, v'(x_0) = v_0'$ hat die Gleichung (JHGl) eine globale Lösung auf $[a, b]$. Es existiert ein Fundamentalsystem, sodass $v = \alpha_1 v_1 + \alpha_2 v_2$ für jede Lösung v gilt.

Lemma 2.11. *Sei v ein Jacobi-Feld. Dann gilt*

(i) *Wenn $v(x_0) = 0$, dann $v'(x_0) \neq 0$.*

(ii) *Die Nullstellen von v sind isoliert.*

(iii) *Gilt $v(x_0) = 0$ und ist $\tilde{v}$ eine weitere Lösung, dann: $v, \tilde{v}$ linear abhängig genau dann, wenn $\tilde{v}(x_0) = 0$.*

Für zwei Lösungen v_1, v_2 bezeichnet $W(x) := v_1 v_2' - v_1' v_2$ die Wronski-Determinante.

Lemma 2.12. (i) *$\{v_1, v_2\}$ ist Basis genau dann, wenn $W(x) \neq 0$.*

(ii) *$P(x)W(x) = P(x_0)W(x_0)$ für alle x.*

(iii) *Ist $\{v_1, v_2\}$ eine Basis, dann liegt zwischen zwei benachbarten Nullstellen von v_1 genau eine von v_2. Diese Aussage heißt* Sturmscher Oszillationssatz.

Beweis. (ii) $(PW)' = 0$.
(iii) Da $\{v_1, v_2\}$ eine Basis ist, gilt $W(x) \neq 0$, und wir können $W(x) > 0$ annehmen. Bezeichnen ξ, ξ^* zwei benachbarte Nullstellen von v_1, d. h. $v_1(x) \neq 0$ für $x \in (\xi, \xi^*)$. Wegen $v_1'(\xi) \neq 0$ und $v_1'(\xi^*) \neq 0$ haben diese Werte verschiedene Vorzeichen. Aus $W(x) > 0$ folgt $-v_1'(\xi)v_2(\xi) > 0$ und $-v_1'(\xi^*)v_2(\xi^*) > 0$. Das bedeutet $v_2(\xi) > 0 > v_2(\xi^*)$ oder $v_2(\xi) < 0 < v_2(\xi^*)$. Somit hat v_2 eine Nullstelle in (ξ, ξ_*). Gäbe es eine weitere Nullstelle von v_2 in (ξ, ξ_*), so liefert das gleiche Argument eine weitere Nullstelle von v_1 in (ξ, ξ_*); das ist nicht möglich. □

Definition. Wir sagen, der Punkt $x_0^* \in \mathbb{R} \setminus \{x_0\}$ ist *konjugiert* zu x_0 (entlang von u), wenn es ein Jacobi-Feld v entlang u gibt mit $v(x_0^*) = v(x_0) = 0$.

Wir kommen nun zur wichtigsten Aussage über konjugierte Punkte. Dieser Satz zeigt, dass die Abwesenheit konjugierter Punkte notwendig und hinreichend für die Existenz eines positiven Jacobi-Felds, mithin eines Minimierers ist (in einer recht allgemeinen Situation). Somit ist die Frage, ob eine Extremale Minimierer ist, auf das Auffinden konjugierter Punkte zurückgeführt, was ein „diskretes Problem“ ist.

Theorem 2.13. *Angenommen, u ist Extremale und erfüllt $P(x) \equiv L_{pp}(x, u(x), u'(x)) > 0$ für alle $x \in [a, b]$.*

(i) *Wenn in $[a, b]$ kein zu a konjugierter Punkt existiert, so ist u striktes lokales Minimum.*

(ii) *Wenn in (a, b) ein zu a konjugierter Punkt existiert, so ist u kein lokales Minimum.*

(iii) *Wenn b in $[a, b]$ der nächste zu a konjugierte Punkt ist, so ist keine Aussage über Art des Extremums möglich.*

Wir verwenden im Folgenden die Bezeichnung $\Delta(x, x_0)$ für die Lösung der Jacobi-Hilfsgleichung mit Anfangsdaten $v(x_0) = 0$ und $v'(x_0) = 1$.

Beweis. Zu (i). Idee: Wir konstruieren ein Jacobi-Feld v mit $v(x) > 0$ auf $[a, b]$ und wenden Satz 2.10 an. SKIZZE!

1. Schritt: Da $P(x) > 0$ auf ganz $[a, b]$ gilt, können wir annehmen, dass dies auf $(a - \delta, b + \delta)$ mit $\delta > 0$ gilt. Nach Voraussetzung gilt $\Delta(x, a) > 0$ auf $(a, b]$. Außerdem sind $\Delta(x, a)$ und $\Delta(x, b)$ linear unabhängig, da $\Delta(b, b) = 0$ (Lm. 2.11).

2. Schritt: Es gibt ein $x_0 \in (a - \delta, a)$, sodass $\Delta(x, b) \neq 0$ auf $[x_0, b)$ gilt. Andernfalls gäbe es in $[a, b)$ einen größten zu b konjugierten Punkt b^* und nach Sturm besitzt $\Delta(x, a)$ dann eine Nullstelle in (b^*, b), was nicht geht.
Wir betrachten nun $v(x) := \Delta(x, x_0)$. Da $v(x_0) = 0$ und $\Delta(x_0, b) \neq 0$, sind v und $\Delta(\cdot, b)$ uanbhängig. Dann gilt $v(x) > 0$ in $(x_0, b]$; andernfalls liefert Sturm, dass $\Delta(x, b)$ ein Nullstelle in (x_0, b) hat, was nicht möglich ist.
Somit ist $v(x)$ ein positives Jacobi-Feld auf $[a, b]$. Satz 2.10 liefert die Behauptung.

Zu (ii). Idee: Wir konstruieren eine Funktion φ, für die $\delta^2 I[u]\varphi < 0$ gilt, und wenden dann Satz 2.7 an. SKIZZE!

1. Schritt: Sei $a^* > a$ der nächste zu a konjugierte Punkt, o.E. bedeutet dies, das Jacobi-Feld $v_1(x) := \Delta(x, a)$ erfüllt $v_1(a) = v_1(a^*) = 0$ und $v_1(x) \neq 0$ in (a, a^*). Wir wählen $\beta \in (a^*, b)$ so, dass (a^*, β) keinen zu a konjugierten Punkt enthält, und betrachten zusätzlich das Jacobi-Feld $v_2(x) := -\Delta(x, \beta)$. Da $v_1(\beta) \neq 0$ und $v_2(\beta) = 0$, sind v_1, v_2 unabhängig. Somit besitzt v_2 (nach Sturm) genau eine Nullstelle $\alpha \in (a, a^*)$ und es gilt $v_2(a^*) \neq 0$ (da v_1, v_2 unabhängig). Außerdem gilt $v_2(x) \neq 0$ für $x \in (a^*, \beta)$; andernfalls hätte v_1 nach Sturm eine Nullstelle in (a^*, β) – im Widerspruch zur Wahl von β. Somit ist α die einzige Nullstelle von v_2 in $[a, \beta)$ und $v_2'(\beta) = -1$ impliziert $v_2(x) > 0$ auf (α, β) sowie $v_2(x) < 0$ auf $[a, \alpha)$.

Nun betrachten wir die Wronski-Determinante $W = v_1 v_2' - v_1' v_2$. Nach Lemma 2.12 gilt

$$P(x)W(x) \equiv C \qquad \text{mit } C = P(a)W(a) = P(a)(-v_2(a)) > 0.$$

Auch $v_1, v_2 - v_1$ sind unabhängig, sodass nach Sturm eine Nullstelle $\gamma \in (a, a^*)$ existiert, d. h. $v_1(\gamma) = v_2(\gamma)$.

2. Schritt: Wir definieren nun

$$\varphi(x) := \begin{cases} v_1(x) & \text{auf } [a, \gamma], \\ v_2(x) & \text{auf } [\gamma, \beta], \\ 0 & \text{auf } [\beta, b]. \end{cases}$$

φ ist stetig und stückweise C^2, löst (JHGl) stückweise und erfüllt $\varphi(a) = \varphi(b) = 0$. φ', φ'' sind nur bei $x \in \{\gamma, \beta\}$ unstetig. Wir berechnen jetzt $\mathcal{D} := \delta^2 I[u]\varphi$. Für φ gilt:

$$\begin{aligned} \mathcal{D} &= \int_a^b Q(x)\varphi(x)^2 + P(x)(\varphi'(x))^2 \,\mathrm{d}x \\ &= \int_a^\gamma Q v_1^2 + P(v_1')^2 \,\mathrm{d}x + \int_\gamma^\beta Q v_2^2 + P(v_2')^2 \,\mathrm{d}x \\ &= v_1 P v_1' \big|_a^\gamma + v_2 P v_2' \big|_\gamma^\beta , \end{aligned}$$

denn v_1, v_2 sind Jacobi-Felder. Aufgrund von $v_1(a) = 0, v_2(\beta) = 0, v_1(\gamma) = v_2(\gamma)$ folgt weiter

$$\begin{aligned} \mathcal{D} &= P v_1(\gamma) v_1'(\gamma) - P v_1(a) v_1'(a) + P v_2(\beta) v_2'(\beta) - P v_2(\gamma) v_2'(\gamma) \\ &= P(v_2(\gamma) v_1'(\gamma) - v_1(\gamma) v_2'(\gamma)) = -PW(\gamma) = -C < 0. \end{aligned}$$

Durch Glätten von φ finden wir auch ein $\tilde{\varphi} \in C_0^\infty$ mit $\mathcal{D}(\tilde{\varphi}) < 0$. Satz 2.7 liefert die Behauptung.

Zu (iii). Beispiele zeigen, dass in dieser Situation beides eintreten kann. Siehe Referenzen in [7, p. 284]. $\square$

Beispiel. Wir kommen auf das Beispiel $I[w] := \frac{1}{2}\int_0^l (w')^2 - kw^2$ mit $w(0) = w(l) = 0$ zurück. Hier ist $P \equiv 1 > 0$ global erfüllt. Die Jacobi-Gleichung lautet $v'' + kv = 0$. Somit erhalten wir

$$\Delta(x, 0) = \begin{cases} \frac{1}{\sqrt{-k}} \sinh(\sqrt{-k}x), & \text{wenn } k < 0, \\ x, & \text{wenn } k = 0, \\ \frac{1}{\sqrt{k}} \sin(\sqrt{k}x), & \text{wenn } k > 0. \end{cases}$$

Wenn $k \leq 0$, so gibt es also keinen zu 0 konjugierten Punkt (Minimierer). Wenn $k > 0$, so ist der erste konjugierte Punkt bei $x^* := \pi/\sqrt{k}$. Also gibt es für $u \equiv 0$ keinen konjugierten Punkt (Minimierer), sofern $l < x^*$. Für $l \geq x^*$ liegt in $[0, l]$ ein konj. Punkt (kein Minimierer für $l > x^*$). Für $l = x^*$ ist $u \equiv 0$ noch Minimierer, aber nicht mehr strikter Minimierer.

Bemerkung. (i) Selbst wenn die starke Legendre-Bedingung nur an einer Stelle verletzt ist, kann die Minimierer-Eigenschaft verloren gehen (siehe Aufgabe 4, Serie 1).

(ii) Jacobi-Felder kann man aus der allgemeinen Lösung der Euler-Lagrange-Gleichung erhalten: Wenn $u(x; c_1, c_2)$ allgemeine Lsg. der ELGl ist, dann erfüllen $\frac{\partial u}{\partial c_1}$ und $\frac{\partial u}{\partial c_2}$ die Jacobische Hilfsgleichung. (Siehe Übung).

(d) *Ausblick: Morse-Index

Um die Extremalen zu klassifizieren, ist eine Charakterisierung wie im Falle von Funktionen $f : \mathbb{R}^n \to \mathbb{R}$ wünschenswert. Hier ist bekannt, dass man ein Extremum x anhand seines *Morse-Index* $m(x)$ charakterisieren kann, d. h. der Anzahl negativer Eigenwerte von $H_f(x)$, sofern $\ker H_f(x) = 0$ gilt, d. h. x *nichtdegeneriert* ist. Dann ist x ein Minimum, wenn $m(x) = 0$, ein Maximum, wenn $m(x) = n$, und ein Sattelpunkt mit genau $m(x)$ negativen und $n - m(x)$ positiven Richtungen, wenn $m(x) \in \{1, \dots, n-1\}$.

Eine analoge Charakterisierung ist auch für I gesucht und unter geeigneten Voraussetzungen tatsächlich möglich. In diesem Fall ist der Morse-Index $m(u)$ die größtmögliche Dimension eines Unterraums, auf dem $\delta^2 I[u]$ negativ definit ist (das verallgemeinert den endlichdimensionalen Fall sinnvoll). Zunächst ist nicht einmal klar, ob $m(u) < \infty$ ist. Aber das lässt sich bewerkstelligen und der Höhepunkt dieser Theorie ist die explizite Angabe des Morse-Index:

$$m(u) = \#\text{konjugierte Punkte}.$$

Hierbei müssen die konjugierten Punkte mit Vielfachheit gezählt werden. Für einen zu a konjugierten Punkt a^* versteht man darunter die Anzahl linear unabhängiger Lösungen mit $v(a) = v(a^*) = 0$. In unserem Fall kann die Vielfachheit nur 0 oder 1 sein.

2.5 Variationsprobleme mit Nebenbedingungen

(a) Das klassische isoperimetrische Problem (Königin Dido)

Man finde diejenige Funktion $w \in \mathcal{A} := \{w \in C^1([a,b]) : w(a) = w(b) = 0\}$, die bei gegebener Länge

$$J[w] := \int_a^b \sqrt{1 + (w'(x))^2}\,\mathrm{d}x = l > 0$$

den mit der x-Achse eingeschlossenen Flächeninhalt maximiert:

$$I[w] := \int_a^b w(x)\,\mathrm{d}x.$$

Bemerkung. Wir haben hier schon die vereinfachte Annahme gemacht, dass wir die gesuchte Kurve als Graph schreiben können.

(b) Lagrange-Multiplikator

Probleme mit einer Nebenbedingung der Form

$$J[w] = \int_a^b M(x, w(x), w'(x))\,\mathrm{d}x$$

lassen sich auf ein Problem zurückführen, das sich in gewohnter Weise behandeln lässt. Hierbei tritt ein neuer Parameter auf, der Lagrange-Multiplikator. Das folgende Theorem zeigt, wie die Nebenbedingung eingeht.

Theorem 2.14. *Sei $u \in \mathcal{A}$ ein Minimierer von I unter der Nebenbedingung $J[u] = l$, der nicht gleichzeitig Extremum von J ist, d. h., es gibt ein ψ mit $\delta J[u]\psi \neq 0$. Dann existiert ein* Lagrange-Multiplikator $\lambda \in \mathbb{R}$, *sodass*

$$\delta I[u]\varphi + \lambda \delta J[u]\varphi = 0 \quad \forall \varphi \in C_0^\infty([a,b])$$

gilt.

Bemerkung. Aus der letzten Gleichung ergibt sich die Euler-Lagrange-Gleichung wie in 2.2.

Beweis. Die Beweisidee besteht – in Analogie zur Herleitung der Euler-Lagrange-Gleichung ohne Nebenbedingung – darin, das Problem auf skalare Funktionen zurückzuführen und dann den (klassischen) Satz über Existenz von Lagrange-Multiplikatoren im $\mathbb{R}^2$ anzuwenden.

Nach Voraussetzung existiert ein $\psi \in C_0^\infty([a,b])$ mit $\delta J[u]\psi = 1$. Für beliebiges $\varphi \in C_0^\infty([a,b])$ definieren wir

$$i(\varepsilon,\eta) := I[u + \varepsilon\varphi + \eta\psi] \quad \text{und} \quad j(\varepsilon,\eta) := J[u + \varepsilon\varphi + \eta\psi]$$

mit $(\varepsilon,\eta) \in [-\varepsilon_0, \varepsilon_0] \times [-\eta_0, \eta_0] =: Q$, wobei $\varepsilon_0, \eta_0 > 0$ hinreichend klein seien. Nun gilt

$$i(\varepsilon,\eta) \geq i(0,0) \quad \forall (\varepsilon,\eta) \in Q : j(\varepsilon,\eta) = l;$$

d. h. $i : Q \subset \mathbb{R}^2 \to \mathbb{R}$ hat bei $(0,0)$ ein Extremum unter der Nebenbedingung $j = l$. Da $\partial_\eta j(0,0) = 1 \neq 0$ ist $\nabla j(0,0) \neq 0$. Somit können wir aus dem Satz über Extrema unter Nebenbedingungen auf die Existenz eines Lagrange-Multiplikators $\lambda \in \mathbb{R}$ schließen, sodass $i(\varepsilon,\eta) + \lambda j(\varepsilon,\eta)$ ein Extremum bei $(0,0)$ hat. Ausführlich gilt:

$$\begin{aligned} i_\varepsilon(0,0) + \lambda j_\varepsilon(0,0) &= 0 \\ i_\eta(0,0) + \lambda j_\eta(0,0) &= 0 \end{aligned}$$

Mit $j_\eta(0,0) = 1$ folgt aus der zweiten Zeile $\lambda = -i_\eta(0,0) = \delta I[u]\psi$; man beachte, dass λ nicht von φ abhängt. Die erste Zeile bedeutet nun

$$\delta I[u]\varphi + \lambda \delta J[u]\varphi = 0,$$

also ausführlich

$$\int_a^b (L_z + \lambda M_z)\,\varphi + (L_p + \lambda M_p)\,\varphi'\,\mathrm{d}x = 0.$$

□

(c) Lösung des klassischen isoperimetrischen Problems

Das bewiesene Theorem wenden wir nun auf das Problem der Dido an. Die erweiterte Lagrange-Funktion lautet

$$\tilde{L}(x, z, p, \lambda) = L(x, z, p) + \lambda M(x, z, p) = z + \lambda\sqrt{1 + p^2}.$$

Die zugehörige ELGl ist

$$1 = \lambda \frac{d}{dx} \left[\frac{u'}{\sqrt{1 + (u')^2}} \right];$$

hieraus ergibt sich

$$\frac{1}{\lambda} = \frac{u''(x)}{(1 + u'(x)^2)^{3/2}}.$$

Die rechte Seite ist die Krümmung der Kurve $(x, u(x))$ bei x, und die linke Seite ist konstant und $\neq 0$. Somit ist die Kurve ein Kreisbogen. (Man kann auch anders argumentieren.)

2.6 *Natürliche Randbedingungen

Wir haben bisher stets Randbedingungen der Form

$$w(a) = A, \quad w(b) = B \tag{2.10}$$

vorgegeben. Alle zulässigen Funktionen $w \in \mathcal{A}$ müssen sie erfüllen und die zugelassenen „Störungen“ φ müssen so beschaffen sein, dass $w + \varepsilon\varphi \in \mathcal{A}$ für kleines ε gilt.

Nun betrachten wir folgendes Problem: *Finde die kürzeste Verbindung (in der Ebene) zwischen den Geraden $x = a$ und $x = b$.* Hier ist keine Randbedingung der Form (2.10) gegeben. Dementsprechend wählen wir

$$\mathcal{A} := \{w \in C^1([a, b])\}.$$

Wir können alle Störungen $\varphi \in C^\infty([a, b])$ zulassen (ohne Nullrandwerte!), denn $w+\varepsilon\varphi \in \mathcal{A}$ ist offenbar gewährleistet. Das hat interessante Konsequenzen für die notwendige Bedingung $\delta I[u]\varphi = 0$. In Analogie zu Satz 2.2 erhalten wir den folgenden Satz.

Satz 2.15. *Sei $L \in C^1([a, b] \times \mathbb{R} \times \mathbb{R})$ und $u \in \mathcal{A}$ ein Minimierer des Funktionals*

$$I[w] := \int_a^b L(x, w(x), w'(x))\, \mathrm{d}x.$$

Dann gilt $L_p(x, u(x), u'(x)) \in C^1([a, b])$ und u erfüllt die Euler-Lagrange-Gleichung (2.3) *sowie die* natürlichen Randbedingungen

$$L_p(a, u(a), u'(a)) = 0 \quad \textit{und} \quad L_p(b, u(b), u'(b)) = 0. \tag{2.11}$$

Beweis. Wie im Beweis von Satz 2.2 erhalten wir aus $\delta I[u]\varphi = 0$ mit $\varphi \in C^1([a,b])$ die Formel

$$\int_a^b L_z(x,u,u')\varphi + L_p(x,u,u')\varphi' \, \mathrm{d}x = 0.$$

Von dort wissen wir, dass $L_p(x,u,u') \in C^1$ ist, also können wir partiell integrieren:

$$0 = \int_a^b \left[L_z(x,u,u') - \frac{d}{dx} L_p(x,u,u') \right] \varphi \, \mathrm{d}x + L_p(b,u(b),u'(b))\varphi(b) - L_p(a,u(a),u'(a))\varphi(a),$$

wobei jetzt Randterme zu berücksichtigen sind. Aufgrund der ELGl verschwindet das Integral und somit bleibt nur

$$L_p(b,u(b),u'(b))\varphi(b) - L_p(a,u(a),u'(a))\varphi(a).$$

Wählen wir ein φ mit $\varphi(a) \neq 0, \varphi(b) = 0$ bzw. mit $\varphi(a) = 0, \varphi(b) \neq 0$ so erhalten wir die natürliche Randbedingung am linken bzw. rechten Rand. $\square$

Bemerkung. Eine entsprechende Aussage gilt, wenn nur ein Randwert vorgegeben ist. Am anderen Rand gilt dann die natürliche Randbedingung.

Beispiel. Wir betrachten das oben erwähnte Beispiel: In diesem Fall ist

$$I[w] = \int_a^b \sqrt{1 + (w'(x))^2} \, \mathrm{d}x,$$

also $L = \sqrt{1+p^2}$. Die natürlichen Randbedingungen besagen demnach $u'(a) = 0$ und $u'(b) = 0$. Das ist anschaulich klar, weil man im Falle $u'(a) \neq 0$ bzw. $u'(b) \neq 0$ leicht eine kürzere Verbindung findet.

2.7 *Analytische Mechanik

Bezeichnet $r(t) = (x(t), y(t), z(t)), t \in [t_0, t_1]$, die Bahnkurve eines Massepunktes m, so ist

$$T[r] = \frac{m}{2} |\dot{r}(t)|^2$$

die kinetische Energie zur Zeit t. Die auf m wirkende Kraft f besitze eine Potential $V = V(t,r)$, s.d. $f = -\nabla V$. Dann heißt $L := T - V$ *Lagrange-Funktion* und das *Hamiltonsche Prinzip* besagt: Die Bahnkurve $r(t)$ ist so beschaffen, dass das *Wirkungsintegral*

$$I[r] := \int_{t_0}^{t_1} L(t, r(t), \dot{r}(t)) \, \mathrm{d}t$$

stationär wird, also bei r ein Extremum oder Sattelpunkt besitzt.

Das gilt entsprechend auch für mehrere Massepunkte. Hierbei müssen nicht kartesische Koordinaten benutzt werden, sondern es können beliebige dem Problem angepasste, sogenannte „verallgemeinerte“ Koordinaten verwendet werden (deren Anzahl mit der

Zahl der Freiheitsgrade übereinstimmt); diese werden klassischerweise mit $(q_1, \dots, q_m)$ bezeichnet. Das mathematische Pendel als Beispiel:

$$T = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) = \frac{1}{2}ml^2\dot{\varphi}^2, V = -mgl\cos(\varphi) \Longrightarrow L = \frac{ml^2}{2}\dot{\varphi}^2 + mgl\cos(\varphi).$$

$$I[\varphi] := \int_{t_0}^{t_1} \frac{ml^2}{2}\dot{\varphi}^2 + mgl\cos(\varphi)\,\mathrm{d}t$$

3 Die direkte Methode in L^p, Teil I: Skalare Probleme

3.1 Grundbegriffe: Erste und zweite Variation, Euler-Lagrange-Gleichung

(a) Bezeichnungen

Wir betrachten eine beschränkte, offene Menge $U \subset \mathbb{R}^n$ mit glattem Rand ∂U und eine gegebene glatte Funktion, die *Lagrange-Funktion*,

$$L : \bar{U} \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}.$$

Wir schreiben $L = L(x, z, p) = L(x_1, \ldots, x_n, z, p_1, \ldots, p_n)$ und wir setzen $D_p L = (L_{p_1}, \ldots, L_{p_n})$, $D_z L = L_z$, $D_x L = (L_{x_1}, \ldots, L_{x_n})$. Im Folgenden untersuchen wir Funktionale von der Form

$$I[w] := \int_U L(x, w(x), Dw(x))\, \mathrm{d}x \tag{3.1}$$

zunächst für glatte Funktionen $w : \bar{U} \to \mathbb{R}$, die der Randbedingung

$$w = g \quad \text{auf } \partial U \tag{3.2}$$

genügen.

Bereits in der Einleitung haben wir die *erste Variation*, d. h. $i'(0)$ von

$$i(\varepsilon) := I[u + \varepsilon\varphi] = \int_U L(x, (u + \varepsilon\varphi)(x), D(u + \varepsilon\varphi)(x))\, \mathrm{d}x \quad \text{mit } \varphi \in C_0^\infty(U),$$

berechnet:

$$i'(0) = \int_U \sum_{i=1}^n L_{p_i}(x, u, Du)\varphi_{x_i} + L_z(x, u, Du)\varphi\, \mathrm{d}x.$$

Aus $i'(0) = 0$ ergibt sich (mit dem Fundamentallemma) die *Euler-Lagrange-Gleichung*

$$-\sum_{i=1}^n \left(L_{p_i}(x, u, Du)\right)_{x_i} + L_z(x, u, Du) = 0.$$

Beispiel (*)**.** Für

$$L(p, z, x) = \frac{1}{2}\, |p|^2 \tag{3.3}$$

ist $L_{p_i} = p_i$ und $L_z = 0$. Also gehört zum Funktional

$$I[w] := \frac{1}{2} \int_U |Dw|^2\, \mathrm{d}x$$

die ELGl

$$\Delta u = 0 \quad \text{in } U.$$

Man spricht vom *Dirichlet-Prinzip*.

(b) Zweite Variation

In Analogie zum 1D-Fall berechnen wir die *zweite Variation*

$$i''(0) = \left.\frac{\mathrm{d}^2}{\mathrm{d}\varepsilon^2}\right|_{\varepsilon=0} I[u + \varepsilon\varphi].$$

Ist u ein genügend glatter Minimierer und $i \in C^2$, dann folgt $i''(0) \geq 0$. Also gilt

(3.4)
$$0 \leq i''(0) = \int_U \sum_{i,j=1}^n L_{p_i p_j}(x,u,Du)\varphi_{x_i}\varphi_{x_j} + 2\sum_{i=1}^n L_{p_i z}(x,u,Du)\varphi_{x_i}\varphi + L_{zz}(x,u,Du)\varphi^2 \,\mathrm{d}x$$

für alle Testfunktionen $\varphi \in C_0^\infty(U)$. Daraus können wir wie folgt eine punktweise Aussage über L: Durch Approximation gilt diese Aussage auch für alle Lipschitz-Funktionen, die auf dem Rand verschwinden. (Zu diesem Zweck beachte, dass $W^{1,\infty}(U) \subset W^{1,q}(U)$ die Menge der Lipschitz-Funktionen ist und man eine solche durch eine Folge von C^∞-Funktionen $u^\varepsilon \to u$ bez. $\|\cdot\|_{W^{1,q}}$ approximieren kann.) Für festes $\xi \in \mathbb{R}^n$ definiere

$$\varphi(x) = \varepsilon\rho\left(\frac{x\cdot\xi}{\varepsilon}\right)\zeta(x),$$

wobei $\zeta \in C_0^\infty(U)$ und $\rho : \mathbb{R} \to \mathbb{R}$ bezeichne die 1-periodische Zickzack-Funktion

$$\rho(x) = \begin{cases} x, & \text{wenn } 0 \leq x \leq \frac{1}{2} \\ 1 - x, & \text{wenn } \frac{1}{2} \leq x \leq 1 \end{cases}, \quad \rho(x+1) = \rho(x).$$

Also gilt $|\rho'| = 1$ f.ü. und $\varphi_{x_i} = \rho'\left(\frac{x\cdot\xi}{\varepsilon}\right)\xi_i\zeta + O(\varepsilon)$. Einsetzen in (3.4) liefert

$$0 \leq \int_U \sum_{i,j=1}^n L_{p_i p_j}(x,u,Du)\xi_i\xi_j\zeta^2 \,\mathrm{d}x + O(\varepsilon)$$

und daraus folgt bei Grenzübergang $\varepsilon \to 0$ unter Beachtung der Tatsache, dass dies für alle ζ gilt:

$$\sum_{i,j=1}^n L_{p_i p_j}(x,u,Du)\xi_i\xi_j \geq 0.$$

Das ist die Begründung dafür, dass später von L Konvexität bezüglich p gefordert wird. Diese Relation heißt auch *Legendre-(Hadamard)-Bedingung.*

3.2 Forderungen an I: Koerzivität, Unterhalbstetigkeit

In diesem Abschnitt führen wir die Eigenschaften von L ein, die später beim Existenzbeweis vorausgesetzt werden.

(a) Koerzivität

Schon die Beispiele $x \mapsto \mathrm{e}^x$ und $x \mapsto (1+x^2)^{-1}$ zeigen, dass das Infimum im Allgemeinen nicht angenommen wird, wenn die Funktion nicht am Rand wächst (SKIZZE). Die allgemeine Bedingung für $I : X \to \mathbb{R}$ lautet: I heißt *koerziv*, wenn aus $\|x_n\| \to \infty$ folgt $I[x_n] \to \infty$. Wir fordern im Folgenden eine schärfere Bedingung. Dabei setzen wir voraus:

$$1 < q < \infty.$$

Dass die Aussage für $q = 1$ falsch ist, sieht man in Aufgabe 12, Serie 3.

Definition. I erfüllt die *Koerzivitätsbedingung*, wenn es Konstanten $\alpha > 0, \beta \geq 0$ gibt, sodass

$$L(x,z,p) \geq \alpha\, |p|^q - \beta \tag{KB}$$

für alle $x \in U, z \in \mathbb{R}, p \in \mathbb{R}^n$ gilt.

Dann ist I tatsächlich koerziv in obigem Sinne, denn wir erhalten direkt

$$I[w] \geq \alpha\, \|Dw\|_{L^q}^q - \gamma$$

mit $\gamma := \beta\, |U|$ und die Poincare-Ungleichung liefert $\|Dw\|_{L^q} \geq C\, \|w\|_{L^q}$, sodass also $I[w] \to \infty$ gilt, wenn $\|w\|_{W_0^{1,q}(U)} \to \infty$. Diese Überlegung wird in Abschnitt 3.3. präzisiert.

Im Folgenden betrachten wir

$$\mathcal{A} := \{w \in W^{1,q}(U) : w|_{\partial U} = g\}$$

als Menge zulässiger Funktionen, wobei die Randwerte im Sinne der Spur angenommen werden. Dann ist $I[w] \in (-\infty, \infty]$ für jedes $w \in \mathcal{A}$ definiert (aber möglicherweise $+\infty$).

(b) Schwache Topologie

Eine stetige Funktion $\mathbb{R} \to \mathbb{R}$, die einer Koerzivitätsbedingung genügt, nimmt ihr Infimum an. Für das Funktional I braucht das nicht zu gelten.

Immerhin ist das Funktional I nach unten beschränkt. Also existiert

$$m := \inf_{w \in \mathcal{A}} I[w].$$

Wähle eine *minimierende Folge* $w_k \in \mathcal{A}$ mit

$$\lim_{k \to \infty} I[w_k] = m.$$

Um zu zeigen, dass eine Teilfolge von (w_k) tatsächlich konvergiert, benötigt man Kompaktheitseigenschaften (der Folge bzw. des Raumes). Das ist ein Problem, da $W^{1,q}$ unendlichdimensional ist.

Aus (KB) folgt immerhin, dass (w_k) beschränkt ist. Um Kompaktheit zu bekommen, müssen wir eine andere „schwächere" Topologie wählen. Da wir $1 < q < \infty$ gewählt haben, ist $L^q(U)$ reflexiv, deshalb existieren eine Teilfolge w_{k_j} und eine Funktion $w \in W^{1,q}$, sodass

$$w_{k_j} \rightharpoonup w \quad \text{schwach in } L^q(U) \quad \text{und} \quad Dw_{k_j} \rightharpoonup Dw \quad \text{schwach in } L^q(U;\mathbb{R}^n).$$

Dazu sagen wir kurz: „$w_{k_j} \rightharpoonup w$ schwach in $W^{1,q}(U)$".

Bemerkung. Es handelt sich dabei tatsächlich um die schwache Konvergenz von Funktionenfolgen in $W^{1,q}(U)$. Um das zu sehen, muss man sich daran erinnern, dass sich jedes Funktional $l \in (W^{1,q}(U))'$ als

$$l(u) = \int_U \sum_{i=1}^n g_i(x) u_{x_i}(x) + g(x)u(x)\,\mathrm{d}x \quad \text{für alle } u \in W^{1,q}(U)$$

mit $g, g_i \in L^{q'}(U)$ schreiben lässt, d. h., dass $(W^{1,q}(U))' \cong W^{-1,q'}(U)$ gilt. (Siehe Literatur, z. B. Mazya)

Es wird sich als wahr herausstellen, dass $w = g$ am Rand gilt, sodass tatsächlich $w \in \mathcal{A}$ ist. Durch die Betrachtung der schwachen Topologie haben wir demnach die nötige Kompaktheit gewonnen.

(c) Halbstetigkeit von unten

Nun ergibt sich ein weiteres Problem. Das Funktional I ist im Allgemeinen nicht stetig bezüglich der schwachen Konvergenz (Beispiel siehe Übung); das Problem ist, $Dw_{k_j} \rightharpoonup Dw$ impliziert nicht $Dw_{k_j} \to Dw$ f.ü. (Beispiel siehe Übung).

Andererseits ist Stetigkeit nicht nötig. Es genügt zu wissen, dass

$$I[w] \le \liminf_{j\to\infty} I[w_{k_j}].$$

Denn daraus folgt $I[w] \le m$. Zusammen mit $m \le I[w]$ folgt $m = I[w]$, d. h., w ist ein Minimierer.

Definition. Die Funktion I heißt *schwach (folgen-)halbstetig von unten* auf $W^{1,q}$, wenn aus

$$u_k \rightharpoonup u \quad \text{schwach in } W^{1,q}$$

folgt:

$$I[u] \le \liminf_{k\to\infty} I[u_k].$$

(d) Konvexität impliziert sUHS

Wir zeigen jetzt, dass I schwach unterhalbstetig ist, wenn L konvex bezüglich p ist. Diese Bedingung hatten wir ohnehin schon als notwendige Bedingung für die Existenz von genügend glatten Minimierern identifiziert.

Theorem 3.1 (Schwache Unterhalbstetigkeit). *Wir nehmen an, L ist glatt, nach unten beschränkt und die Abbildung*

$$p \mapsto L(x,z,p) \quad \textit{ist konvex für jedes } z \in \mathbb{R}, x \in U.$$

Dann ist $I : W^{1,q}(U) \to \mathbb{R}$ schwach unterhalbstetig.

- **Hilfsmittel aus Maßtheorie und Funktionalanalysis:**

(H1) Jede schwach konvergente Folge ist beschränkt.

(H2) Jede stark konvergente Folge in L^q besitzt eine Teilfolge, die f.ü. punktweise konvergiert.

(H3) Die Einbettung $W^{1,q} \subset L^q$ ist kompakt, d. h. jede beschränkte Folge in $W^{1,q}$ besitzt eine konvergente Teilfolge in L^q.

(H4) Theorem von Egorov: Sind $f, f_k : X \to \mathbb{R}$ messbare Funktionen mit $f_k \to f$ f.ü. auf der messbaren Menge A, $|A| < \infty$. Dann gibt es zu jedem $\varepsilon > 0$ eine messbare Menge $E \subset A$ mit: (i) $|A \setminus E| \le \varepsilon$, (ii) $f_k \to f$ gleichmäßig auf E.

Beweis. **1. Schritt:** Wir wählen eine Folge $(u_k)_{k\in\mathbb{N}}$ mit $u_k \rightharpoonup u$ in $W^{1,q}(U)$ und setzen $l := \liminf_{k\to\infty} I[u_k]$. Wir müssen zeigen, dass $I[u] \le l$ gilt. Evtl. nach Übergang zu einer Teilfolge können wir $l := \lim_{k\to\infty} I[u_k]$ annehmen.

Nach (H1) gilt $\sup_k \|u_k\|_{W^{1,q}} < \infty$. Aufgrund von (H3) gilt $u_k \to u$ in L^q. Wegen (H2) gilt für eine Teilfolge $u_k \to u$ f.ü. in U.

2. Schritt: Wähle $\varepsilon > 0$. Nach (H4) gilt dann $u_k \to u$ gleichmäßig auf E_ε, einer messbaren Menge mit $|U \setminus E_\varepsilon| \le \varepsilon$. Nun betrachten wir

$$F_\varepsilon := \{x \in U : |u(x)| + |Du(x)| \le \frac{1}{\varepsilon}\}.$$

Dann gilt $|U \setminus F_\varepsilon| \to 0$ für $\varepsilon \to 0$. Für $G_\varepsilon := E_\varepsilon \cap F_\varepsilon$ gilt nun auch $|U \setminus G_\varepsilon| \to 0$.

3. Schritt: Da L nach unten beschränkt ist, können wir $L \ge 0$ annehmen. Also gilt

$$\begin{aligned} I[u_k] &= \int_U L(x,u_k,Du_k)\,\mathrm{d}x \ge \int_{G_\varepsilon} L(x,u_k,Du_k)\,\mathrm{d}x \\ &\ge \int_{G_\varepsilon} L(x,u_k,Du) + D_pL(x,u_k,Du)\cdot(Du_k - Du)\,\mathrm{d}x \end{aligned}$$

aufgrund der Konvexität. Nun gilt

$$\lim_{k\to\infty} \int_{G_\varepsilon} L(x,u_k,Du)\,\mathrm{d}x = \int_{G_\varepsilon} L(x,u,Du)\,\mathrm{d}x.$$

Da $D_pL(x,u_k,Du) \to D_pL(x,u,Du)$ gleichmäßig auf G_ε, gilt

$$\int_{G_\varepsilon} |D_pL(x,u_k,Du) - D_pL(x,u,Du)| \cdot |Du_k - Du|\ \mathrm{d}x \to 0,$$

denn der erste Faktor lässt sich durch eine beliebig kleine Konstante abschätzen und das übrig bleibende Integral ist beschränkt. Da $Du_k \rightharpoonup Du$ in L^q, gilt

$$\left| \int_{G_\varepsilon} D_p L(x, u, Du) \cdot (Du_k - Du) \, \mathrm{d}x \right| \to 0,$$

denn $D_p L(x, u, Du) \in L^\infty(U) \subset L^{q'}$. Insgesamt sehen wir damit

$$\lim_{k\to\infty} \int_{G_\varepsilon} D_p L(x, u_k, Du) \cdot (Du_k - Du) \, \mathrm{d}x = 0.$$

Daraus folgt

$$l = \lim_{k\to\infty} I[u_k] \geq \int_{G_\varepsilon} L(x, u, Du) \, \mathrm{d}x$$

für jedes $\varepsilon > 0$. Wegen $L \geq 0$ liefert der Satz über monotone Konvergenz

$$l \geq \int_U L(x, u, Du) \, \mathrm{d}x = I[u],$$

wie behauptet. □

Bemerkung. Es ist sehr wichtig zu verstehen, wie die schwache Konvergenz $Du_k \rightharpoonup Du$ im Beweis eingeht: Entscheidend ist die Konvexitätsungleichung, auf deren rechter Seite Du_k nur linear auftaucht. Nach Definition ist schwache Konvergenz mit linearen Ausdrücken kompatibel, sodass man zum Grenzwert übergehen kann. Da die Konvergenz $u_k \to u$ in L^q viel stärker ist, benötigen wir keine Konvexität von L bezüglich z.

3.3 Existenz und Eindeutigkeit von Minimierern

Theorem 3.2. *Angenommen, L erfüllt die Koerzivitätsbedingung und ist konvex bezüglich p. Die Menge $\mathcal{A}$ der zulässigen Funktionen sei nicht leer. Dann existiert eine Funktion $u \in \mathcal{A}$ mit*

$$I[u] = \min_{w\in\mathcal{A}} I[w].$$

Weitere Hilfsmittel: (H5) Theorem von Mazur: Ist X ein Banach-Raum und $Y \subset X$ abgeschlossen und konvex, dann ist Y schwach abgeschlossen.
(H6) Poincaré Ungleichung: Für $w \in W_0^{1,q}(U)$ gilt $\|w\|_{L^q(U)} \leq C \|Dw\|_{L^q(U;\mathbb{R}^n)}$.

Beweis. **1. Schritt:** Setze $m := \inf_{w\in\mathcal{A}} I[w]$. Wenn $m = +\infty$, ist nichts zu tun. Andernfalls wählen wir eine minimierende Folge $(u_k)_{k\in\mathbb{N}}$ mit $I[u_k] \to m$.

2. Schritt: In der Koerzivitätsbedingung (KB) können wir $\beta = 0$ annehmen. Andernfalls würden wir zu $\tilde{L} := L + \beta$ übergehen. Also $L \geq \alpha |p|^q$ und somit

$$I[w] \geq \alpha \int_U |Dw|^q \, \mathrm{d}x.$$

Da $m < \infty$, schließen wir

$$\sup_k \|Du_k\|_{L^q} < \infty.$$

3. Schritt: Nun wählen wir ein $w \in \mathcal{A}$. Da u_k und w auf dem Rand gleich g sind, ist $u_k - w \in W_0^{1,q}(U)$. Die Poincaré-Ungleichung impliziert

$$\|u_k\|_{L^q} \le \|u_k - w\|_{L^q} + \|w\|_{L^q} \le C\,\|Du_k - Dw\|_{L^q} + C \le C.$$

Somit ist auch $\sup_k \|u_k\|_{L^q} < \infty$, also ist (u_k) beschränkt in $W^{1,q}(U)$.

4. Schritt: Folglich existiert eine Teilfolge

$$u_{k_j} \rightharpoonup u \quad \text{in } W^{1,q}(U).$$

Bleibt zu zeigen, dass $u \in \mathcal{A}$ ist. Um das einzusehen, beachten wir, dass für $w \in \mathcal{A}$ $u_k - w \in W_0^{1,q}(U)$ gilt. $W_0^{1,q}(U)$ ist ein linearer abgeschlossener Unterraum von $W^{1,q}(U)$, also schwach abgeschlossen nach (H5). Also gilt $u - w \in W_0^{1,q}(U)$, d. h., u hat Spur g auf ∂U. Nach Theorem 3.1 gilt dann $I[u] \le \liminf_{j\to\infty} I[u_{k_j}] = m$. Da aber $u \in \mathcal{A}$ ist, folgt

$$I[u] = m = \min_{w\in\mathcal{A}} I[w].$$

□

Im Allgemeinen gibt es keinen eindeutigen Minimierer. Wir geben weitere Bedingungen für L an, die Eindeutigkeit garantieren. Wir nehmen an:

$$L = L(x,p) \quad \text{hängt nicht von } z \text{ ab} \tag{3.5}$$

und

$$\exists \theta > 0 : \sum_{i,j=1}^{n} L_{p_i p_j}(x,p)\xi_i\xi_j \ge \theta\,|\xi|^2 \quad \text{für } p,\xi \in \mathbb{R}^n;\ x \in U \tag{3.6}$$

Die letzte Bedingung besagt, $p \mapsto L(x,p)$ ist gleichmäßig konvex für jedes x.

Theorem 3.3. *Angenommen, die Bedingungen* (3.5) *und* (3.6) *gelten. Dann besitzt* $I[\cdot]$ *einen eindeutigen Minimierer* $u \in \mathcal{A}$.

Beweis. **1. Schritt** Angenommen, es gäbe zwei Minimierer $u, \tilde{u} \in \mathcal{A}$ von I. Dann ist $v := \frac{1}{2}(u + \tilde{u}) \in \mathcal{A}$. Wir behaupten, dass

$$I[v] \le \frac{1}{2}(I[u] + I[\tilde{u}]) \tag{3.7}$$

gilt, wobei strikte Ungleichheit gilt, wenn nicht $u \equiv \tilde{u}$ f.ü. ist.

2. Schritt Um das zu zeigen, verwenden wir folgende Ungleichung, die aus der gleichmäßigen Konvexität folgt:

$$L(x,p) \ge L(x,q) + D_pL(x,q)\cdot(q-p) + \frac{\theta}{2}\,|p-q|^2, \quad x \in U, p, q \in \mathbb{R}^n. \tag{3.8}$$

Wir setzen $q := \frac{1}{2}(Du + D\tilde{u}), p := Du$ und integrieren über U; dann:

$$I[u] \ge I[v] + \int_U D_pL\left(\frac{Du + D\tilde{u}}{2}, x\right)\cdot\frac{Du - D\tilde{u}}{2}\,\mathrm{d}x + \frac{\theta}{8}\int_U |Du - D\tilde{u}|^2\,\mathrm{d}x.$$

Analog erhalten wir aus (3.8) mit $q := \frac{1}{2}(Du + D\tilde{u}), p := D\tilde{u}$:

$$I[\tilde{u}] \geq I[v] + \int_U D_p L\left(\frac{Du + D\tilde{u}}{2}, x\right) \cdot \frac{D\tilde{u} - Du}{2} \, \mathrm{d}x + \frac{\theta}{8} \int_U |Du - D\tilde{u}|^2 \, \mathrm{d}x.$$

Beide Ungleichungen addiert und halbiert ergeben

$$\frac{I[u] + I[\tilde{u}]}{2} \geq I[v] + \frac{\theta}{8} \int_U |Du - D\tilde{u}|^2 \, \mathrm{d}x.$$

Dies impliziert (3.7).

3. Schritt Gilt in der Ungleichung (3.7) Gleichheit, so folgt aus der letzten Ungleichung $Du = D\tilde{u}$ f.ü. Der Randwert $u = \tilde{u} \equiv g$ auf ∂U impliziert $u \equiv \tilde{u}$ f.ü.

□

3.4 Schwache Lösungen der Euler-Lagrange-Gleichung

Wir zeigen nun, dass jeder Minimierer $u \in \mathcal{A}$ von I eine schwache Lösung der ELGl ist. Zu diesem Zweck setzen wir zusätzlich voraus: Es gibt ein $C > 0$, sodass

$$|L(x, z, p)| \leq C\left(|p|^q + |z|^q + 1\right) \tag{3.9a}$$

$$|D_p L(x, z, p)| \leq C\left(|p|^{q-1} + |z|^{q-1} + 1\right) \tag{3.9b}$$

$$|D_z L(x, z, p)| \leq C\left(|p|^{q-1} + |z|^{q-1} + 1\right) \tag{3.9c}$$

für alle $x \in U$, $z \in \mathbb{R}$ und $p \in \mathbb{R}^n$ gilt.

(a) Motivation

Wir hatten uns bereits in der Einleitung davon überzeugt, dass genügend glatte Minimierer die Euler-Lagrange-Gleichung

$$\left\{ -\sum_{i=1}^{n} (L_{p_i}(x, u, Du))_{x_i} + L_z(x, u, Du) = 0 \quad \text{in } U, u \quad = g \quad \text{auf } \partial U \right. \tag{3.10}$$

erfüllen. Gehen wir umgekehrt von der ELGl aus, so finden wir durch Multiplikation mit einer Testfunktion $v \in C_0^\infty(U)$ und Integration die Gleichung

$$\int_U \sum_{i=1}^{n} L_{p_i}(x, u, Du) v_{x_i} + L_z(x, u, Du) v \, \mathrm{d}x = 0. \tag{3.11}$$

Nehmen wir $u \in W^{1,q}(U)$ an, so folgt mit (3.9)

$$|D_p L(x, u, Du)| \leq C\left(|Du|^{q-1} + |u|^{q-1} + 1\right) \in L^{q'}(U)$$

mit $q' = q/(q-1)$, sodass $1/q' + 1/q = 1$. Entsprechend ist

$$|D_z L(x, u, Du)| \leq C\left(|Du|^{q-1} + |u|^{q-1} + 1\right) \in L^{q'}(U).$$

Folglich gilt (3.11) durch Approximation für alle $v \in W_0^{1,q}(U)$. Wie üblich bezeichnen wir dies als schwache Formulierung der betrachteten Gleichung.

Definition (Schwache Lösung der ELGl)**.** $u \in \mathcal{A}$ heißt schwache Lösung von (3.10), wenn

$$\int_U \sum_{i=1}^n L_{p_i}(x,u,Du)v_{x_i} + L_z(x,u,Du)v\,\mathrm{d}x = 0$$

für alle $v \in W_0^{1,q}(U)$ gilt.

Theorem 3.4. *Angenommen, L erfüllt die Wachstumsbedingungen* (3.9) *und für $u \in \mathcal{A}$ gilt*

$$I[u] = \min_{w\in\mathcal{A}} I[w].$$

Dann ist u eine schwache Lösung von (3.10)*.*

Beweis. Idee: Wir verfahren wie in der Einleitung, rechtfertigen aber die einzelnen Schritte – insbesondere das Ableiten unterm Integral – genau.

Für $v \in W_0^{1,q}(U)$ setzen wir

$$i(\tau) := I[u+\tau v] \qquad \text{für } t \in \mathbb{R}.$$

Wegen (3.9a) ist $i(\tau) < \infty$ für alle τ. Für $\tau \neq 0$ betrachten wir den Differenzenquotienten

$$\frac{i(\tau)-i(0)}{\tau} = \int_U \underbrace{\frac{1}{\tau}\left(L(x,u+\tau v, Du+\tau Dv) - L(x,u,Du)\right)}_{=:L^\tau(x)} \mathrm{d}x.$$

Offenbar gilt für $\tau \to 0$

$$L^\tau(x) \to \sum_{i=1}^n L_{p_i}(x,u,Du)v_{x_i} + L_z(x,u,Du)v \quad \text{f.ü.}$$

Außerdem haben wir

$$\begin{aligned} L^\tau(x) &= \frac{1}{\tau}\int_0^\tau \frac{d}{ds} L(x,u+sv,Du+sDv)\,\mathrm{d}s \\ &= \frac{1}{\tau}\int_0^\tau \sum_{i=1}^n L_{p_i}(x,u+sv,Du+sDv)v_{x_i} + L_z(x,u+sv,Du+sDv)v\,\mathrm{d}s. \end{aligned}$$

Um den Satz von der majorisierten Konvergenz (Lebesgue) anzuwenden, benötigen wir eine Abschätzung von $|L^\tau(x)|$ gegen eine integrierbare Majorante. Wegen (3.9) gilt:

$$\begin{aligned} |L^\tau(x)| &\leq \frac{1}{\tau}\int_0^\tau |D_pL(x,u+sv,Du+sDv)|\,|Dv(x)| + |D_zL(\cdot)|\,|v|\,\mathrm{d}s \\ &\leq \frac{1}{\tau}\int_0^\tau C\left(|Du+sDv|^{q-1} + |u+sv|^{q-1} + 1\right)|Dv| \end{aligned}$$

$$+\,C\left(|Du+sDv|^{q-1}+|u+sv|^{q-1}+1\right)|v|\;\mathrm{d}s.$$

Im Integranden verwenden wir nun Abschätzungen von folgender Art:

$$|Du+sDv|^{q-1}\le(|Du|+s\,|Dv|)^{q-1}\le(|Du|+|Dv|)^{q-1}\le 2^{q-1}\left(|Du|^{q-1}+|Dv|^{q-1}\right),$$

wobei wir $(a+b)^s\le 2^s(a^s+b^s)$ ausgenutzt haben. Mittels der Youngschen Ungleichung $ab\le\frac{a^r}{r}+\frac{b^{r'}}{r'}$ mit $1/r+1/r'=1$ finden wir

$$|L^\tau(x)|\le C\left(|Du|^q+|u|^q+|Dv|^q+|v|^q+1\right).$$

Da $u,v\in W^{1,q}(U)$, folgt daraus, dass $|L^\tau(x)|\in L^1(U)$. Nun können wir den Satz von der majorisierten Konvergenz anwenden, um

$$i'(0)=\int_U\sum_{i=1}^n L_{p_i}(x,u,Du)v_{x_i}+L_z(x,u,Du)v\,\mathrm{d}x$$

zu erhalten. Dieser Ausdruck verschwindet, da $i(\tau)$ ein Minimum bei $\tau=0$ hat. □

Bemerkung. Im Allgemeinen erhält man keine Eindeutigkeit des Minimierers. Man bekommt sie jedoch (wie in 1D), wenn L konvex bezüglich (z,p) ist.

(b) Regularität

Wir machen einige Anmerkungen zu dem schwierigen Gebiet der Regularität von Minimierern. Hierfür nehmen wir zur Vereinfachung an, das Funktional habe die Form

$$I[w]:=\int_U L(Dw)-wf\,\mathrm{d}x$$

mit $f\in L^2(U)$. Wir nehmen weiter an $q=2$ und $D_pL(p)\le C(|p|+1)$. Ein Minimierer von I ist dann schwache Lösung von

$$-\sum_{i=1}^n\left(L_{p_i}(Du)\right)_{x_i}=f. \tag{3.12}$$

Das erste Ziel ist zu zeigen, dass aus $u\in H^1(U)$ sogar $u\in H^2_{\mathrm{loc}}(U)$ folgt. Wir nehmen an $\left|D^2L(p)\right|\le C$ und L ist gleichmäßig konvex, d. h.

$$\sum_{i,j=1}^n L_{p_ip_j}(p)\xi_i\xi_j\ge\theta\,|\xi|^2\,;$$

das ist eine nichtlineare Version der gleichmäßigen Elliptizität für lineare PDE. Deshalb liegt es nahe, die entsprechende Beweisidee zu übertragen.

Theorem 3.5 (Zweite Ableitung für Minimierer). *(i) Ist $u\in H^1(U)$ eine schwache Lösung von* (3.12), *dann gilt $u\in H^2_{loc}(U)$.*

(ii) Ist außerdem $u \in H_0^1(U)$ und $\partial U \in C^2$, dann gilt $u \in H^2(U)$ mit

$$\|u\|_{H^2(U)} \leq C \|f\|_{L^2(U)} .$$

Beweis. Idee: Die Vorgehensweise aus dem Beweis für die entsprechende Aussage für lineare PDE kann im Wesentlichen übertragen werden (via Differenzenquotienten). Statt der Elliptizitätsbedingung verwendet man die gleichmäßige Konvexität. □

Bemerkung. Für höhere Regularität kann man sich nicht mehr am linearen Fall orientieren: Dort differenziert man die Gleichung und erhält wiederum eine lineare Gleichung, sodass man durch *bootstrapping* höhere Regularität gewinnt, sofern die rechte Seite sie besitzt. Das funktioniert hier nicht mehr, weil die Gleichung nichtlinear ist und deshalb beim Differenzieren zunehmend kompliziertere Gestalt annimmt. Um das zu veranschaulichen, betrachten wir $f \equiv 0$. Es gilt also:

$$\int_U \sum_{i=1}^n L_{p_i}(Du) v_{x_i} \, \mathrm{d}x = 0$$

Für eine Testfunktion $w \in C_0^\infty(U)$ und $k \in \{1, \dots, n\}$ setzen wir $v := -w_{x_k}$. Da $u \in H^2(U)$, können wir partiell integrieren und erhalten:

$$0 = -\int_U \sum_{i=1}^n L_{p_i}(Du) w_{x_i x_k} \, \mathrm{d}x = \int_U \sum_{i,j=1}^n L_{p_i p_j}(Du) u_{x_j x_k} w_{x_i} \, \mathrm{d}x.$$

Mit $\tilde{u} := u_{x_k}, a^{ij}(x) := L_{p_i p_j}(Du)$ heißt das

$$\int_U \sum_{i,j=1}^n w_{x_i} \, \mathrm{d}x = 0.$$

(nach Approximation) für alle $v \in H_0^1(V)$ mit $V \subset\subset U$, also ist $\tilde{u} \in H^1(V)$ schwache Lösung von

$$\int_U \sum_{i,j} \left(a^{ij}(x) \tilde{u}_{x_j}\right)_{x_i} = 0.$$

Hierauf können wir die Regularitätstheorie linearer Gleichungen *nicht* anwenden, denn es ist nur $a^{ij} \in L^\infty(U)$ bekannt. Nach einem Theorem von DeGiorgi und Nash gilt $\tilde{u} \in C^{0,\gamma}(W)$ mit $\gamma > 0$ und $W \subset\subset V$, also $u \in C^{1,\gamma}_{\mathrm{loc}}(U)$. Wenn L glatt ist, dann gilt $a^{ij} \in C^{0,\gamma}_{\mathrm{loc}}(U)$. Nach einem Theorem von Schauder folgt nun $u \in C^{2,\gamma}_{\mathrm{loc}}(U)$. An dieser Stelle lässt sich ein Bootstrap-Argument beginnen, das bei $u \in C^\infty(U)$ endet. Siehe Literatur.

4 Die direkte Methode in L^p, Teil II: Systeme

4.1 Systeme: Existenz und Eindeutigkeit von Minimierern

(a) Notation

Wir betrachten nun eine (ausreichend) glatte Lagrange-Funktion

$$L : \bar{U} \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}, \quad (x, z, P) \mapsto L(x, z, P).$$

Wir bezeichnen die Einträge von $P \in \mathbb{R}^{m\times n}$ mit $P = (p_i^k)$, wobei $1 \le k \le m, 1 \le i \le n$, d.h., der obere Index k bezeichnet die Zeilennummer, der untere Index i die Spaltennummer. Zu L betrachten wir das entsprechende Funktional

$$I[w] := \int_U L(x, w(x), Dw(x))\,\mathrm{d}x$$

für Funktionen $w : \bar{U} \to \mathbb{R}^m, w = (w_1, \ldots, w_m),$, von denen wir voraussetzen, dass sie vorgegebene Randwerte $w = g$ auf ∂U (im Spursinne) annehmen. Dabei ist

$$Dw(x) = \begin{pmatrix} w^1_{x_1} & \cdots & w^1_{x_n} \\ \vdots & \ddots & \vdots \\ w^m_{x_1} & \cdots & w^m_{x_n} \end{pmatrix} \in \mathbb{R}^{m\times n}$$

die Jacobi-Matrix von w an der Stelle x.

(b) Euler-Lagrange-Gleichungen

Analog zu früher erhalten wir für einen genügend glatten Minimierer $u : U \to \mathbb{R}^m$ die zugehörige Euler-Lagrange-Gleichung, die jetzt ein *System quasilinearer partieller Differentialgleichungen 2. Ordnung* bilden:

$$-\sum_{i=1}^{n} \Big(L_{p_i^k}(x, u, Du) \Big)_{x_i} + L_{z^k}(x, u, Du) = 0 \quad (k = 1, \ldots, m). \tag{4.1}$$

(c) Existenz für konvexe, koerzive Funktionale

Wir nehmen eine zu (KB) analoge Ungleichung an.

Definition. I erfüllt die *Koerzivitätsbedingung*, wenn es Konstanten $\alpha > 0, \beta \ge 0$ gibt, sodass

$$L(x, z, P) \ge \alpha\, |P|^q - \beta \tag{KB2}$$

für alle $x \in U, z \in \mathbb{R}, P \in \mathbb{R}^{m\times n}$ gilt.

Der Existenzsatz aus dem skalaren Fall lässt sich nun fast wörtlich übertragen, nachdem wir die Begriffe Konvexität und Koerzivität entsprechend übertragen haben.

Theorem 4.1. *L erfülle die Koerzivitätsbedingung und sei konvex bezüglich P. Ferner sei $\mathcal{A} \neq \emptyset$. Dann existiert ein $u \in \mathcal{A}$ mit*

$$I[u] = \min_{w \in \mathcal{A}} I[w].$$

Beweis. Im Wesentlichen wie im Fall $m = 1$. □

(d) Eindeutigkeit

Auch der Beweis der Eindeutigkeit überträgt sich problemlos auf den Systemfall.

Theorem 4.2. *L erfülle die Koerzivitätsbedingung und sei konvex bezüglich P. Ferner sei $\mathcal{A} \neq \emptyset$. Dann existiert genau ein $u \in \mathcal{A}$ mit*

$$I[u] = \min_{w \in \mathcal{A}} I[w].$$

Beweis. Im Wesentlichen wie im Fall $m = 1$. □

(e) Lösung der ELGl

In Analogie zum skalaren Fall setzen wir Wachstumsbedingungen voraus. Wir nehmen an: Es gibt ein $C > 0$, sodass

$$|L(x, z, P)| \leq C\left(|P|^q + |z|^q + 1\right) \tag{4.2a}$$

$$|D_P L(x, z, P)| \leq C\left(|P|^{q-1} + |z|^{q-1} + 1\right) \tag{4.2b}$$

$$|D_z L(x, z, P)| \leq C\left(|P|^{q-1} + |z|^{q-1} + 1\right) \tag{4.2c}$$

für alle $x \in U$, $z \in \mathbb{R}$ und $P \in \mathbb{R}^{m \times n}$ gilt.

Theorem 4.3. *Angenommen, L erfüllt die Wachstumsbedingungen* (4.2) *und für $u \in \mathcal{A}$ gilt*

$$I[u] = \min_{w \in \mathcal{A}} I[w].$$

Dann ist u eine schwache Lösung von (4.1)*, d. h., es gilt*

$$\sum_{k=1}^{m} \int_U \sum_{i=1}^{n} L_{p_i^k}(x, u, Du) w_{x_i}^k + L_{z^k}(x, u, Du) w^k \,\mathrm{d}x = 0 \quad (k = 1, \ldots, m)$$

für alle $w \in W_0^{1,q}(U; \mathbb{R}^m)$.

Beweis. Im Wesentlichen wie im Fall $m = 1$. □

4.2 Polykonvexität

Bei einigen interessanten Anwendungen der Physik ist L nicht konvex bezüglich P, sodass Theorem 4.1 über die Existenz eines Minimierers nicht angewendet werden kann. Dennoch lassen sich die zugehörigen mathematischen Probleme mit der Variationsrechnung behandeln, weil sich herausstellt, dass I dennoch schwach halbstetig von unten ist.

Eine solche Anwendung ist die *nichtlineare Elastizität* (SKIZZE). Hier betrachten wir einen Körper, der das Gebiet $U \subset \mathbb{R}^3$ besetzt, die Referenzkonfiguration. Bei Verformung wird jeder Randpunkt $x \in \partial U$ auf einen Punkt $g(x) \in \mathbb{R}^3$ abgebildet und wir suchen die Verschiebung $u(x)$ für jeden inneren Punkt $x \in U$; die neue Position des Punktes x ist damit $x + u(x)$. Die lokale Verformung ist im *Deformationsgradienten* $Du(x)$ codiert. Die Spalten von $\mathrm{Id} + Du(x)$ beschreiben die deformierte Basis bei $x + u(x)$ (sofern x bezüglich der kanonischen Einheitsbasis gegeben ist). Für eine ausführliche Diskussion aktueller Probleme, siehe [3].

Ein Material heißt *hyperelastisch*, wenn es eine Energie-Dichte L gibt, sodass die realisierte Verschiebung u das zugehörige Funktional der inneren Energie

$$I[w] := \int_U L(Dw)\,\mathrm{d}x$$

unter allen zulässigen Verschiebungen $w \in \mathcal{A}$ minimiert. Hier haben wir das Material als *homogen* angenommen, d.h., L hängt nicht von x ab. Aus physikalischen Gründen ist es natürlich, dass die Energie-Dichte L, die die Änderung der inneren Energie bei Dehnung und Kompression beschreibt, auch explizit von der lokalen Volumenänderung, also $\det Dw$, abhängt. Das bedeutet, es ist naheliegend anzunehmen, dass die Lagrange-Funktion von der Form

$$L(x, P) = F(x, P, \det P)$$

ist. Beispielsweise könnte die Funktion F so aussehen:

$$F(x, P, \det P) = |P|^4 - |P|^2 \det P.$$

Dann beschreibt F im zweiten Argument Änderungen der Energie durch Verformung der Linienelemente und im dritten Argument Änderungen der Energie durch Verformung der Volumenelemente. Aus physikalischen Gründen fordert man typischerweise, dass $L(P) \to \infty$, wenn $\det P \to 0$ oder $\det P \to \infty$. Das impliziert bereits, dass L nicht konvex sein kann.

Definition (Polykonvexität)**.** Eine Lagrange-Funktion von der Form

$$L(x, z, P) = F(x, z, P, \det P) \tag{4.3}$$

heißt *polykonvex*, wenn die Abbildung

$$\mathbb{R}^{n\times n} \times \mathbb{R} \to \mathbb{R}, \quad (P, r) \mapsto F(x, z, P, r)$$

konvex ist für alle $x \in U, z \in \mathbb{R}$.

Bemerkung. Die Definition der Polykonvexität gibt es noch in einer allgemeineren Form; dabei lässt man zu, dass F in (4.3) von allen Unterminoren der Matrix Dw abhängt.

Es wird sich herausstellen, dass I in diesem Fall schwach halbstetig von unten ist. Der Existenzbeweis lässt sich dann übertragen, wobei wir statt Konvexität bezüglich eines Arguments (damals bezüglich p) nun Konvexität bezüglich (P, r) ausnutzen. Um diesen Beweis schließlich umsetzen zu können, benötigen wir die schwache Konvergenz der Determinanten.

Ist A eine $n \times n$-Matrix, so bezeichnen wir mit $\operatorname{cof} A$ die Kofaktor-Matrix von A, die an der Stelle (k, i) den Eintrag

$$(\operatorname{cof} A)_i^k = (-1)^{i+k} d(A)_i^k$$

hat, wobei $d(A)_i^k$ die Determinante der Untermatrix ist, die durch Streichen der k-ten Zeile und i-ten Spalte aus A entstanden ist.

Lemma 4.4 ($\operatorname{cof} Du$ hat divergenzfreie Zeilen)**.** *Sei $u : \mathbb{R}^n \to \mathbb{R}^n$ eine glatte Funktion. Dann gilt:*

$$\sum_{i=1}^n (\operatorname{cof} Du)_{i,x_i}^k = 0 \qquad (k = 1, \ldots, n).$$

Beweis. **1. Schritt:** Wir gehen von der Identität $(\det P)\mathrm{Id} = P^{\mathsf{T}}(\operatorname{cof} P)$ aus. Also gilt

$$(\det P)\delta_{ij} = \sum_{k=1}^n p_i^k (\operatorname{cof} P)_j^k, \tag{4.4}$$

und damit insbesondere

$$\frac{\partial \det P}{\partial p_m^k} = (\operatorname{cof} P)_m^k,$$

denn p_m^k kommt in keinem der Minoren vor (die k-te Zeile wurde ja gestrichen).

2. Schritt: Setze jetzt $P = Du$ in (4.4) ein, leite nach x_j ab und summiere über $j = 1, \ldots, n$. Das liefert

$$\sum_{j,m,k} \delta_{ij} (\operatorname{cof} Du)_m^k u_{x_m x_j}^k = \sum_{j,k} u_{x_i x_j}^k (\operatorname{cof} Du)_j^k + u_{x_i}^k (\operatorname{cof} Du)_{j,x_j}^k,$$

wobei wir die linke Seite zu

$$LHS \equiv \sum_{j,m,k} \delta_{ij} (\operatorname{cof} Du)_m^k u_{x_m x_j}^k = \sum_{j,k} (\operatorname{cof} Du)_j^k u_{x_i x_j}^k$$

vereinfachen können (setze $i = j$ und ersetze den Summationsindex m durch j), sodass dieser Ausdruck der ersten Summe auf der rechten Seite gleicht. Folglich gilt

$$\sum_{k=1}^n u_{x_i}^k \left(\sum_{j=1}^n (\operatorname{cof} Du)_{j,x_j}^k \right) = 0, \quad i = 1, \ldots, n,$$

d. h., bei jedem x_0 ist folgendes homogenes LGS erfüllt:

$$Du(x_0)v = 0, \quad \text{wobei } v = (v^1, \dots, v^n) \quad \text{mit } v^k := \sum_{j=1}^{n} (\operatorname{cof} Du)^k_{j,x_j}.$$

3. Schritt: Wenn $\det Du(x_0) \neq 0$ gilt, folgt sofort

$$\sum_{j=1}^{n} (\operatorname{cof} Du)^k_{j,x_j} = 0$$

bei x_0. Falls $\det Du(x_0) = 0$ gilt, dann wähle $\varepsilon > 0$ so klein, dass $\det(Du(x_0) + \varepsilon \mathrm{Id}) \neq 0$ gilt, führe die Argumentation für $\tilde{u} = u + \varepsilon x$ durch und schicke $\varepsilon \to 0$. □

Satz 4.5 (Schwache Stetigkeit von Determinanten)**.** *Sei $n < q < \infty$ und $u_k \rightharpoonup u$ in $W^{1,q}(U;\mathbb{R}^n)$. Dann gilt*

$$\det Du_k \rightharpoonup \det Du \quad \text{in } L^{q/n}(U).$$

- **Hilfsmittel**

(H6) Ungleichung von Morrey: $\|u\|_{C^{0,\gamma}(\overline{U})} \leq C \|u\|_{W^{1,q}(U)}$, $n < q < \infty$, $\gamma = 1 - \frac{n}{q}$.

(H7) Satz von Arzela-Ascoli: Wenn die Folge $(f_k)_{k\in\mathbb{N}}$ stetiger Funktionen $f_k : \mathbb{R}^n \to \mathbb{R}$ beschränkt und gleichgradig stetig ist, dann existiert eine Teilfolge, die gleichmäßig auf Kompakta gegen eine stetige Funktion konvergiert.

(H8) Für $P \in \mathbb{R}^{n\times n}$ gilt die Ungleichung: $|\det P| \leq \prod_{i=1}^{n} |p^i|_2 \leq |P|^n$.

Beweis. **1. Schritt:** Aufgrund der Matrixidentität $(\det P)\mathrm{Id} = P(\operatorname{cof} P)^{\mathsf{T}}$ gilt

$$\det P = \sum_{j=1}^{n} p^i_j (\operatorname{cof} P)^i_j \quad \text{für } i = 1, \dots, n.$$

2. Schritt: Nun sei $w \in C^\infty(U;\mathbb{R}^n), w = (w^1, \dots, w^n)$. Dann gilt

$$\det Dw = \sum_{j=1}^{n} w^i_{x_j} (\operatorname{cof} Dw)^i_j.$$

Nach Lemma 4.4 gilt $\sum_{j=1}^{n} (\operatorname{cof} Dw)^i_{j,x_j} = 0$, also folgt

$$\det Dw = \sum_{j=1}^{n} (w^i (\operatorname{cof} Dw)^i_j)_{x_j},$$

d.h., die Determinante von Dw lässt sich als Divergenz schreiben. Deshalb gilt für $v \in C_0^\infty(U)$

$$\int_U v \det Dw \,\mathrm{d}x = -\sum_{j=1}^{n} \int_U v_{x_j} w^i (\operatorname{cof} Dw)^i_j \,\mathrm{d}x \quad \text{für } i = 1, \dots, n. \tag{4.5}$$

3. Schritt: Gleichung (4.5) haben wir für glatte Funktionen gezeigt. Durch Standard-Approximation gilt sie auch in $W^{1,q}(U)$, insbesondere

$$\int_U v \det Du_k \,\mathrm{d}x = -\sum_{j=1}^n \int_U v_{x_j} u_k^i (\operatorname{cof} Du_k)_j^i \,\mathrm{d}x \tag{4.6}$$

für $k \in \mathbb{N}$. Da $n < q < \infty$ und $u_k \rightharpoonup u$ in $W^{1,q}(U;\mathbb{R}^n)$ folgt aus der Morreyschen Ungleichung, dass $(u_k)_{k\in\mathbb{N}}$ beschränkt ist in $C^{0,1-n/q}(U;\mathbb{R}^n)$. Nach dem Satz von Arzela-Ascoli gilt – nach evtl. Übergang zu einer Teilfolge – daher $u_k \to u$ gleichmäßig (gleichgradige Stetigkeit folgt aus der Beschränktheit in der Hölder-Norm). Aus (4.6) könnten wir schließen

$$\lim_{k\to\infty} \int_U v \det Du_k \,\mathrm{d}x = -\sum_{j=1}^n \int_U v_{x_j} u^i (\operatorname{cof} Du)_j^i \,\mathrm{d}x = \int_U v \det Du \,\mathrm{d}x, \tag{4.7}$$

wenn wir schon wüssten, dass

$$\lim_{k\to\infty} \int_U \psi (\operatorname{cof} Du_k)_j^i \,\mathrm{d}x = \int_U \psi (\operatorname{cof} Du)_j^i \,\mathrm{d}x$$

für alle i, j und $\psi \in C_0^\infty(U)$ gilt. Nun ist aber $(\operatorname{cof} Du_k)_j^i$ die Determinante einer $(n-1)\times(n-1)$-Matrix, die man wie oben behandeln könnte, sodass sie die Summe geeigneter $(n-2)\times(n-2)$-Untermatrizen mit konvergenten Faktoren ist. Induktiv so fortfahrend, muss schließlich nur die triviale Aussage, dass die Einträge von Du_k schwach gegen die Einträge von Du konvergieren, gezeigt werden (folgt aus der schwachen Konvergenz von (u_k) in $W^{1,q}$). Also gilt (4.2), mithin (4.7).

4. Schritt: Da $(u_k)_{k\in\mathbb{N}}$ in $W^{1,q}$ beschränkt ist und $|\det Du_k| \le C\,|Du_k|^n$ nach (H8) gilt, sehen wir dass $(\det Du_k)_{k\in\mathbb{N}}$ in $L^{q/n}(U)$ beschränkt ist. Also gibt es eine schwach konvergente Teilfolge, die aufgrund von (4.7) nur gegen $\det Du$ konvergieren kann. □

Dieses Lemma benutzen wir nun, um die schwache Halbstetigkeit von unten eines polykonvexen Funktionals zu zeigen.

Theorem 4.6. *Angenommen, $n < q < \infty$. Angenommen, L ist nach unten beschränkt und polykonvex. Dann ist I schwach halbstetig von unten auf $W^{1,q}(U;\mathbb{R}^n)$.*

Beweis. Wähle eine Folge $(u_k)_{k\in\mathbb{N}}$, die schwach in $W^{1,q}(U;\mathbb{R}^n)$ konvergiert. Gemäß dem Lemma gilt dann

$$\det Du_k \rightharpoonup \det Du \quad \text{schwach in } L^{q/n}(U).$$

Nun können wir fast wie im Beweis von Theorem 3.1 argumentieren. In der dortigen Notation gilt:

$$\begin{aligned} I[u_k] &= \int_U L(x, u_k, Du_k)\,\mathrm{d}x \ge \int_{G_\varepsilon} L(x, u_k, Du_k)\,\mathrm{d}x \\ &= \int_{G_\varepsilon} F(x, u_k, Du_k, \det Du_k)\,\mathrm{d}x \end{aligned}$$

$$\geq \int_{G_\varepsilon} F(x, u_k, Du, \det Du)\,\mathrm{d}x + \int_{G_\varepsilon} F_P(x, u_k, Du, \det Du) \cdot (Du_k - Du)$$
$$+ F_r(x, u_k, Du, \det Du)(\det Du_k - \det Du)\,\mathrm{d}x,$$

wobei in der letzten Zeile die Konvexität bezüglich (P, r) verwendet wurde. Wie damals können wir schließen, dass das zweite Integral für $k \to \infty$ verschwindet. □

Daraus ergibt sich unmittelbar folgender Existenzsatz, dessen Beweis wie der von Theorem 3.2 funktioniert.

Theorem 4.7. *Angenommen, $n < q < \infty$. Angenommen, L erfüllt die Koerzivitätsbedingung* (KB2) *und ist polykonvex. Sei $\mathcal{A} \neq \emptyset$. Dann existiert ein $u \in \mathcal{A}$ mit*

$$I[u] = \min_{w \in \mathcal{A}} I[w].$$

Beweis. Wie der von Theorem 3.2. □

Bemerkung. An dieser Stelle gibt es einen riesigen Unterschied zwischen dem skalaren Fall $m = 1$ und dem Fall $m > 1$ von Systemen. Im skalaren Fall haben wir gesehen, dass die Existenz eines Minimierers die Legendre-Hadamard-Bedingung impliziert. Diese ist erfüllt, wenn wir die Konvexität der Lagrange-Funktion $L(x, z, p)$ bezüglich p voraussetzen. Unter dieser Bedingung ist das Funktional schwach unterhalbstetig und wir konnten damit die Existenz eines Minimierers beweisen. Es gilt sogar die Umkehrung, dass schwache Unterhalbstetigkeit von I die Konvexität von L impliziert (siehe Dacorogna).

Im Fall von Systemen, also $m > 1$, ist das anders. Hier taucht eine Reihe von Konvexitätsbegriffen auf, die alle schwächer als Konvexität sind. Zunächst ergibt die Existenz eines Minimierers hier die *Rang-1-Konvexität* von $L(x, z, P)$ bezüglich P, das heißt, es gilt die Legendre-Hadamard-Bedingung

$$\sum_{i,j,k,l} \frac{\partial^2 L}{\partial p_i^k \partial p_j^l} \xi_i \xi_j \eta_k \eta_l \geq 0$$

für alle $x \in U, z \in \mathbb{R}^m, P \in \mathbb{R}^{m \times n}, \xi \in \mathbb{R}^n, \eta \in \mathbb{R}^m$, oder, äquivalenterweise, die skalare Funktion

$$f(t) := L(x, z, P + t\eta \otimes \xi)$$

ist konvex. Es ist aber an dieser Stelle nicht klar, ob diese Bedingung für schwache Unterhalbstetigkeit des zugehörigen Funktionals ausreicht.

Interessante Funktionale (etwa der Elastizitätstheorie) sind polykonvex. Unter dieser Bedingung ist das Funktional schwach unterhalbstetig. Aber die Umkehrung gilt hier nicht: Die richtige Bedingung an $L(x, z, P)$, die äquivalent zur schwachen Unterhalbstetigkeit ist, lautet *Quasikonvexität*:

$$\forall P \in \mathbb{R}^{m \times n}, v \in C_0^\infty(U; \mathbb{R}^m): \quad \int_U L(x, z, P)\,\mathrm{d}x \leq \int_U L(x, z, P + Dv)\,\mathrm{d}x.$$

Diese Bedingung ist jedoch für konkrete Anwendungen im Allgemeinen schwer zu prüfen. Es gelten die Implikationen:

$$L \text{ konvex} \Longrightarrow L \text{ polykonvex} \Longrightarrow L \text{ quasikonvex} \Longrightarrow L \text{ Rang-1-konvex}$$

Im Allgemeinen sind alle Umkehrungen falsch; dabei war die Frage, ob Rang-1-Konvexität Quasikonvexität impliziert lange offen (und ist es etwa für $n = m = 2$ immer noch). Ausführlichste Diskussion all dieser Begriffe findet man etwa bei Dacorogna. Insbesondere findet man dort auf Seite 221 die Funktion

$$f_\gamma : \mathbb{R}^{2\times 2} \to \mathbb{R}, P \mapsto |P|^2 \left(|P|^2 - 2\gamma \det P\right),$$

die jede der Eigenschaften für bestimmte Werte von γ hat.

Im Falle $m = 1$ oder $n = 1$ fallen alle genannten Konvexitätsbegriffe zusammen.

4.3 Null-Lagrangians

Eine interessante Klasse von Lagrange-Funktionen erhält man, wenn man fordert, dass jede (genügend) glatte Funktion eine Lösung der zugehörigen Euler-Lagrange-Gleichung ist.

Definition. Die Funktion L heißt *Null-Lagrangian*, wenn das System der Euler-Lagrange-Gleichungen,

$$(4.8) \qquad -\sum_{i=1}^{k} \left(L_{p_i^k}(x,u,Du)\right)_{x_i} + L_{z^k}(x,u,Du) = 0 \quad \text{in } U,\ k = 1,\dots,m,$$

für alle glatten Funktionen $u : U \to \mathbb{R}^m$ erfüllt ist.

Die Bedeutung von Null-Lagrangians kommt daher, dass die zugehörige „Energie“

$$I[w] = \int_U L(x,w,Dw)\,\mathrm{d}x$$

nur von den Randwerten abhängt, was das folgende Theorem zeigt.

Theorem 4.8. *Angenommen, L ist ein Null-Lagrangian und $u, \tilde{u} \in C^2(\bar{U},\mathbb{R}^m)$ erfüllen*

$$(4.9) \qquad u \equiv \tilde{u} \quad \text{auf } \partial U.$$

Dann ist

$$(4.10) \qquad I[u] = I[\tilde{u}].$$

Beweis. Wir definieren $i(\tau) := I[\tau u + (1-\tau)\tilde{u}]$ für $\tau \in [0,1]$. Dann gilt:

$$i'(\tau) = \int_U \sum_{i=1}^{n}\sum_{k=1}^{m} L_{p_i^k}(x, \tau u + (1-\tau)\tilde{u}, \tau Du + (1-\tau)D\tilde{u})(u^k_{x_i} - \tilde{u}^k_{x_i})$$

$$
\begin{aligned}
&\quad + \sum_{k=1}^{m} L_{z^k}(x, \tau u + (1-\tau)\tilde{u}, \tau Du + (1-\tau)D\tilde{u})(u^k - \tilde{u}^k)\,\mathrm{d}x \\
&= \sum_{k=1}^{m} \int_U \Big[-\sum_{i=1}^{n} \Big(L_{p_i^k}(x, \tau u + (1-\tau)\tilde{u}, \tau Du + (1-\tau)D\tilde{u}) \Big)_{x_i} \\
&\quad + L_{z^k}(x, \tau u + (1-\tau)\tilde{u}, \tau Du + (1-\tau)D\tilde{u}) \Big] (u^k - \tilde{u}^k)\,\mathrm{d}x \\
&\quad + \int_{\partial U} L_{p_i^k}(x, \tau u + (1-\tau)\tilde{u}, \tau Du + (1-\tau)D\tilde{u})(u^k - \tilde{u}^k)\nu^i\,\mathrm{d}S \\
&= 0.
\end{aligned}
$$

Die letzte Gleichheit folgt, da einerseits auch $\tau u + (1-\tau)\tilde{u}$ die Euler-Lagrange-Gleichung erfüllt und da andrerseits $u \equiv \tilde{u}$ auf ∂U gilt. □

Im skalaren Fall $m = 1$ ist jeder Null-Lagrangian automatisch linear. Für $m > 1$ jedoch gibt es eine Reihe von interessanten Anwendungen. Die Situation, dass ein Integral nur von den Randwerten abhängt ist aus anderen Kontexten bekannt, z. B. Abbildungsgrad, Homotopieinvarianz der Kurvenintegrale, Satz von Gauss-Bonnet. Ein interessantes Beispiel schauen wir uns nun an, samt einer interessanten Anwendung – dem Brouwerschen Fixpunktsatz.

Satz 4.9. *Die Determinanten-Funktion*

$$L(P) = \det P \quad \text{für } P \in \mathbb{R}^{n \times n}$$

ist ein Null-Lagrangian.

Zum Beweis benötigen wir nun erneut Lemma 4.4: cof Du *hat divergenzfreie Zeilen.*

Beweis. Wir müssen zeigen, dass für jede glatte Funktion $u : U \to \mathbb{R}^n$ die Gleichungen

$$\sum_{i=1}^{n} \Big(L_{p_i^k}(Du) \Big)_{x_i} = 0, \quad k = 1, \ldots, n,$$

erfüllt sind. Es gilt $L_{p_i^k}(P) = (\operatorname{cof} P)_i^k$. Damit gilt gemäß dem Lemma:

$$\sum_{i=1}^{n} \Big(L_{p_i^k}(Du) \Big)_{x_i} = \sum_{i=1}^{n} (\operatorname{cof} A)_{i,x_i}^k = 0.$$

□

Daraus erhalten wir einen kurzen analytischen Beweis des Brouwerschen Fixpunktsatzes.

Theorem 4.10. *Sei $u : B(0,1) \to B(0,1)$ eine stetige Abbildung, wobei $B(0,1) \in \mathbb{R}^n$ die abgeschlossene Einheitskugel bezeichnet. Dann besitzt u einen Fixpunkt, d. h., es existiert ein $x \in B(0,1)$ mit $u(x) = x$.*

Beweis. **1. Schritt:** Wir schreiben $B = B(0,1)$ und behaupten zunächst, dass es keine glatte Funktion

$$w : B \to \partial B \quad \text{mit } w(x) = x \quad \forall x \in \partial B \tag{4.11}$$

(Eine solche Abbildung heißt *Retraktion.*) Angenommen, es gäbe eine solche Funktion. Wir bezeichnen mit $\tilde{w}$ die Identität auf B, also $\tilde{w}(x) = x$ für alle $x \in B$. Es gilt $w \equiv \tilde{w}$ auf ∂B und die Determinante ist ein Null-Lagrangian. Deshalb impliziert Theorem 4.8:

$$\int_B \det Dw \,\mathrm{d}x = \int_B \det D\tilde{w} \,\mathrm{d}x = |B| \neq 0. \tag{4.12}$$

Andererseits folgt aus der Definition von w, dass $|w(x)|^2 \equiv 1$ gilt. Differenzieren liefert $Dw(x)^\mathsf{T} w(x) = 0$. Da $|w(x)| = 1$, also $w(x) \neq 0$, heißt das, 0 ist ein Eigenwert von $Dw(x)^\mathsf{T}$ für jedes $x \in B$. Deshalb gilt $\det Dw(x) \equiv 0$ in B. Das ist ein Widerspruch zu (4.12). Also existiert keine glatte Funktion w wie in (4.11).

2. Schritt: Als nächstes zeigen wir, dass es auch keine stetige Funktion w der Art (4.11) gibt. Wenn es eine solche Funktion w gäbe, dann setzen wir sie auf $\mathbb{R}^n$ fort durch $w(x) = x$ für $x \in \mathbb{R}^n \setminus B$. Dann gilt $w(x) \neq 0$ für alle $x \in \mathbb{R}^n$. Wähle $\varepsilon > 0$ so klein, dass $w_1 := \eta_\varepsilon * w$ ebenfalls $w_1(x) \neq 0$ für alle $x \in \mathbb{R}^n$ erfüllt. Da η_ε radial ist, gilt $w_1(x) = x$ für $x \in \mathbb{R}^n \setminus B(0,2)$ für hinreichend kleines ε. (Hierfür beachte, dass

$$(\eta_\varepsilon * w)(x) = \int_{B_\varepsilon} \eta_\varepsilon(y)(x-y)\,\mathrm{d}y = \int_{B_\varepsilon} \eta_\varepsilon(y) x \,\mathrm{d}y - \int_{B_\varepsilon} \eta_\varepsilon(y) y \,\mathrm{d}y = x - 0,$$

wobei das zweite Integral 0 ergibt – nach Transformationssatz oder anschaulich.) Dann ist

$$w_2(x) := \frac{w_1(2x)}{|w_1(2x)|}$$

eine Abbildung $w_2 : B \to \partial B$ mit $w_2|_{\partial B} = \mathrm{Id}|_{\partial B}$, also von der Art (4.11) – im Widerspruch zum 1. Schritt.

3. Schritt: Angenommen, $u : B \to B$ habe keinen Fixpunkt. Dann definieren wir eine Abbildung $w : B \to \partial B$, indem x auf den Randpunkt abgebildet wird, wo der Strahl ausgehend von $u(x)$ durch x auf ∂B trifft. Diese Abbildung ist wohldefiniert, da $u(x) \neq x$, und stetig. Somit ist sie von der Art (4.11) – im Widerspruch zum 2. Schritt.
□

5 *Spezielle Themen

Da dieser Abschnitt aus Zeitgründen nur angerissen wurde, sind größtenteils lediglich die Aussagen angegeben. Für umfassendes Material: Siehe Literaturliste.

5.1 Extrema unter Nebenbedingungen

In diesem Abschnitt betrachten wir exemplarisch Minimierungsprobleme, bei denen zusätzlich zur Randbedingung noch weitere Einschränkungen an die Lösung bestehen. Wir werden sehen, wie hier Lagrange-Multiplikatoren ins Spiel kommen und wie sie sich in der ELGl wiederfinden.

(a) Nichtlineare Eigenwertprobleme

Als erstes betrachten wir eine Nebenbedingung, die anhand eines Integrals gegeben ist. Der Konkretheit halber betrachten wir

$$I[w] := \frac{1}{2} \int_U |Dw|^2 \, \mathrm{d}x \tag{5.1}$$

mit $w = 0$ auf ∂U. Zusätzlich fordern wir, dass

$$J[w] := \int_U G(w) \, \mathrm{d}x = 0 \tag{5.2}$$

gilt, wobei $G : \mathbb{R} \to \mathbb{R}$ eine gegebene glatte Funktion ist. Wir schreiben $g = G'$ und nehmen an, dass mit einer Konstanten $C > 0$ die Ungleichungen

$$|g(z)| \leq C(|z| + 1), \text{ also auch } \quad |G(z)| \leq C(|z|^2 + 1) \tag{5.3}$$

gelten. Im Folgenden suchen wir einen Minimierer in:

$$\mathcal{A} := \{w \in H_0^1(U) : J[w] = 0\}.$$

Wir nehmen zudem an, dass $U \subset \mathbb{R}^n$ offen, beschränkt und zusammenhängend ist.

Theorem 5.1 (Existenz eines Minimierers)**.** *Angenommen, $\mathcal{A} \neq \emptyset$. Dann existiert ein $u \in \mathcal{A}$ mit*

$$I[u] = \min_{w \in \mathcal{A}} I[w].$$

Beweis. Wie üblich wählen wir eine minimerende Folge $(u_k)_{k \in \mathbb{N}} \subset \mathcal{A}$ mit

$$I[u_k] \to m := \inf_{w \in \mathcal{A}} I[w].$$

Dann finden wir eine schwach konvergente Teilfolge

$$u_{k_j} \rightharpoonup u \quad \text{ in } H_0^1(U)$$

mit $I[u] \le m$ (I ist schwach unterhalbstetig). Wir sind fertig, wenn wir zeigen können, dass $J[u] = 0$ gilt. Da die Einbettung $H_0^1(U) \subset L^2(U)$ kompakt ist, gilt $u_{k_j} \to u$ in $L^2(U)$. Folglich ist

$$\begin{aligned}
|J[u]| = \left|J[u] - J[u_{k_j}]\right| &\le \int_U \left|G(u) - G(u_{k_j})\right| \mathrm{d}x \\
&\le \int_U \underbrace{\left|G'(\xi)\right|}_{=|g(\xi)|\le C(1+|\xi|)\le C(1+|u|+\left|u_{k_j}\right|)} \left|u - u_{k_j}\right| \mathrm{d}x \\
&\le C \left\|u - u_{k_j}\right\| \left\|1 + |u| + \left|u_{k_j}\right|\right\| \to 0.
\end{aligned}$$

Dabei haben wir den Mittelwertsatz, die Abschätzung (5.3), die Cauchy-Schwarz-Ungleichung und

$$\left\|1 + |u| + \left|u_{k_j}\right|\right\| \le C(1 + \|u\| + \left\|u_{k_j}\right\|) \le C, \text{da } u, u_{k_j} \in H_0^1(U),$$

verwendet. □

Wir betrachten als nächstes die zugehörige Euler-Lagrange-Gleichung.

Theorem 5.2 (Lagrange-Multiplikatoren)**.** *Angenommen, $u \in \mathcal{A}$ erfüllt*

$$I[u] = \min_{w \in \mathcal{A}} I[w].$$

Dann existiert $\lambda \in \mathbb{R}$, sodass gilt:

$$\int_U Du \cdot Dv \,\mathrm{d}x = \lambda \int_U g(u) v \,\mathrm{d}x \qquad \forall v \in H_0^1(U). \tag{5.4}$$

Bemerkung. Das bedeutet, u ist eine schwache Lösung des nichtlinearen Randwertproblems

$$\begin{cases} -\Delta u = \lambda g(u) & \text{in } U, \\ \quad u = 0 & \text{auf } \partial U, \end{cases} \tag{5.5}$$

wobei λ den Lagrange-Multiplikator bezeichnet, der zur Nebenbedingung $J[u] = 0$ gehört.

Beweis. **1. Schritt:** Wir wählen $v \in H_0^1(U)$. Zunächst nehmen wir an,

$$g(u) \not\equiv 0 \quad \text{in } U, \tag{5.6}$$

d.h., $g(u)$ ist nicht f.ü. gleich 0. Dann können wir eine Funktion $w \in H_0^1(U)$ so wählen, dass

$$\int_U g(u) w \,\mathrm{d}x \ne 0. \tag{5.7}$$

Nun betrachten wir $j(\tau,\sigma) := J[u+\tau v+\sigma w] = \int_U G(u+\tau v+\sigma w)\,\mathrm{d}x$ für $\tau,\sigma\in\mathbb{R}$. Offenbar ist $j(0,0)=\int G(u)=0$. Außerdem ist j eine C^1-Funktion mit

$$\frac{\partial j}{\partial\tau}(\tau,\sigma) = \int_U g(u+\tau v+\sigma w)v\,\mathrm{d}x, \quad \frac{\partial j}{\partial\sigma}(\tau,\sigma) = \int_U g(u+\tau v+\sigma w)w\,\mathrm{d}x.$$

Gleichung (5.7) impliziert, dass $\frac{\partial j}{\partial\sigma}(0,0)\neq 0$. Also können wir den Satz über implizite Funktionen anwenden und erhalten eine C^1-Funktion $\varphi:\mathbb{R}\to\mathbb{R}$ mit

$$\varphi(0)=0, \quad j(\tau,\varphi(\tau))=0 \quad \text{für } |\tau|\ll 1.$$

Leiten wir die Relation ab, so erhalten wir

$$\frac{\partial j}{\partial\tau}(\tau,\varphi(\tau)) + \frac{\partial j}{\partial\sigma}(\tau,\varphi(\tau))\varphi'(\tau) = 0,$$

also

$$\varphi'(0) = -\frac{\int_U g(u)v\,\mathrm{d}x}{\int_U g(u)w\,\mathrm{d}x}.$$

2. Schritt: Nun setze $w(\tau) := \tau v+\varphi(\tau)w$ (für kleines τ) und schreibe $i(\tau) := I[u+w(\tau)]$. Es ist $J[u+w(\tau)]=0$, also $w(\tau)\in\mathcal{A}$. Also hat die C^1-Funktion $i(\cdot)$ ein Minimum bei $\tau=0$. Daher gilt:

$$\begin{aligned} 0 = i'(0) &= \int_U (Du+\tau Dv+\varphi(\tau)Dw)\cdot(Dv+\varphi'(\tau)Dw)\,\mathrm{d}x|_{\tau=0} \\ &= \int_U Du\cdot(Dv+\varphi'(0)Dw)\,\mathrm{d}x. \end{aligned}$$

Wir definieren jetzt

$$\lambda := \frac{\int_U Du\cdot Dw\,\mathrm{d}x}{\int_U g(u)w\,\mathrm{d}x}$$

und erhalten die gesuchte Gleichung

$$\int_U Du\cdot Dw\,\mathrm{d}x = \lambda\int_U g(u)v\,\mathrm{d}x \quad \forall v\in H_0^1(U).$$

3. Schritt: Nun nehmen wir an, anstelle von (5.7) gelte

$$g(u)=0 \quad \text{f.ü. in } U.$$

Indem wir g durch beschränkte Funktionen approximieren, finden wir $DG(u)=g(u)Du=0$ f.ü. Daher ist $G(u)=\text{const}$ f.ü., denn U ist zusammenhängend. Aus $J[u]=\int_U G(u)=0$ folgt somit $G(u)=0$ f.ü. Da $u=0$ auf dem Rand gilt, folgt nun $G(0)=0$. Aber dann muss $u=0$ f.ü. sein, denn andernfalls wäre $I[u]>I[0]=0$, also u kein Minimierer. Da $g=0$ f.ü. gilt, ist Gleichung (5.4) trivialerweise für jedes $\lambda\in\mathbb{R}$ erfüllt. $\square$

(b) Inkompressibilität: Stokes-Problem

Nebenbedingungen können sich auf recht unterschiedliche Weise auf die ELGl auswirken. Wir schauen uns ein weiteres Beispiel an. Für die stationäre Verteilung einer homogenen Flüssigkeit im Gebiet $U \subset \mathbb{R}^3$ betrachtet man

$$I[w] := \int_U \frac{1}{2} |Dw|^2 - fw \, \mathrm{d}x,$$

wobei w das Geschwindigkeitsfeld und f eine gegebene externe Kraft ist. Eine inkompressible Flüssigkeit erfüllt $\operatorname{div} w = 0$, also ist die Menge der zulässigen Funktionen

$$\mathcal{A} = \left\{ w \in H_0^1(U; \mathbb{R}^3) : \operatorname{div} w = 0 \right\}.$$

Mit den üblichen Methoden lässt sich die Existenz eines eindeutigen Minimierers zeigen. Besonders interessant ist hierbei, wie sich die Nebenbedingung auf die ELGl auswirkt.

Theorem 5.3 (Druck als Lagrange-Multiplikator)**.** *Es existiert eine skalare Funktion $p \in L^2_{loc}(U)$, sodass*

$$\int_U Du : Dv \, \mathrm{d}x = \int_U p \operatorname{div} v + fv \, \mathrm{d}x$$

für alle $v \in H^1(U; \mathbb{R}^3)$ mit kompaktem Träger in U gilt.

Das bedeutet, (u, p) ist eine schwache Lösung des Stokes-Problems

$$\begin{aligned} -\Delta u &= f - Dp \quad \text{in } U, \\ \operatorname{div} u &= 0 \quad \text{in } U, \\ u &= 0 \quad \text{auf } \partial U. \end{aligned}$$

Mit anderen Worten: Die Funktion p ist der zur Nebenbedingung $\operatorname{div} u = 0$ gehörige Lagrange-Multiplikator.

Beweis. Siehe [5, 8.4.4]. □

5.2 Sattelpunkte und das Mountain-Pass-Theorem

(a) Motivation: Morse-Index und Gestalt der Levelsets

Wir betrachten die Oberfläche eines Torus $H := T^2 \subset \mathbb{R}^3$, der aufrecht auf der x-y-Ebene steht, und die entsprechende *Höhenfunktion* $I : T \to \mathbb{R}, u := (x, y, z) \mapsto z$.

Wir untersuchen nun *kritische Punkte* dieses Funktionals, d.h. Punkte $u \in T^2$ mit $I'[u] = 0$. (Um das konkret zu tun, müssen zunächst geeignete Koordinaten gewählt werden.) Der Funktionswert $I[u]$ bei einem solchen u heißt *kritischer Wert*. Bei einem kritischen Wert $c \in \mathbb{R}$ gilt also

$$K_c := \{u \in H : I[u] = c, I'[u] = 0\} \neq \emptyset.$$

Wir sehen hier, dass I genau vier kritische Punkte $u_i, i = 1, \ldots, 4$, hat und diese entsprechen vier kritischen Werten $c_1 < c_2 < c_3 < c_4, c_i := I[u_i]$. Wir untersuchen nun den Zusammenhang zwischen dem Morse-Index von I bei u_i und den *Sublevelsets*

$$A_c := \{u \in H : I[u] \leq c\}$$

für $c \in \mathbb{R}$. Natürlich ist $A_c = \emptyset$, wenn $c < 0$. Nun betrachten wir die Mengen A_c für wachsendes c. Hierfür stellen wir uns vor, der Torus werde mit Wasser bis zur Höhe c gefüllt, und wir betrachten die „topologische Gestalt" des Teils der Torusoberfläche, die von Wasser bedeckt ist, d.h., wir betrachten den Homöomorphie-Typ der Sublevelsets. Offensichtlich (siehe Skizze) passiert genau dann eine topologische Veränderung von A_c, wenn c einen kritischen Wert durchläuft. (k-Zelle $= B^k =$ Ball in $\mathbb{R}^k$) In der folgenden Tabelle beschreiben wir das: Links findet sich jeweils der Typ von A_c ($\simeq$ bedeutet homotopieäquivalent), rechts finden sich die Morse-Indizes.

$c_4 < c$	$A_c = T$
$c_3 < c < c_4$	$A_c \simeq$ „Rohr mit 1 Henkel"
$c_2 < c < c_3$	$A_c \simeq$ „Korb mit 1 Henkel"
$c_1 < c < c_2$	$A_c = B^2$
$c < c_1$	$A_c = \emptyset$

c_4	2
c_3	1
c_2	1
c_1	0

Wir können hiermit den „Hauptsatz der Morse-Theorie" bestätigen, nämlich:

- Ist c kein kritischer Wert, dann sind $A_{c+\varepsilon}$ und $A_{c-\varepsilon}$ homotopieäquivalent (das ist schwächer als homöomorph).
- Für einen kritischen Wert c mit einem kritischen Punkt mit Morse-Index k entsteht $A_{c+\varepsilon}$ aus $A_{c-\varepsilon}$ durch „Anheften einer $n-k$-Zelle B^{n-k}" (bis auf Homotopieäquivalenz).

Um die erste Aussage zu beweisen, wird folgende Anschauung zu einem rigorosen Argument ausgebaut: Ist c kein kritischer Wert, dann untersuche, wie sich die Punkte in $A_{c-\varepsilon}$ verhalten, wenn „der Wasserpegel steigt", d.h., wenn wir von $c - \varepsilon$ zu $c + \varepsilon$ übergehen. Mathematisch entspricht dies der Untersuchung der Trajektorien der ODE $\dot{x} = \nabla I(x)$ (sogenannter *Gradientenfluss*). Da das Vektorfeld abseits von Fixpunkten parallelisiert werden kann, ist der zugehörige Fluss ein Diffeomorphismus. Dann sind die Mengen $A_{c-\varepsilon}$ und $A_{c+\varepsilon}$ homöomorph. Für eine Einführung in die Morse-Theorie: Siehe Milnor [9].

(b) Kritische Punkte und Deformationen

Wir wenden uns jetzt der Untersuchung von kritischen Punkten zu, die im Allgemeinen keine Minimierer sind. Solche Probleme treten insbesondere bei nach oben und unten *unbeschränkten* Funktionalen I auf. Hintergrundwissen findet sich bei Struwe [10].

Beispiel: Die PDE

$$-\Delta u = |u|^{r-1} u$$

ist die ELGl zu dem nach oben und unten unbeschränkten Funktional

$$I[w] := \int_U \frac{1}{2}|Dw|^2 - \frac{1}{r+1}|w|^{r+1}\,\mathrm{d}x \quad \text{mit } L(x,z,p) = \frac{1}{2}|p|^2 - \frac{z^{r+1}}{r+1}.$$

Im Folgenden sei H ein Hilbertraum mit Skalarprodukt $(\cdot,\cdot)$ und Norm $\|\cdot\|$. Wir betrachten ein nichtlineares Funktional $I : H \to \mathbb{R}$.

Definition. I ist *differenzierbar* bei $u \in H$, wenn es ein $v \in H$ gibt, sodass gilt:

$$\forall w \in H: \quad I[w] = I[u] + (v, w-u) + o(\|w-u\|).$$

Das Element v ist eindeutig, falls es existiert, und wir schreiben dann: $I'[u] = v$.

Wir schreiben $I \in \mathcal{C}$, wenn: (i) $I'[u]$ existiert für alle $u \in H$, (ii) die Abbildung $I' : H \to H$ ist stetig, (iii) die Abbildung $I' : H \to H$ ist Lipschitz-stetig auf beschränkten Teilmengen von H.

Bemerkung. Die Forderung (iii) wäre nicht notwendig, aber vereinfacht die Beweise. Erfüllt I nur die Forderungen (i) und (ii) schreibt man $I \in C^1(H;\mathbb{R})$.

Für $c \in \mathbb{R}$ schreiben wir weiterhin

$$K_c := \{u \in H : I[u] = c, I'[u] = 0\} \quad \text{und} \quad A_c := \{u \in H : I[u] \leq c\}.$$

Definition. (i) Wir nennen $u \in H$ einen *kritischen Punkt*, wenn $I'[u] = 0$.
(ii) Wir nennen $c \in \mathbb{R}$ einen *kritischen Wert*, wenn $K_c \neq \emptyset$.

Wir wollen im Folgenden zeigen, dass in der vorliegenden Hilbertraum-Situation immer noch $A_{c-\varepsilon}$ und $A_{c+\varepsilon}$ die gleiche topologische Gestalt haben, wenn c kein kritischer Wert ist. Die Beweisidee besteht wiederum in der Betrachtung des Gradientenflusses.

Definition. Ein Funktional $I \in \mathcal{C}$ erfüllt die *Palais-Smale-Kompaktheitsbedingung (PSK)*, wenn jede Folge $(u_k)_{k\in\mathbb{N}} \subset H$ mit

- $\{I[u_k]\}_{k\in\mathbb{N}}$ ist beschränkt
- $I'[u_k] \to 0$ in H

eine konvergente Teilfolge besitzt.

Theorem 5.4 (Deformationstheorem)**.** *Angenommen, $I \in \mathcal{C}$ erfüllt die (PSK) und $K_c = \emptyset$. Dann gibt es für jedes hinreichend kleine $\varepsilon > 0$ eine Konstante $\delta \in (0,\varepsilon)$ und eine Funktion („Homotopie")*

$$\eta \in C([0,1] \times H; H),$$

sodass die Abbildungen $\eta_t(u) := \eta(t,u)$ folgende Eigenschaften haben:

(i) $\eta_0(u) = u$ *für alle* $u \in H$,

(ii) $\eta_1(u) = u$ *für alle* $u \notin I^{-1}[c-\varepsilon, c+\varepsilon]$,

(iii) $I[\eta_t(u)] \leq I[u]$,

(iv) $\eta_t(A_{c+\delta}) \subset A_{c-\delta}$.

(c) Das Mountain-Pass-Theorem

Als Anwendung des Deformationstheorems werden wir nun einen Satz behandeln, der die Existenz eines Sattelpunkts garantiert.

Theorem 5.5. *Angenommen, $I \in \mathcal{C}$ erfüllt (PSK). Außerdem gelten:*

(i) $I[0] = 0$.

(ii) *Es gibt Konstanten $r, a > 0$, sodass $I[u] \geq a$ gilt, falls $\|u\| = r$.*

(iii) *Es gibt ein $v \in H$ mit $\|v\| > r, I[v] \leq 0$.*

Wir definieren

$$\Gamma := \{g \in C([0,1]; H) : g(0) = 0, g(1) = v\}.$$

Dann ist

$$c = \inf_{g \in \Gamma} \max_{0 \leq t \leq 1} I[g(t)]$$

ein kritischer Wert von I.

Anschaulich bedeuten die Voraussetzungen, dass es in der I-Landschaft bei 0 eine Senke gibt, die von einer Bergkette umgeben ist. Außerdem gibt es „jenseits der Berge“ eine weitere Senke v, die nicht höher als 0 liegt. Die Wege, die 0 und v verbinden, führen notwendig über den Bergpass. Betrachtet man entlang jedes Wegs das Maximum von I, dann liefert das Infimum über alle Wege einen kritischen Wert von I.

(d) Anwendung auf eine semilineare elliptische PDGl

Wir kommen nun zu einer Anwendung des Mountain-Pass-Theorems auf eine semilineare elliptische Gleichung, die zu einem indefinite Funktional gehört.

Theorem 5.6. *Sei f eine glatte Funktion und mit einem $1 < p < \frac{n+2}{n-2}$ und einer Konstanten $C > 0$ gelte*

$$|f(z)| \leq C(1 + |z|^p), \quad |f'(z)| \leq C(1 + |z|^{p-1})$$

für alle $z \in \mathbb{R}$. Außerdem gelte für ein $0 < \gamma < \frac{1}{2}$

$$0 \leq F(z) \leq \gamma f(z) z, \qquad F(z) := \int_0^z f(s)\,\mathrm{d}s.$$

Schließlich gebe es $0 < a \leq A$, sodass

$$a\,|z|^{p+1} \leq |F(z)| \leq A\,|z|^{p+1}.$$

Dann besitzt das Randwertproblem

$$\begin{cases} -\Delta u = f(u) & \text{in } U, \\ u = 0 & \text{auf } \partial U, \end{cases}$$

eine schwache Lösung $u \not\equiv 0$.

Insbesondere die oben erwähnte Gleichung $-\Delta u = |u|^{r-1}\,u$ fällt unter dieses Theorem.

5.3 Anwendung auf travelling waves der KdV-Gleichung

(a) Vorbemerkung und Hamiltonsche Struktur

Im Kontext von travelling waves treten Variationsprobleme auf, zum Beispiel bei der Korteweg-deVries-Gleichung

$$\partial_t U = -\partial_x(U^2) - \partial_{xxx} U, \tag{5.8}$$

die unter anderem als Modellgleichung für nichtlineare Wasserwellen in einem Kanal dient. Wir werden nun die Existenz und Stabilität von Solitonen betrachten. Dabei handelt es sich um beidseits abklingende travelling waves.

Praktischerweise gehen wir zu mitbewegten Koordinaten $(t, \xi = x - ct)$ über, d.h., wir setzen $U(t,x) = u(t,\xi)$, wobei c die noch nicht bestimmte Geschwindigkeit eines gesuchten Solitons ist. Dadurch wird aus (5.8)

$$\partial_t u = c\partial_\xi u - \partial_\xi(u^2) - \partial_{\xi\xi\xi} u = \partial_\xi(cu - u^2 - \partial_{\xi\xi} u). \tag{5.9}$$

Diese Gleichung ist ein Beispiel eines unendlich-dimensionalen Hamiltonschen Systems oder einer Hamiltonschen PDGl. Die Gleichung lässt sich nämlich als

$$u_t = JH_c'[u] \tag{5.10}$$

schreiben[1] mit dem (bezüglich des L^2-Skalarprodukts) schiefsymmetrischen Operator $J = \partial_\xi$ und der Energie

$$H_c[u] = \int_{\mathbb{R}} \frac{1}{2}(u_x)^2 - \frac{1}{3}u^3 + \frac{c}{2}u^2 \,\mathrm{d}x.$$

Dann ist tatsächlich

$$H_c'[u] = -(\partial_{\xi\xi} u + u^2 - cu).$$

(b) Existenz von Solitonen

Solitonen sind abklingende travelling waves, also Lösungen der Form $u(t,\xi) = \varphi^c(\xi)$, d.h., in den Koordinaten (t,ξ) sind sie stationär. Also erfüllen sie

$$0 = JH_c'[\varphi^c].$$

Durch Integration und unter Verwendung des Abklingens (d.h. Integrationskonstante $= 0$) findet man also die (sogenannte *Profil*-)Gleichung

$$0 = H_c'(\varphi^c),$$

[1] Ein endlich-dimensionales Hamiltonsches System hat die Gestalt

$$\begin{pmatrix} q \\ p \end{pmatrix}_t = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \nabla_{q,p} H(q,p).$$

Setzt man $u = (q,p)$ und $J = \left(\begin{smallmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{smallmatrix}\right)$, dann findet man formal dieselbe Struktur in (5.10) wieder. Wenn J die symplektische Standardmatrix ist, nennt man das Hamiltonsche System *kanonisch.* In der angegbenen Version ist die KdV-Gleichung kein kanonisches Hamiltonsches System.

d.h., Solitonen sind stationäre Punkte des Funktionals H_c! Mit der unten eingeführten Bezeichnung $H_c = H - cI$ kann man das auch so ausdrücken: Solitonen sind Extrema von H unter der Nebenbedingung $I = \text{const}$, wobei c als Lagrange-Multiplikator auftritt.

Um diese Gleichung konkret zu lösen, betrachtet man das entsprechende ebene dynamische System und stellt fest, dass dieses selbst hamiltonsch ist. Ein Soliton entspricht dann einem homoklinen Orbit, der $(0,0)$ mit sich selbst verbindet. Auf diese Weise findet man für jedes $c > 0$ eine Schar von Solitonen $\varphi^{c,\delta} = \overline{\varphi}^c(\cdot + \delta)$ mit $\delta \in \mathbb{R}$, wobei $\overline{\varphi}^c$ das Profil des einzigen symmetrischen Solitons sei. In Formeln:

$$\overline{\varphi}^c(\xi) = \frac{3c}{2} \operatorname{sech}^2\left(\frac{\sqrt{c}}{2}\xi\right).$$

Im Folgenden lassen wir den Querstrich weg.

(c) Stabilität von Solitonen

Um die Stabilität (im Sinne von: Bleiben Lösungen, die in der „Nähe" von φ^c starten, für alle Zeiten in der Nähe?) zu untersuchen, kommt zunächst die Frage auf, ob das Soliton ein Minimum von H_c ist, d.h., ob $H_c''[\varphi^c]$ positiv definit ist. In diesem Fall wären Lösungen, die in der Nähe starten, innerhalb des Paraboloids gefangen und würden in der Nähe bleiben.

Jedoch ist das nicht der Fall, denn: Leitet man die Profilgleichung $H_c'[\varphi^c] = 0$ nach ξ ab, so erhält man $H_c''[\varphi^c]\partial_\xi \varphi^c = 0$. Also ist 0 ein Eigenwert von $H_c''[\varphi^c]$. Da es sich um einen Sturm-Liouville-Operator handelt und die zugehörige Eigenfunktion $\partial_\xi \varphi^c$ genau eine Nullstelle besitzt (φ^c besitzt genau ein lokales Extremum), gibt es genau einen kleineren Eigenwert $\mu < 0$, der zudem einfach ist.

Andererseits sind die Bestandteile von $H_c \equiv H - cI$, nämlich

$$H[u] = \int_{\mathbb{R}} \frac{1}{2}(u_x)^2 - \frac{1}{3}u^3 \, \mathrm{d}x, \qquad I[u] = \int_{\mathbb{R}} -\frac{1}{2}u^2 \, \mathrm{d}x,$$

Erhaltungsgrößen der PDGl. Deshalb genügt es für die Betrachtung der Stabilität, das Verhalten von H_c in der Nähe von φ^c eingeschränkt auf die Menge $I = \text{const}$ zu studieren. Für die Stabilität genügt es zu wissen:

Wenn $\langle I'[\varphi], \psi \rangle = 0$, dann $\langle H_c'' \psi, \psi \rangle \geq 0$.

Das bedeutet, die Hyperfläche $(I'[\varphi])^\perp$ „trifft" den Kegel $\{\psi : \langle H_c'' \psi, \psi \rangle < 0\}$ nicht. Gemäß Grillakis et al. [12] ist das – unter hier erfüllten Annahmen, insbesondere, dass Kern und negativer Eigenraum eindimensional sind – genau dann der Fall, wenn das sogenannte *moment of instability*

$$m(c) := H_c[\varphi^c] \quad \text{konvex ist, d. h., wenn} \quad m''(c) > 0.$$

Für die KdV-Solitonen lässt sich das in der Tat nachrechnen (das Soliton ist explizit gegeben). Modifiziert man die Nichtlinearität der KdV-Gleichung und betrachtet die *verallgemeinerte KdV-Gleichung*

$$\partial_t U = -\partial_x (U^p) - \partial_{xxx} U \tag{5.11}$$

mit $p \geq 2$, dann bleibt die Existenz von Solitonen bestehen, aber ihre Stabilität hängt von p ab. Durch Berechnung von $m(c)$ findet man, dass die Solitonen genau dann stabil sind, wenn $p < 5$ ist. Das wurde von Pego & Weinstein [13] gezeigt.

Interessanterweise hat man hier also die Untersuchung der Stabilität von Solitonen auf die Untersuchung einer *skalaren* Funktion $m(c)$ zurückführen können. Verallgemeinerungen auf Oberflächenwasserwellen oder interne Wasserwellen sind denkbar.

Literatur

[1] H. D. Alber: *Skript zur Vorlesung „Variationsrechnung und Sobolev-Räume"*, siehe: `http://www3.mathematik.tu-darmstadt.de/ags/analysis/ag-partielle-differentialgleichungen-und-anwendungen/lehrmaterial.html`.

[2] B. van Brunt: *The Calculus of Variations*, Springer, 2004.

[3] J. Ball: *Some open problems in elasticity* (Buchkapitel), Springer, 2002.

[4] B. Dacorogna: *Direct Methods in the Calculus of Variations*, Springer, 2004.

[5] L. C. Evans: *Partial Differential Equations*, AMS, 2002.

[6] M. Giaquinta: *Multiple Integrals in the Calculus of Variations and Nonlinear Elliptic Systems*, Princeton University Press, 1983.

[7] M. Giaquinta, S. Hildebrandt: *Calculus of Variations I*, Springer, 2004.

[8] H. Kielhöfer: *Variationsrechnung*, Vieweg+Teubner, 2010.

[9] M. Milnor: *Morse Theory*, Princeton University Press, 1963.

[10] M. Struwe: *Variational methods*, Springer, 1996.

[11] D. Werner: *Funktionalanalysis*, Springer, 2011.

[12] M. Grillakis, J. Shatah, W. Strauss: *Stability Theory of Solitary Waves in the Presence of Symmetry, I**, J. Funct. Analysis 74, 160-197 (1987)

[13] R. L. Pego, M. I. Weinstein: *Eigenvalues, and Instabilities of Solitary Waves*, Phil. Trans. R. Soc. Lond. A (1992) 340, 47-94.