1 Der „Hauptsatz“ der Differential- und Integralrechnung für das eindimensionale Lebesguessche Integral

Der sogenannte Hauptsatz der Differential- und Integralrechnung besagt, daß das eindimensionale Riemann-Integral $\int_a^x f\,dx$ bezüglich x stetig differenzierbar ist, und daß gilt

$$\frac{d}{dx}\left(\int_a^x f\,d\xi\right) = f(x)\,.$$

Ein ähnliches Resultat gilt, wenn f im Sinne von Lebesgue integrabel ist.

Satz 1.1 *Es sei $f \in L^1[a,b]$ und $F(x) := \int_a^x f\,d\xi$. Dann ist F fast überall in $[a,b]$ differenzierbar und es gilt $\frac{d}{dx}F = f$ fast überall.*

Anmerkung: F ist außerdem stetig (siehe Lemma 1.1).

Der Beweis dieses Satzes benötigt größere Vorbereitungen. Zur Motivation für die Wichtigkeit dieses Satzes erwähnen wir:
(i) In Anwendungen (z.B. Steuerungstheorie) betrachtet man häufig auf gewöhnliche Differentialgleichungen $\dot{x}(t) = f(t, x(t))$, bei denen f bezüglich t Sprünge haben darf, etwa also nur meßbar ist. Die Ableitung auf der linken Seite existiert dann nur fast überall.
(ii) Der Raum der stetig differenzierbaren Funktionen mit der Sobolev-Norm $\|\cdot\|_{1,2}$

$$\tilde{H}^1(a,b) = \left\{ u \in C[a,b] \cap C^1(a,b) \Big| \|u\|_{1,2} = \left(\int_a^b ((u')^2 + u^2)\,dx\right)^{1/2} < \infty \right\}$$

ist bezüglich der Norm $\|\cdot\|_{1,2}$ nicht vollständig. Um einen vollständigen Raum zu erhalten, muß der Ableitungsbegriff im Sinne von Satz 1.1 abgeschwächt werden.

Beweisstrategie für Satz 1.1:
(i) Sei $f \in L^1[a,b]$. Man erinnere sich, daß das Lebesgue-Integral absolutstetig ist, zu $\varepsilon > 0 \exists \delta > 0$, so daß für alle Mengen $E_\delta \subset [a,b]$ mit $\mu(E_\delta) < \delta$ gilt:

$$\Big|\int_{E_\delta} f\,d\xi\Big| < \varepsilon\,.$$

Ist $E_\delta = \bigcup\limits_{i=1}^{N} [t_i, t_i']$, so gilt

$$\sum_{i=1}^{N} (F(t_i') - F(t_i)) = \int\limits_{E_\delta} f \, d\xi \,.$$

Dies gilt in Analogie für f_+ und f_-. Da $t_i' > t_i$ gilt, falls $f \geq 0$

$$\sum_{i=1}^{N} |F(t_i) - F(t_i)| = \int\limits_{E_\delta} f \, d\xi < \varepsilon \text{ falls } \sum_{i=1}^{N} (t_i' - t_i) < \delta(\varepsilon)$$

und, da wir diese Uberlegung für f_+ und f_- durchführen können, letztlich

$$\sum_{i=1}^{N} |F(t_i') - F(t_i)| < \varepsilon \text{ falls } \sum_{i=1}^{N} (t_i' - t_i) < \varepsilon \,. \tag{1.1}$$

Die Eigenschaft (1.1) ist eine Verschärfung des Stetigkeitsbegriffes; man nennt F dann absolutstetig.

Definition 1.1 *Eine Funktion $g : [a, b] \to \mathbb{R}$ heißt absolutstetig $\Leftrightarrow$. Zu jeden $\varepsilon > 0$ gibt es ein $\delta > 0$, so daß für alle $N \in \mathbb{N}$ und alle $2N$-Tupel $(t_1, t_1', \dots t_N, t_N')$ mit $t_i < t_i' < t_{i+1} < t_{i+1}'$ und*

$$\sum_{i=1}^{N} (t_i' - t_i) < \delta$$

die Ungleichung

$$\sum_{i=1}^{N} |g(t_i') - g(t_i)| < \varepsilon$$

stattfindet. Man schreibt $g \in AC$ oder $g \in AC[a, b]$.

Unsere Überlegungen unter (i) ergeben somit

Lemma 1.1 *Sei $f \in L^1[a, b]$. Dann ist die durch $F(x) = \int\limits_a^x f \, d\xi$ definierte Funktion F absolutstetig.*

(ii) Es gilt nun der

Satz 1.2 *Sei $F \in AC[a, b] \Rightarrow F$ ist fast überall differenzierbar.*

Dies ergibt zunächst wenigstens den ersten Teil der Aussage von Satz 1.1. Der Nachweis von Satz 1.2 ist schwierig. Man überlegt sich, daß jede Funktion $F \in AC$ von beschränkter Variation ist.

Definition 1.2 *Eine Funktion $f : [a,b] \to$ ist von beschränkter Variation, wenn für alle $N \in \mathbb{N}$ und alle N-Tupel $a = x_1^N < x_2^N < \ldots < x_{N-1}^N < x_N^N = b$ die Summen*

$$\sum_{i=1}^{N-1} |f(x_{i+1}) - f(x_i)|$$

gleichmäßig beschränkt sind. Die Größe

$$\mathbf{V_a^b(f)} = \sup_{N\in\mathbb{N}} \left\{ \sum_{i=1}^{N-1} |f(x_{i+1}) - f(x_i)| \,\middle|\, a = x_1^N < \ldots < x_N^N = b \right\}$$

heißt Variation von f. Offensichtlich ist nicht jede stetige Funktion von endlicher Variation.

Man beweist nun nacheinander die folgenden Sätze:

Satz 1.3 *Jede Funktion von beschränkter Variation ist Differenz zweier monotoner Funktionen.*

Satz 1.4 *Jede monotone Funktion ist fast überall differenzierbar.*

Damit sind auch Funktionen aus AC fast überall differenzierbar, und der erste Teil von Satz 1.1 ist bewiesen.

Dem Beweis der Sätze 1.3 und 1.4 ist Kapitel 2 gewidmet.

Zum Nachweis der zweiten Aussage von Satz 1.1, nämlich daß $\frac{F(x+h)-F(x)}{h} \to f(x)$ für alle $x \in [a,b] - E$ mit $\mu(E) = 0$ ist die Konvergenz

$$\frac{1}{h}\int_x^{x+h} f\,dx \to f(x)\,, \qquad x \in [a,b] - E \text{ äquivalent} \tag{1.2}$$

Begnügt man sich, daß die Konvergenz (1.2) für eine Teilfolge ($h_k \to 0$) richtig ist, so reicht ein Approximationsargument aus: Die Faltung

$$w_h * f(x) = \int_{-\infty}^{\infty} w_h(x-t) f(t)\,dt\Big(= \frac{1}{h}\int_x^{x+h} f\,dx\Big) \text{ mit } w_h = \begin{cases} 1/h & \text{auf} \quad [-h,0] \\ 0 & \text{sonst} \end{cases}$$

hat die Eigenschaft, daß

$$w_h * f(x) = \frac{1}{h} \int\limits_x^{x+h} f \, d\xi \text{ und } w_h * f \to f \text{ in } L^1 .$$

Für eine Teilfolge h_h erhält man dann Aussage (1.2) - aber wir wollen diese für jede Nullfolge ($h_h > 0$) wissen. Da jedoch bereits bewiesen ist, daß die Folge $\frac{F(x+h)-F(x)}{h}$ überhaupt konvergiert für $h \to 0$, erhalten wir die gewünschte Konvergenz

$$\frac{F(x+h) - F(x)}{h} \to f(x) \text{ für jede Nullfolge } \{h_k\} .$$

2 Funktionen von endlicher Variation

Die Ausführungen in diesem Kapitel halten sich an das Buch von Natanson, „ Theorie der reellen Funktionen einer reellen Veränderlichen“ (mit - hoffentlich - didaktischen Verbesserungen). Wir wiederholen die Definition der Variation $V_a^b f$ einer Funktion $f : [a,b] \to \mathbb{R}$.

Definition 2.1 $V_a^b f = \sup \left\{ \sum\limits_{k=1}^{N} \left| f(x_k) - f(x_{k-1}) \right| \;\middle|\; a = x_0^N < x_1^N < \ldots < x_N^N = b\,;\, N \in \mathbb{N} \right\}$.
$V_a^b f < \infty \Leftrightarrow$ „f besitzt endliche Variation“.

Beispiel: Lipschitz-stetige und monotone Funktionen sind von endlicher Variation. Die Funktion $x \cos \frac{1}{x}$ ist von unbeschränkter Variation. Wir erwähnen, daß Summe, Produkt und Quotienten von Funktionen endlicher Variation wiederum von endlicher Variation sind. Der Raum der Funktionen von beschränkter Variation wird mit BV bezeichnet. Wir erwähnen die Identität

$$V_a^b f = V_a^c f + V_c^b f \qquad \text{für } c \in [a,b] \qquad \text{(Übungsaufgabe)}$$

und werden diese auch benutzen.

In n-Dimensionen ist die Definition

$$V_\Omega f = \sup \left\{ \int\limits_\Omega f \operatorname{div} \varphi \, dy | \varphi \in C^1(\bar{\Omega}) \,, \quad |\varphi| \le 1 \right\}$$

üblich. Hierbei ist $f \in L^1(\Omega)$. Im eindimensionalen ist diese Definition äquivalent zu der eingangs getroffenen.

Satz 2.1 *Sei $f : [a,b] \to \mathbb{R}$ stetig und von endlicher Variation. Ist f an der Stelle $x_0 \in [a,b]$ stetig, so ist auch die Funktion g, welche definiert ist durch*

$$g(x) = V_a^x f$$

an der Stelle x_0 stetig.

Beweis: Man zeigt, daß g in x_0 rechts- und linksstetig ist, aber wir beschränken uns auf den Nachweis der Rechtsstetigkeit. Sei $\varepsilon > 0$ vorgegeben. Nach Definition des Begriffs „sup“ gibt es Zahlen $x_1, \ldots, x_N$ mit $a = x_0 < x_1 < x_2 < \ldots x_N = b$, so daß

$$V := \sum_{k=1}^{N} |f(x_k) - f(x_{k-1})| > V_{x_0}^b(f) - \frac{\varepsilon}{2} \tag{2.1}$$

Sei $\xi \in]x_0, x_1[$ mit $|f(x_0) - f(\xi)| < \frac{\varepsilon}{2}$. Aus (2.1) folgt

$$|f(x_0) - f(\xi)| + |f(\xi) - f(x_1)| + \sum_{k=2}^{N} |f(x_k) - f(x_{k-1}) > |V_{x_0}^b f - \frac{\varepsilon}{2}\,.$$

Andererseits ist

$$|f(\xi) - f(x_1)| + \sum_{k=2}^{N} |f(x_k) - f(x_{k-1})| \leq V_\xi^b f\,.$$

Somit folgt

$$|f(x_0) - f(\xi)| > V_{x_0}^b f - V_\xi^b f - \frac{\varepsilon}{2}$$

und damit

$$V_{x_0}^b f - V_\xi^b f < \varepsilon\,,\ \text{da } f \text{ stetig ist}\ .$$

Aus der Beziehung $V_a^\xi f + V_\xi^b f = V_a^{x_0} f + V_{x_0}^b f$ folgt $V_a^\xi f - V_a^{x_0} f < \varepsilon$ für alle $\xi \in]x_0, x_1[$. Da $V_a^\xi f - V_a^{x_0} f > 0$, ist die Rechtsstetigkeit damit bewiesen. □

Satz 2.2 *Sei $f : [a, b] \to \mathbb{R}$ von beschränkter Variation. Dann ist f Differenz zweier monoton wachsender Funktionen. Ist f zusätzlich stetig, ist f Differenz zweier stetiger, monoton wachsender Funktionen.*

Beweis: Sei $g(x) = V_a^x f$. Aufgrund der Definition von V_a^x ist g monoton wachsend. Der Clou ist nun, daß die Funktion

$$h(x) = g(x) - f(x)$$

auch monoton in x wächst. Damit ist dann $f(x) = g(x) - h(x)$, und der Satz ist bewiesen. Es gilt für $a \leq x \leq y < b$

$$\begin{aligned} h(y) = g(y) - f(y) = V_a^y f - f(y) = V_x^y f + V_a^x f - f(y) &= g(x) + V_x^y f - f(y) \\ &= g(x) - f(x) + V_x^y f + (f(x) - f(y)) \end{aligned}$$

und daher

$$h(y) - h(x) = V_x^y f - (f(y) - f(x)) \geq 0\,.$$

Das letzte Ungleichheitszeichen folgt aus der Definition von $V_x^y f$. Die Behauptungen betreffs der Stetigkeit folgen, da h und g nach Satz 2.1 stetig sind. □

Eine interessante Anwendung von Funktionen mit beschränkter Variation ist die Banachsche Indikatrix.

Sei $f : [a,b] \to \mathbb{R}$ stetig, $m = \min\limits_{[a,b]} f$, $M = \max\limits_{[a,b]} f$. Sei $N(y)$ die Anzahl der Lösungen (evtl. $= \infty$) der Gleichung

$$f(x) = y\,.$$

N heißt die Banachsche Indikatrix. Man kann beweisen (s. Natanson)

Satz 2.3 *Die Banachsche Indikatrix ist meßbar, und es gilt*

$$\int\limits_m^M N(y)\,dy = V_a^b f\,.$$

Folgerung: Ist f stetig und von endlicher Variation, so hat das Bild der Menge der Werte, die von f unendlich oft angenommen werden, das Maß Null.

3 Differentiation monotoner Funktionen

In diesem Kapitel wollen wir den vergleichsweise schwierigen Satz beweisen, daß monotone Funktionen fast überall differenzierbar sind.

Satz 3.1 *Sei* $[a, b] \subset \mathbb{R}$ *und* $f : [a, b] \to \mathbb{R}$ *monoton. Dann existiert eine Menge* $E \in [a, b]$ *mit* $\mu(E) = 0$*, so daß* f *in allen Punkten von* $[a, b] - E$ *differenzierbar ist.*

Da eine monotone Funktion höchstens abzählbar viele Unstetigkeitsstellen hat, genügt es, den Satz für stetige monotone Funktionen zu beweisen. Ferner genügt es, strikt monoton wachsende Funktionen zu behandeln, da wir statt $f(x)$ die Funktion $f(x) + x$ behandeln können. Zur Vorbereitung wird der Begriff des Deriviertenwertes eingeführt.

Definition 3.1 *Ein Element* $\lambda \in \mathbb{R} \cup \{-\infty, \infty\}$ *heißt Deriviertenwert der Funktion* $f : [a, b] \to \mathbb{R}$ *an der Stelle* x*, wenn eine Nullfolge* (h_k)*,* $h_k \neq 0$ *existiert mit*

$$[f(x + h_k) - f(x)]/h_k \to \lambda \qquad k \to \infty\,.$$

Offensichtlich ist f genau dann differenzierbar an der Stelle x, wenn alle Deriviertenwerte an der Stelle x gleich sind. Es gibt immer einen Deriviertenwert, und wenn f monoton steigend ist, müssen alle Deriviertenwerte positiv sein. λ bezeichnet man intuitiv mit $Df(x)$. Entscheidend für den Beweis von Satz 3.1 ist folgendes Resultat.

Hilfssatz 3.1 *Sei* $f : [a, b] \to \mathbb{R}$ *strikt monoton wachsend. In jedem Punkt* $x \in S \subset [a, b]$ *existiere ein Deriviertenwert* $Df(x) \leq p$*. Dann gilt*

$$\mu^*(f(S)) \leq p \cdot \mu^*(S)\,.$$

μ^* *bedeutet hierbei das äußere Lebesgue-Maß.*

Beweis: Sei $G \supset S$, G offen mit $\mu(G) \leq \mu^*(S) + \varepsilon$ (siehe Definition des äußeren Maßes). Ist $x_0 \in S$, $p_0 > p$, so gilt für $p_0 > p$ und eine Nullfolge (h_k)

$$\frac{f(x_0 + h_k) - f(x_0)}{h_k} < p_0\,, \qquad k \geq k_0\,. \tag{3.1}$$

□

Wir betrachten die folgenden Intervalle im Werte- und Bildbereich.

$$d_k(x_0) := [x_0, x_0 + h_k]\,, \quad \Delta_k(x_0) := [f(x_0), f(x_0 + h_k)]\,.$$

Da f monoton wächst, gilt $f(d_k(x_0)) \subset \Delta_k(x_0)$. Wir schreiben die Ungleichung (3.1) in der Form

$$\mu(\Delta_k(x_0)) < p_0 \mu(d_k(x_0))\,. \tag{3.2}$$

Da $h_k \to 0$, gibt es unter den Strecken $\Delta_k(x_0)$ beliebig kleine. (Beachte, daß wir o.B.d.A. f als stetig vorausgesetzt haben.) Da $f(x_0) \in f(S)\,,\ f(x_0) \in \Delta_h(x_0)$, wird die Menge $f(S)$ im Sinne von Vitali überdeckt (s. folgenden Satz). Daher gibt es eine abzählbare disjunkte Teilfolge $\Delta_{k_i}(x_i)$ mit

$$\mu(f(S) - \bigcup_{i=1}^{\infty} \Delta_{k_i}(x_i)) = 0\,.$$

Ferner sind auch die $d_{k_i}(x_i)$ disjunkt, da f strikt monoton ist. Aus (3.2) folgt daher mit $x_0 = x_i; k = k_i$ durch Aufsummieren

$$\mu^*(f(S)) \le p_0 \mu\left(\bigcup_{i=1}^{\infty} d_{k_i}(x_i)\right) \le p_0 \mu(G) = p_0 \mu^*(S) + \varepsilon\,.$$

Durch Grenzübergang $p_0 \to p\,,\ \varepsilon \to 0$ ergibt sich die Behauptung. Ähnlich beweist man

Hilfssatz 3.2 *Sei $f : [a,b] \to \mathbb{R}$ stetig und strikt monoton wachsend und $S \subset [a,b]$. In jedem Punkt $x \in S$ existiere ein Deriviertenwert $Df(x) \ge q$. Dann gilt*

$$\mu^*(f(S)) \ge q\mu^*(S)\,.$$

Beweis von Satz 3.1:

Wir bezeichnen mit $E_{p,q} \subset [a,b]$ die Menge aller x, in denen es zwei Deriviertenwerte $D_1 f(x)\,,\ D_2 f(x)$ mit $D_1 f(x) < p < q < D_2 f(x)$ gibt. Es gilt dann:

$$\begin{aligned} \mu^*(f(E_{p,q})) &\le p\mu^*(E_{p,q}) \\ \mu^*(f(E_{p,q})) &\ge q\mu^*(E_{p,q}) \\ \Rightarrow (p-q)\mu^*(E_{p,q}) &\le 0 \quad \Rightarrow \mu^*(E_{p,q}) = 0\,. \end{aligned}$$

Wir lassen nun p und q alle rationalen Zahlen p_i, q_i durchlaufen. Da $\mu^*(E_{p_i,q_i}) = 0$, gilt auch $\mu^*\left(\bigcup\limits_{i=1}^{\infty} E_{p_i,q_i}\right) = 0$. Da man zwischen zwei reelle Zahlen eine rationale Zahl schachteln kann, gilt

$$\bigcup_{p,q\in\mathbb{R}} E_{p,q} = \bigcup_{i=1}^{\infty} E_{p_i,q_i}$$

und damit

$$\mu^*(\bigcup_{p,q\in\mathbb{R}} E_{p,q}) = 0\,.$$

Dies bedeutet aber, daß die Menge E aller x, für die es zwei Deriviertenwerte gibt, das Maß Null hat. □

Wir benötigen zur Vervollständigung des Beweises somit noch den Vitalischen Überdeckungssatz. Im Folgenden wird häufig mit Intervallen aus $\mathbb{R}$ gearbeitet. Unter einem Intervall verstehen wir Mengen der Form $[a,b], [a,b), (a,b]$ oder (a,b). Nur letzteres ist ein offenes Intervall.

Definition 3.2 *Sei $E \subset \mathbb{R}$ und $\mathfrak{F} = \{I_\iota\}$ eine (eventuell überabzählbare) Familie von Intervallen I_ι. Zu jedem $x \in E$ und jedem $\varepsilon > 0$ gebe es ein $I_\iota \in \mathfrak{F}$ mit $\mu(I_\iota) < \varepsilon$ und $x \in I_\iota$. Man sagt dann, daß E von $\mathfrak{F}$ im Sinne von Vitali überdeckt wird. Grob gesprochen wird E im Sinne von Vitali überdeckt, wenn jeder Punkt aus E in beliebig kleinen Intervallen enthalten ist.*

Bemerkung: Zum Beispiel bildet die Menge $\{(q-r, q+r); r \in \mathbb{Q}, q \in \mathbb{Q}, r > 0\}$ eine Vitali-Überdeckung von ganz $\mathbb{R}$.

Satz 3.2 (Überdeckungssatz von Vitali) *Es sei $E \subset \mathbb{R}$ eine beschränkte Menge. E werde von einer Familie $\mathfrak{F} = \{I_\iota\}$ von Intervallen I_ι im Sinne von Vitali überdeckt, wobei $\operatorname{int}(I_\iota) \neq \emptyset$ gelte. Dann gibt es abzählbar viele $I_j \in \mathfrak{F}$, so daß*
$\mu^*(E - \bigcup\limits_{j=1}^{\infty} I_j) = 0$ und $I_k \cap I_j = \emptyset$ für $j \neq k$.

Wir bereiten den Beweis des Vitalischen Überdeckungssatzes durch einige Lemmata vor.

Die Beweise in vielen Lehrbüchern sind schwer zu lesen. Wir folgen hier einer relativ eleganten Darstellung aus dem Buch „Measure Theory“ von Doob.

Lemma 3.1 *Sei $\mathfrak{F} = \{I_\iota\}$ eine nicht notwendigerweise abzählbare Familie von nicht notwendigerweise offenen Intervallen I_ι mit $\operatorname{int} I_\iota \neq \emptyset$. Dann ist die Menge*

$$B := \bigcup_{\mathfrak{F}} I_\iota - \bigcup_{\mathfrak{F}} \operatorname{int} I_\iota$$

abzählbar.

Beweis: Sei B' die Menge aller Punkte von B, die rechte Randpunkte von Intervallen $I_\iota \in \mathfrak{F}$ sind. Zu jedem Punkt P aus B' wähle man ein einziges Intervall $I_\iota(P) \in \mathfrak{F}$, welches P als rechten Randpunkt hat. Die Menge dieser $I_\iota(P)$ ist disjunkt. Da $\operatorname{int} I_\iota \neq \emptyset$, muß die Menge der $I_\iota(P), P \in B'$ abzählbar sein. Daher ist B' abzählbar. Entsprechend ist die Menge aller linken Randpunkte von Intervallen I_ι abzählbar. □

Anmerkung: Eine Familie disjunkter offener Mengen aus $\mathbb{R}$ ist abzählbar, da jede einen rationalen Punkt enthält.

Aufgrund des Lemmas 3.1 dürfen wir im Überdeckungssatz von Vitali annehmen, daß die I_ι offene Intervalle sind, da wir E um die abzählbare Menge B abändern dürfen.

Lemma 3.2 *Sei $\mathfrak{F} = \{I_\iota\}$ eine Familie offener Intervalle mit $\mu\Big(\bigcup_{\mathfrak{F}} I_\iota\Big) < \infty$. Dann existiert zu jedem $c < 1/2$ eine Folge $(I_i \in \mathfrak{F} | i \in \mathbb{N})$ mit $I_i \cap I_h = \emptyset, i \neq h$ und $\sum_{i=1}^{\infty} \mu(I_i) \geq c\mu\Big(\bigcup_{\mathfrak{F}} I_\iota\Big)$.*

Beweis: Zunächst gilt nach einem Satz von Lindelöf, daß es eine abzählbare Menge $\mathfrak{F}' \subset \mathfrak{F}$ gibt, so daß $\bigcup_{\mathfrak{F}'} I_\iota = \bigcup_{\mathfrak{F}} I_\iota$. (Eine interessante Übungsaufgabe, die man z.B. durch Ausschöpfen der offenen Menge $\bigcup_{\mathfrak{F}'} I_\iota$ durch eine Folge von Kompakta und dem Satz von Heine/Borel löst.) Wir dürfen daher o.B.d.A. $\mathfrak{F}$ als abzählbar annehmen und schreiben $I_i, i = 1, 2, \ldots$. Wir wählen n so groß, daß

$$\sum_{j=1}^{n} \mu(I_j) \geq 2c\mu\Big(\bigcup_{\mathfrak{F}} I_\iota\Big) .$$

Die in der Summe vorkommenden I_j bezeichnen wir mit J_j. Wir dürfen annehmen, daß die J_j mindestens einen Punkt s_j enthalten, der nicht in den anderen J_l enthalten ist. Wir ändern die Numerierung, so daß die s_j monoton steigen. Es gilt dann $J_i \subset (-\infty, s_j)$, $J_k \subset (s_j, \infty)$ für $i < j < k$. Die Intervalle mit geraden Indizes sind daher disjunkt, entsprechend die mit ungeradem Index. Wegen

$$\sum_{j=1}^{n} \mu(J_j) > 2c\mu\Big(\bigcup_{\mathfrak{F}} I_\iota\Big)$$

gilt entweder für die gerade bzw. ungerade Indexauswahl:

$$\sum_{\substack{j=1 \\ j \text{ (un-)gerade}}}^{n} \mu(J_j) > c\mu\Big(\bigcup_{\mathfrak{F}} I_\iota\Big).$$

Das Lemma ist damit bewiesen. □

Beweis des Überdeckungssatzes von Vitali:

Nach dem oben Gesagten dürfen wir die überdeckenden Intervalle I_ι als offen sowie abzählbar annehmen. Sei G offene Menge mit $\mu(G) \leq \frac{10}{9}\mu^*(E)$. Sei $\mathfrak{F}'$ die abzählbare Unterfamilie der I_ι, welche in G enthalten ist. Diese Unterfamilie $\mathfrak{F}'$ ist immer noch eine Überdeckung von E im Sinne von Vitali, da mit $x \in E \subset G$ wegen der Offenheit von G immer noch kleine Intervalle (s. Vor. im Satz von Vitali) in G liegen. Nach Lemma 3.2 gibt es eine disjunkte Folge $\{I_j\}$ von paarweise disjunkten Intervallen mit

$$\sum_{j=1}^{\infty} \mu(I_j) \geq c\mu\Big(\bigcup_{\mathfrak{F}'} I_j\Big).$$

Wir wählen etwa $c = \frac{1}{3} < \frac{1}{2}$ und beachten $\mu\Big(\bigcup_{\mathfrak{F}'} I_j\Big) \geq \mu^*(E)$. Ferner gilt $\mu\Big(\bigcup_{\mathfrak{F}'} \bar{I}_j\Big) = \mu\Big(\bigcup_{\mathfrak{F}'} I_j\Big)$. Wir bilden die Menge

$$H = G - \bigcup_{\mathfrak{F}'} I_j \text{ sowie } H_N = G - \bigcup_{\mathfrak{F}'} I_j\,.$$

H_N ist offen und für $N_{\to\infty}$ gilt $\mu\Big(\bigcup_{\mathfrak{F}'} \bar{I}_j\Big) = o(1) + \mu\Big(\bigcup_{\substack{\mathfrak{F}' \\ j=1,\dots N}} \bar{I}_j\Big)$ $j = 1, N$. Daher ist

$$\mu(H_N) = \mu(G) - \mu\Big(\bigcup_{\substack{\mathfrak{F}' \\ j=1,\dots N}} \bar{I}_j\Big) \leq \frac{10}{9}\mu^*(E) - \frac{1}{3}\mu^*(E) + o(1) = \frac{7}{9}\mu^*(E) + o(1)\,.$$

Daher ist auch $\mu^*\Big(E - \bigcup_{\mathfrak{F}'} I_j\Big) \leq \mu\Big(G - \bigcup_{\mathfrak{F}'} I_j\Big) \leq \frac{7}{9}\mu^*(E)$. Nunmehr wird dieselbe Konstruktion für die Menge $E - \bigcup_{\mathfrak{F}'} I_j$ durchgeführt. Die überdeckenden Intervalle sind $I_j \in \mathfrak{F} - \mathfrak{F}'$. Man erhält eine neue abzählbare Menge $\mathfrak{F}''$ von Intervallen I_j die paarweise disjunkt sind, so daß

$$\mu^*\Big(E - \bigcup_{\mathfrak{F}'} I_j - \Big(\bigcup_{\mathfrak{F}''} I_j\Big)\Big) \leq \Big(\tfrac{7}{9}\Big)^2 \mu^*(E)\,.$$

Die Konstruktion wird abzählbar oft durchgeführt, man erhält

$$\mu^*\Big(E - \bigcup_{l=1}^{N} \bigcup_{\mathfrak{F}^{(l)}} I_j\Big) \le \Big(\frac{7}{9}\Big)^N \mu^*(E) \to 0\,.$$

Da $\bigcup\limits_{l=1}^{\infty} \bigcup\limits_{\mathfrak{F}^{(l)}} I_j$ ebenfalls abzählbar ist, ist der Satz bewiesen.

4 Gewöhnliche Differentialgleichungen mit meßbaren Koeffizienten, Satz von Peano

In diesem Kapitel betrachten wir das Anfangswertproblem gewöhnlicher Differentialgleichungen. Für zahlreiche Anwendungen von gewöhnlichen Differentialgleichungen in der Steuerungstheorie (siehe das Beispiel am Schluß des Kapitels) ist es notwendig, nur meßbare Koeffizientenfunktionen zuzulassen. Gesucht ist eine Funktion $x \in AC([t_0, t_1[)$ (AC = Raum der absolutstetigen Funktionen) mit Werten in $\mathbb{R}^n$, so daß

(i) die Anfangsbedingung $x(t_0) = x_0$ bei vorgegebenen $x_0 \in \mathbb{R}^n$ sowie

(ii) die Differentialgleichung

$$\dot{x}(t) = f(t, x(t)) \quad \text{für alle} \quad t \in]t_0, t_1[-E$$

mit einer Ausnahmemenge $E \subset]t_0, t_1[$ vom Maß Null erfüllt ist. Wir schreiben kürzer

$$\begin{aligned} \dot{x} &= f(., x) \quad \text{fast überall in} \quad]t_0, t_1[\\ x(t_0) &= x_0 \end{aligned} \tag{4.1}$$

Hierbei ist $f : [t_0, t_1] \times \mathbb{R}^n \to \mathbb{R}^n$ oder $f : [t_0, t_1] \times U(x_0) \to \mathbb{R}^n$ eine Funktion, welche den sogenannten Carathéodory-Bedingungen genügt.

Definition 4.1 *Sei $I \subset \mathbb{R}$ ein Intervall und $M \subset \mathbb{R}^n$. Die Funktion $f : I \times M \to \mathbb{R}^n$ genügt den Carathéodory-Bedingungen, wenn gilt:*

(i) Für jedes festgehaltene $\xi \in M$ ist die Funktion $f(., \xi)$ meßbar („Meßbarkeit bezüglich t“).

(ii) Für jedes festgehaltene $t \in I - E$ ist die Funktion $f(t, .)$ stetig, mit einer Ausnahmemenge E vom Maß Null („Stetigkeit in ξ für fast alle ξ“).

Es erhebt sich zunächst die Frage, ob für absolutstetige Funktinen bzw. meßbare Funktionen die durch

$$g(t) := f(t, x(t))$$

definierte Funktion meßbar ist. Dies ist der Fall.

Satz 4.1 *Die Funktion $f : [t_0, t_1] \times \mathbb{R}^n \to \mathbb{R}^n$ genüge den Carathéodory-Bedingungen. Dann ist für jede meßbare Funktion $x : [t_0, t_1] \to \mathbb{R}^n$ auch die Funktion $f(., x(.))$ meßbar.*

Beweis: Aus der Maßtheorie wissen wir, daß es eine Folge von Treppenfunktionen $x_k : [t_0, t_1] \to \mathbb{R}^n$ gibt, die punktweise fast überall gegen x konvergiert. Die x_k sind konstant auf meßbaren Mengen $E_j^k, j = 1, 2, \dots k$, d.h. $x_k(t) = C_{jk}$ auf $E_j^k, E_j^k \cap E_l^k = \emptyset, j \neq l\,,\ \bigcup\limits_{j=1}^{k} E_j^k = [t_0, t_1]$. Wir zeigen, daß die Funktion $f(., x_k(.))$ meßbar ist. Hierzu genügt es zu zeigen, daß $f(., x_k(.))\big|_{E_j^k}$ meßbar ist, denn es gilt

$$f(., x_k) = \sum_{j=1}^{k} \chi(E_j^k) f(., x_k)\,,$$

und wir können dann benutzen, daß die Summe meßbarer Funktionen meßbar ist. Es gilt aber $f(., x_k(.))\big|_{E_j^k} = f(., C_{jk})$, und diese Funktion ist nach Voraussetzung meßbar. Es gilt nun $f(., x_k(.)) \to f(., x(.))$ punktweise fast überall, da f im zweiten Argument stetig ist und $x_k \to x$ punktweise fast überall. Nach einem Satz aus der Maßtheorie ist der punktweise Limes meßbarer Funktionen wieder meßbar. Damit ist der Satz bewiesen.
□

Nun zurück zum Anfangswertproblem. Dieses ist äquivalent zur Integralgleichung:

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))\, ds =: (\mathfrak{A}x)(t) \tag{4.2}$$

Wir interpretieren Gleichung (4.2) als Fixpunktgleichung der Form $x = \mathfrak{A}x$. Unsere Aufgabe besteht also im Nachweis der Existenz eines Fixpunktes. Damit wir das Integral für $x \in AC$ hinschreiben dürfen, benötigen wir noch die (vergleichsweise harmlose) Bedingung:
Zu $C > 0$ existiert ein positives K_C, so daß gilt:

$$|f(t,\xi)| \le K_C \quad \text{für alle} \quad t \in [t_0, t_1] - E\,, \quad \xi \in \mathbb{R}^n \text{ bzw. } U(x_0) \text{ mit } |\xi| \le C\,, \tag{4.3}$$

d.h. $f(t, \xi)$ bleibt beschränkt, wenn ξ in beschränkten Mengen variiert.

Da für $x \in AC$ die Funktion $f(., x)$ meßbar ist und wegen (4.3) der Integrand in (4.2) beschränkt ist, gilt aufgrund der Absolutstetigkeit des Lebesgue-Integrals folgende Aussage:

Lemma 4.1 *Die in (4.2) definierte Abbildung $\mathfrak{A}$ bildet $AC([t_0, t_1])$ in sich ab.*

Aus den Sätzen der Kapitel 1-3 folgt, daß man (4.2) fast überall differenzieren darf, und daß aus (4.2) die Beziehung (4.1) folgt, d.h. die Integralgleichung (4.2) und die Differentialgleichung (4.1) sind äquivalent.

Ist f nur auf $[t_0, t_1] \times U(x_0)$ definiert, so ist es von Wichtigkeit zu wissen, daß die Abbildung $\mathfrak{A}$ die Menge aller absolutstetigen Funktionen x mit Werten in $U(x_0)$ in sich selbst abbildet, sofern man das Intervall $[t_0, t_1]$ verkleinert:

Lemma 4.2 *Es existiert ein $\delta > 0$, welches von der Konstante K_C in (4.3) abhängt, so daß die in (4.2) definierte Abbildung $\mathfrak{A}$ die Menge aller $x \in AC[t_0, t_0 + \delta]$ mit Werten in $U(x_0)$ in sich selbst abbildet.*

Beweis:

$$|x_0 - (\mathfrak{A}x)(t)| = \Big|\int_{t_0}^{t} f(s, x(s))\, ds\Big| \leq (t - t_0) K_C \leq \varepsilon$$

falls $\delta = \frac{\varepsilon}{K_C}$ und $t \in [t_0, t_0 + \delta]$, d.h. $\mathfrak{A}(x)(t)$ liegt in $U(x_o)$ für alle t. □

Die Frage der Lösbarkeit des Anfangswertproblems (4.1) ist somit auf die Frage der Existenz eines Fixpunktes zurückgeführt.

$$x = \mathfrak{A}x\,, \qquad x \in \{y \in C([t_0, t_0 + \delta]) : \quad y(t) \in U(x_0) \quad \forall t\} =: \mathcal{C}_\delta \tag{4.4}$$

Ist x Lösung dieser Gleichung, folgt automatisch $x \in AC$. Wie im Fall von Anfangswertproblemen mit stetigem f läßt sich die Existenz eines eindeutigen Fixpunktes mit Hilfe des Satzes von der kontrahierenden Abbildung beweisen, wenn man eine lokale Lipschitz-Bedingung für $f(t, x)$ bezüglich des zweiten Argumentes fordert:

$$|f(t, \xi) - f(t, \eta)| \leq L_C |\xi - \eta| \quad \text{für alle } t \in [t_0, t_1]\,. \tag{4.5}$$

Anmerkung: Die Lipschitz-Bedingung impliziert die lokale Beschränktheit (4.3). Wählt man das Intervall $[t_0, t_0 + \delta]$ genügend klein, erhält man die Kontraktionsbedingung

$$\|\mathfrak{A}x - \mathfrak{A}y\|_\infty \leq q\|x - y\|_\infty\,, \qquad x, y \in C_\delta$$

mit einer Zahl $q < 1$.

Wie im klassischen Fall bei stetigem f erhält man, daß die „iterierten Funktionen" $\mathfrak{A}^m x_0$ konvergieren und der Limes dieser Folge löst (4.4). Man benötigt hierzu lediglich folgendes Resultat:

Lemma 4.3 *Die Carathéodory-Funktion f genüge der Lipschitz-Bedingung (4.5). Dann gilt: $\mathfrak{A}$ ist bezüglich der $\|\cdot\|_\infty$-Konvergenz stetig.*

Beweis: Es gelte $\|x_m - x\|_\infty \to 0$. Dann gilt wegen des Satzes von der majorisierten Konvergenz

$$\sup_t |\mathfrak{A}x_m(t) - \mathfrak{A}x(t)| = \sup_t \left| \int_{t_0}^{t} (f(s, x_m(s)) - f(s, x(s))\, ds \right|$$

$$\leq \int_{t_0}^{t_0+\delta} |f(s, x_m(s)) - f(s, x(s))|\, ds \to 0 \qquad (m \to \infty)$$

□

Wir halten das Ergebnis der vorangegangenen Diskussion in folgendem Satz fest:

Satz 4.2 *Sei $x_0 \in \mathbb{R}^n$ und $U(x_0) \subset \mathbb{R}^n$ eine ε-Umgebung von x_0. Die Funktion $f : [t_0, t_1] \times U(x_0) \to U(x_0)$ erfülle die Carathéodory-Bedingungen und die lokale Lipschitz-Bedingung (4.5). Dann existieren ein $\delta > 0$ und eine eindeutige Lösung $x \in AC([t_0, t_0 + \delta])$ des Anfangswertproblemes (4.1).*

Wegen der u.U. notwendigen Verkleinerung des Intervalles von $[t_0, t_1]$ auf $[t_0, t_0 + \delta]$ spricht man auch von einem „lokalen Existenzsatz“.

Verzichtet man auf die Eindeutigkeitsaussage, läßt sich die lokale Existenz ohne die Lipschitz-Bedingung beweisen, es genügen die Carathéodory-Bedingungen und die lokale Beschränktheit (4.3). Im Fall von stetigem f ist dies die Aussage des Satzes von Peano, aber auch unter den allgemeineren Bedingungen an f versteht man unter dem Schlagwort „Satz von Peano“ die entsprechende Aussage zur lokalen Existenz von Lösungen.

Satz 4.3 *(„Satz von Peano unter Carathéodory-Bedingungen“) Sei $x_0 \in \mathbb{R}^n$ und $U(x_0) \subset \mathbb{R}^n$ eine ε-Umgebung von x_0. Die Funktion $f : [t_0, t_1] \times U(x_0) \to U(x_0)$ genüge den Carathéodory-Bedingungen und der lokalen Beschränktheitsbedingung (4.3). Dann gibt es ein $\delta > 0$ und eine Lösung $x \in AC([t_0, t_0 + \delta])$ von (4.1).*

Eine bekannte Möglichkeit zum Nachweis des Satzes von Peano besteht in der Anwendung des Schauderschen Fixpunkt-Satzes. Zu dessen Formulierung erinnern wir an die Definitionen von Banach-Raum und kompakter Abbildung.

Definition 4.2 *Ein Banachraum ist ein normierter Vektorraum, der bezüglich der Normkonvergenz vollständig ist.*

Definition 4.3 *Eine Abbildung $\mathfrak{A}$ einer Teilmenge $M \subset B$ nach B heißt kompakt, wenn sie beschränkte Mengen in folgenkompakte Mengen überführt. Eine Menge $S \subset B$ heißt folgenkompakt, wenn jede Folge $\{x_k\}, x_k \in S$, eine in B konvergente Teilfolge besitzt, deren Limes in S liegt. Liegt der Limes nicht notwendig in S, sondern in B, so heißt die Menge S relativ folgenkompakt.*

Satz 4.4 *[Schauderscher Fixpunkt-Satz] Eine stetige, kompakte Abbildung einer konvexen Teilmenge eines Banachraumes in sich besitzt einen Fixpunkt.*

Der Beweis ist schwierig. Er läßt sich zwar mit analytischen Methoden auf das endlich-dimensionale Analogon, den Brouwerschen Fixpunkt-Satz, zurückführen, aber dessen Beweis benötigt Methoden aus der algebraischen Topologie. Es gibt Beweise des Brouwerschen Satzes mit Hilfe des Satzes von Stokes - siehe z.B. Dunford-Schwartz - man hat dann die Illusion, keine algebraische Topologie zu benötigen.

Die Voraussetzungen im Satz von Peano sind tatsächlich so, daß der Schaudersche Fixpunkt-Satz für die Abbildung $\mathfrak{A}$ in (4.2) anwendbar ist. Lemma 4.2 und 4.3 sichern einen Teil der Voraussetzungen; es bleibt zu zeigen:

Lemma 4.4 *$\mathfrak{A}$ bildet beschränkte Mengen in folgenkompakte Mengen ab.*

Beweis: Wir zeigen, daß $\mathfrak{A}$ beschränkte Folgen von Funktionen in gleichgradig stetige Funktionenfolgen abbildet. Aus dem Satz von Arzela-Ascoli (s. Infini I) folgt dann die Existenz einer konvergenten Teilfolge.

Nachweis der gleichgradigen Stetigkeit: Zu zeigen ist

$$\forall \varepsilon > 0 \; \exists \delta > 0 : \; \forall t\,,\; t' \quad \text{mit } |t - t'| < \delta \quad \text{und } \forall m \quad \text{gilt } \; |(\mathfrak{A}x_m)(t) - \mathfrak{A}x_m(t')| < \varepsilon\,.$$

Dies ist äquivalent zu

$$\Big| \int_{t'}^{t} f(s, x_m(s))\, ds \Big| < \varepsilon$$

und offensichtlich richtig, wenn x_m beschränkt und f der Bedingung (4.3) genügt. □

Wir bringen nun einen elementaren Beweis des Satzes von Peano, der ein wenig an die numerische Mathematik erinnert.

Beweis: Wenn $f(t,\xi)$ bezüglich t stetig wäre, könnten wir die Differentialgleichung durch Differenzengleichungen

$$\frac{x_h(t+h)-x_h(t)}{h} = f(t, x_h(t))\,, \qquad t = kh\,, \quad k \in \mathbb{N}$$

approximieren. Wenn man Pech hat, ist aber gerade $kh \in E$, der Ausnahmemenge. Deshalb muß im Fall der Carathéodory-Bedingungen anders vorgegangen werde. O.b.d.A. setzen wir im folgenden $t_0 = 0$. Wir lösen rekursiv für $t = 0, h, 2h, \dots, mh$

$$x_h(t) = x_0 + \int_{t_0}^{t} f(s, I_h^0 x_h(s))\, ds\,, \qquad x_h(0) = x_0\,, \tag{4.6}$$

wobei $I_h^0 x_h(s) := x_h(kh-h)$ für $s \in [kh-h, kh[$. Zur Berechnung von $I_h^0 x_h(s), s \in [t_0, t[$, $t = kh$ benötigt man nur die Werte von x_h an den Stellen $t = 0, h, \dots, (k-1)h$, d.h. über (4.6) läßt sich $x_h(kh)$ berechnen. Daher ist x_h für alle $kh \in [t_0, t_1[$ definiert.

Wir wollen nun den Grenzübergang $h \to 0$ rechtfertigen. Zunächst ergibt sich durch Rekursion - ähnlich wie beim Beweis von Lemma 4.2, daß ein $\delta > 0$ existiert, so daß $x_h(t) \in U(x_0), t = kh \in [0, \delta]$. Dies führen wir nicht aus.

Im nächsten Schritt ergibt sich die gleichmäßige Beschränktheit der Differenzenquotienten

$$\frac{x_h(t+h)-x_h(t)}{h} = \frac{1}{h}\int_{t}^{t+h} f(s, I_h^0 x(s))\, ds\,.$$

Die rechte Seite der letzten Gleichung ist offensichtlich beschränkt, da $f(s, I_h^0 x(s))$ auf Grund der Voraussetzung (4.3) und der Aussage des ersten Schrittes gleichmäßig beschränkt ist ($\int_t^{t+h} \omega\, dx \le \|\omega\|_\infty \cdot h$).

Im dritten Schritt setzen wir die Gitterfunktionen $x_h : \{kh | k \in \mathbb{N}\} \to \mathbb{R}^n$ zu stückweise linearen Funktionen $I_h^1 x_h$ fort. Stückweise lineare Funktionen sind absolutstetig, es gilt, fast überall,

$$(I_h^1 x_h)'(s) = \frac{x_h(t+h)-x_h(t)}{h}\,, \qquad s \in [t, t+h]\,, \quad t = hk\,.$$

Die Größen $\|(I_h^1 x_h)'\|_\infty$ sind daher gleichmäßig für $h \to 0$ beschränkt. Es gilt

$$I_h^1 x_h(t) = x_0 + \int_0^t (I_h^1 x_h)'\, ds\,,$$

und wir schließen wie in Lemma 4.4 die gleichgradige Stetigkeit der Funktionen $I_h^1 x_h$ im Intervall $[0,\delta]$. Es gibt dann nach dem Satz von Arzela-Ascoli eine gleichmäßige konvergente Teilfolge, welche gegen eine stetige Funktion $x : [0,\delta] \to U(x_0)$ konvergiert: $x_h \to x \quad \text{in } L^\infty$.

Wir zeigen im vierten Schritt, daß x eine Lösung der Integralgleichung

$$x(t) = x_0 + \int_0^t f(s,x(s))\,ds$$

ist. Hierzu benötigen wir, daß $\|I_h^0 x_h - x\|_\infty \to 0$ ebenfalls gilt. Es gilt für $s \in [(k-1)h, kh]$

$$\begin{aligned}
I_h^0 x_h(s) &= x_h((k-1)h) = x_h(kh) - h\,(I_h^1 x_h)'((k-1)h) \\
&= x(s) + x((k-1)h) - x(s) + x_h(kh) - x(kh) \\
&\quad + x(kh) - x((k-1)h) - h\,(I_h^1 x_h)'((k-1)h) \\
&= x(s) + o(1)\,.
\end{aligned}$$

Da x stetig ist, folgt

$$I_h^0 x_h(s) - x(s) \to 0 \qquad \text{gleichmäßig, .}$$

Die diskrete Gleichung (**??**) kann man in der Form $O_u^0 x_h(t) = x_0 + I_h^0(\int_0^t (s, I_h^0 x_h(s))\,ds|_{t=kh}$ schreiben. Die linke Seite konvergiert, wie gezeigt, gegen $x(t)$. Aus der gleichmäßigen Konvergenz $f(s, I_h^0 x(s)) \to f(s,x(s))$ schließt man $\int_0^t f(s,x(s))\,ds$, und aufgrund gleichgradiger Stetigkeit folgt $I_h^0 \int_0^t f(s, I_h^0 x_h(s))\,ds - \int_0^t f(s, I_u^0 x_h(s))\,ds \to 0$. Somit konvergiert also auch die rechte Seite gegen ihr Pendant in (**??**), und der Satz ist bewiesen. □

Im Anschluß geben wir noch ein typisches Beispiel eines Differentialgleichungssystems, in welchem mit absolutstetigen und nicht überall stetig differenzierbaren abhängigen Variablen $x = (x_1, x_2, x_3)$ gearbeitet wird. Es handelt sich um das Problem, eine (meteorologische) Rakete möglichst hoch zu schießen. Hierbei ist der Rückstoß durch eine Ungleichungsbeschränkung eingeschränkt und die Brennstoffmenge begrenzt. Die Problemformulierung lautet:

Gesucht ist $x = (x_1, x_2, x_3) \in AC$ (Zustandsvariablen) und $u \in L^\infty$ (Steuerungsvariable), so daß

$$\begin{aligned}
\dot{x}_1 &= x_2 && (x_1 \text{ Höhe }, x_2 \text{ Geschwindigkeit}), \\
\dot{x}_2 &= \frac{1}{x_3}(cu - D(x_2,x_1)) - g && (\text{Newtonsche Bewegungsgleichung})\ , \\
\dot{x}_3 &= u && (\text{Massenerhaltung}, x_3 = \text{Masse})\ .
\end{aligned}$$

Die Anfangs- und Endbedingungen lauten

$$x_1(0) = 0\,,\ x_2(0) = 0\,, \qquad \begin{aligned} x_3(0) = x_{30} = &\ \text{Anfangsmasse}\ , \\ x_3(T) = x_{31} = &\ \text{Endmasse}\ . \end{aligned}$$

$x_1(T)$ soll maximiert werden. Die Steuerungsvariable $u : [0, T] \to \mathbb{R}$ soll der Steuerungsbeschränkung

$$0 \leq u \leq u_{\max} \in \mathbb{R}$$

genügen. Gesucht sind die Zeit T und die Funktion $u \in L^\infty(0, T)$ sowie die Funktionen $x_1, \ldots, x_3$, so daß die angegebenen Restriktionen (Anfangs- und Endbedingungen, Differentialgleichungssystem, Steuerungsbeschränkung) erfüllt sind und $x_1(T)$ maximal ist.

Die Steuerung $u(t)$ gibt den Brennstoffverbrauch pro Zeiteinheit an. Hierdurch ändert sich die Masse x_3 der Rakete. In die Newtonsche Bewegungsgleichung „Beschleunigung = Kraft/Masse" geht die Rückstoßkraft $c \cdot u, c$ Konstante, die Erdbeschleunigung g (die hier nicht als von x_1 abhängig angegeben wurde) ein. Der Term $D(x_2, x_1)$ beschreibt (neudeutsch: „modelliert") die „Reibung" der Rakete an der Luft. D (=„Drag") ist eine durch physikalische Gesetze gegebene Funktion. Typisch für $D(x_2, x_1)$ ist quadratisches Verhalten in x_2, d.h. bei hohen Geschwindigkeiten wächst die „Luftreibung" vergleichsweise stark. Bezüglich x_1 verhält sich $D(x_2, x_1)$ wie $e^{-\lambda x_1}$, der Luftdruck ist mit der Höhe durch die barymetrische Höhenformel gekoppelt und proportional zur Reibung.

Typisch für Steuerungsprobleme ist nun, daß die optimalen Steuerungen Srpünge haben können, so daß das zugehörige Differentialgleichungssystem nicht im klassischen Sinne lösbar ist, sondern nur fast überall gilt. Im vorliegenden Fall verläuft die optimale Steuerung in realistischen Situationen derart, daß sie zunächst eine Zeitlang gleich $u_{\max}$ ist, um schnell zu beschleunigen und Gewicht zu verlieren. Nach einiger Zeit ist die Rakete so schnell, daß die Luftreibung großen Einfluß hat. Die optimale Steuerung wird dann stetig auf einen Wert aus $\in]0, u_{\max}[$ zurückgenommen. Ab einem Zeitpunkt T_0 ist dann der Brennstoff verbraucht, die Steuerung fällt sprunghaft nach Null. Die Rakete fliegt dann aufgrund des Eigenimpulses noch ein bißchen weiter, bis sie ihre maximale Höhe erreicht. (Dies kann man alles beweisen, s. etwa das Buch von Lee-Markus, „Foundations of Optimal Control".)

5 Globale Lösbarkeit von Anfangswertproblemen für gewöhnliche Differentialgleichungssysteme

Lösungen von Anfangswertproblemen

$$\dot{x} = f(.,x)\,, \qquad x(t_0) = x_0 \in \mathbb{R}^n \tag{5.1}$$

können in endlicher Zeit „explodieren"; der Satz von Peano sichert nur die lokale Lösbarkeit von (5.1), stetiges f vorausgesetzt. Ein bekanntes Beispiel für eine solche „Explosion" ist das Anfangswertproblem

$$\dot{x} = x^2\,, \qquad x(0) = 1 \in \mathbb{R}\,.$$

Die Lösung lautet

$$x(t) = \frac{1}{1-t}\,,$$

welche an der Stelle $t = 1$ eine Singularität hat. Als Grund für die Ausbildung einer Singularität ist das superlineare Wachstum der rechten Seite x^2. (Allerdings kann man nicht sagen, daß jede in x superlineare rechte Seite $f(.,x)$ zu einer Singularität führt.) Die Frage, ob das Anfangswertproblem (5.1) eine ganz auf R_+ (bzw. auf dem Definitionsbereich von f) definierte Lösung hat - man nennt dies „globale Lösbarkeit" -, ist für physikalisch-technische Anwendungen sehr wichtig. In der Theorie der Planetenbewegung ist es z.B. für die Bewohner der entsprechenden Planeten beruhigend zu wissen, ob die die Bewegung regierenden Anfangswertprobleme (Newtonsches Bewegungs- und Gravitationsgesetz) in endlicher Zeit zu einer Katastrophe führen. Bereits beim Drei-Körperproblem ist diese Frage bisher ungelöst (wobei die jüngste Literatur darauf hindeutet, daß es in endlicher Zeit zu einer Katastrophe kommt).

Wir wollen im folgenden hinreichende Kriterien für die globale Lösbarkeit bereitstellen. Der Einfachheit halber befassen wir uns nur mit dem Fall, daß die Funktion f auf ganz $\mathbb{R}_+ \times \mathbb{R}^n$ definiert ist und wir eine Lösung auf ganz $\mathbb{R}_+$ erhalten. Der Fall $f : I \times \mathbb{R}^n$ wäre ähnlich zu behandeln, man erhält dann höchstens eine Lösung $x : I \to \mathbb{R}^n$.

Die globale Lösbarkeit wird zumeist mit Hilfe des folgenden Beweisprinzips hergeleitet. Sei

$$f : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^n$$

und es gelten die beiden Bedingungen:

$$f \text{ erfülle die Carathéodory-Bedingung. 5.2} \tag{5.2}$$

$$f \text{ sei beschränkt auf beschränkten Mengen von } \mathbb{R}_+ \times \mathbb{R}^n \text{ (s. Bedingung (4.3))}\ . \tag{5.3}$$

Unter $AC([0,a[;\mathbb{R}^n)$, kürzer $AC([0,a[)$ verstehen wir die Menge der Funktionen $x:[0,a[\to\mathbb{R}^n$, welche lokal absolutstetig sind, d.h. für welche die Restriktionen $x|_{[0,c]}$ aus $AC([0,c])$ für $0<c<a$ sind. (Dadurch können wir Funktionen behandeln, die am rechten Intervallende unbeschränkt sind.)

Sei $J=[0,a[$ oder $J=[0,a]\subset\mathbb{R}$, $\quad a>0$ ein „Existenzintervall", d.h. es existiert eine Funktion $x:J\to\mathbb{R}^n$, $\quad x\in AC$, welche das Anfangswertproblem

$$\dot{x}=f(.,x) \text{ fast überall im Inneren von } J\,,\quad x(0)=x_0 \text{ lßt }. \tag{5.4}$$

Wegen des Satzes von Peano ist die Menge der Existenzintervalle nicht leer. Ist das Existenzintervall abgeschlossen, so löst man das Anfangswertproblem zu $\frac{d}{dt}\tilde{x}=f(.,\tilde{x})$ jenseits von a mit der Anfangsbedingung

$$\tilde{x}(a)=x(a)$$

und erhält nach dem Satz von Peano eine Lösung $\tilde{x}$ in $[a,a+\delta_a[$, $\quad\delta_a>0$. Anschließend setzt man aus $x|_{[0,a]}$ und $\tilde{x}:[a,a+\delta_a[\to\mathbb{R}^n$ eine neue Lösung

$$x:[0,a+\delta[\to\mathbb{R}^n\,,$$

zusammen, die also auf einem etwas größerem Existenzintervall erklärt ist. Die Absolutstetigkeit bleibt bei dieser Maßnahme erhalten, da $\tilde{x}(a)=x(a)$. Die Differentialgleichung braucht ohnehin nur fast überall erfüllt zu sein, so daß wir den Punkt a nicht berücksichtigen müssen. (Im klassischen Fall, bei dem x stetig differenzierbar ist, gilt die Differentialgleichung auch im Punkt a aufgrund stetiger Fortsetzung.)

Wir halten die eben gemachten Überlegungen fest:

Lemma 5.1 *(Fortsetzungsprinzip) Ist $x\in AC[J;\mathbb{R}^n]$ eine Lösung von (5.4) mit einem abgeschlossenen Existenzintervall J, so läßt sich x unter den Voraussetzungen (5.2) und (5.3) als Lösung von (5.4) auf ein größeres, nach rechts offenes Intervall $\tilde{J}=[0,a+\delta[$ fortsetzen.*

Gelingt es nun, aufgrund der speziellen Struktur von f zu zeigen, daß für jedes $b>0$ die Funktion $x:[0,b[\to\mathbb{R}$ zu einer Funktion aus $AC([0,b];\mathbb{R})$ fortgesetzt werden kann - also in den rechten Randpunkten b absolutstetig fortgesetzt werden kann -, so erhält man die globale Lösbarkeit. Das „maximale " Existenzintervall muß dazu gleichzeitig nach rechts offen und abgeschlossen sein, also mit $\mathbb{R}_+$ übereinstimmen.

Dieser Gedanke wird in dem Beweis von Lemma 5.2 präziser durchgeführt.

Die Möglichkeit der absolutstetigen Fortsetzung der Funktion x von $[0, b[$ nach $[0, b]$ läßt sich einfacher durch die Bedingung

$$x \text{ sei auf } [0, b[\text{ beschränkt} \tag{5.5}$$

ersetzen. Aus der Differentialgleichung und (5.3) folgt dann, daß $\dot{x}$ beschränkt ist und somit $x \in AC[0, b]$.

Lemma 5.2 *Sei $x : [0, \delta[\to \mathbb{R}^n$ eine Lösung von (5.4). Es mögen die Carathéodory-Bedingung und die Bedingung der lokalen Beschränktheit (5.2) gelten. Für jedes nach rechts offene Intervall $[0, a[$ sei eine mögliche zugehörige Lösung beschränkt. Dann läßt sich $x : [0, \delta[\to \mathbb{R}^n$ als Lösung von (5.4) auf ganz $\mathbb{R}_+$ fortsetzen. (Insbesondere existiert eine globale Lösung.)*

> **Beweis:** Sei $a^* = \sup\{a > 0 | x \in AC([0, \delta[)$ läßt sich zu einer Lösung von (5.4) aus $AC[0, a[$ fortsetzen $\}$. Wir werden zeigen, daß $a^* = \infty$ gilt. Angenommen, $a^* < \infty$, d.h. $a^* \in \mathbb{R}$. Dann läßt sich x zu einer Lösung von (5.4) auf $[0, a^*[$ fortsetzen. Dies ist einfach einzusehen, wenn lokale Eindeutigkeit vorliegt. Es gibt eine monoton steigende Folge mit a_i mit $a_i \to a^*$, $a_0 = \delta$, und eine Lösung $x^{(i)}$ auf den Intervallen $[a_{i-1}, a_i]$, welche $x^{(i)}(a_{i-1})) = \lim\limits_{c \to a_{i-1}} x^{(i-1)}(c)$ erfüllt. Durch Zusammensetzen dieser $x^{(i)}$ erhält man eine Lösung $x : [0, a^*[\to \mathbb{R}^n$. Falls lokale Eindeutigkeit nicht erfüllt ist, also der allgemeine Fall vorliegt, betrachtet man Lösungen $x^{(i)} : [0, a_i] \to \mathbb{R}^n$. Diese sind nach Voraussetzung beschränkt, somit wegen der Differentialgleichung auch ihre Ableitungen. Nach dem Satz von Arzela-Ascoli gibt es dann eine Teilfolge, so daß die $x^{(i)}$ gleichmäßig konvergieren. Man kann dann in den Integralgleichungen zur Grenze übergehen. Aufgrund der Voraussetzung, daß x nach Voraussetzung beschränkt ist, läßt sich x aber absolutstetig nach $[0, a^*]$ fortsetzen, und nach Lemma 5.1 dann aber wiederum als Lösung von (5.4) nach $[0, a^* + \delta_{a^*}[$. Die Zahl a^* war daher nicht maximal. Dies ist ein Widerspruch zur Annahme $a^* < \infty$.
> □

Nach Lemma 5.2 ist es für die globale Lösbarkeit somit entscheidend, ob geeignete Strukturbedingungen an f die Beschränktheit von Lösungen in jedem Intervall $[0, a[$ sichern. Eine solche Bedingung wird z.B. durch das folgende Lemma geliefert.

Lemma 5.3 *Die Funktion $f : [0, a[\times \mathbb{R}^n \to \mathbb{R}^n$ erfülle die Carathéodory-Bedingung und die lokale Beschränktheitsbedingung. Ferner gelte die Bedingung*

$$f(t, \xi) \cdot \xi \leq K|\xi|^2 + g(t) \qquad \forall \xi \in \mathbb{R}^n \textit{ und „fast überall“ } t \in [0, a[\tag{5.6}$$

mit einer Konstanten K und einer integrablen Funktion g. Dann ist jede Lösung $x \in AC[0, a[$ des Anfangswertproblems (5.3) in $[0, a[$ beschränkt.

Anmerkung: „Fast überall“ bedeutet hier „alle mit Ausnahme einer Menge vom Maß Null“.

Beispiele zu Bedingung (5.6):

(i) Linearer Fall:

$$f(t,\xi) = A(t)\xi + b(t)$$

mit einer L^∞-Matrix A und einer L^∞-Funktion b.

(ii)

$$\dot{x} = -x^3 + b(t), \quad b \in L^2$$

Beweis: von Lemma 5.3 Es gilt nach Voraussetzung

$$\dot{x} = f(.,x)\,, \qquad x(0) = x_0\,.$$

Wir multiplizieren skalar mit x und erhalten

$$\dot{x}x = \frac{1}{2}\frac{d}{dt}|x|^2 = f(.,x)\cdot x \le K|x|^2 + g(t)\,,$$

d.h.

$$\frac{d}{dt}|x|^2 \le 2K|x|^2 + 2g(t)\,.$$

Aus dem Gronwallschen Lemma (ein Dauerbrenner, ähnlich wie die Dreiecksungleichung) welches im Anschluß bewiesen wird, folgt dann

$$|x|^2(t) \le e^{2Kt}\Big(|x_0|^2 + 2\int\limits_0^t e^{-2Ks}g^{(s)}\,ds\Big) \le$$

$$\le e^{2Ka}\Big(|x_0|^2 + 2\int\limits_0^a e^{-2Ks}g^{(s)}\,ds\Big) =: C_0\,,$$

also die behauptete Abschätzung. □

Lemma 5.4 (Gronwallsches Lemma) *Sei $\varphi \in AC([t_0,a[), \varphi$ skalar, und es gelte*

$$\dot{\varphi} \le K\varphi + g(.) \qquad \textit{fast überall in } [t_0,a[\textit{ mit } K \in \mathbb{R} \textit{ und } g \in L^1[t_0,a[\,. \tag{5.7}$$

Dann gilt

$$\varphi(t) \le e^{Kt}[e^{-Kt_0}\varphi(t_0) + \int\limits_{t_0}^t e^{-Ks}g(s)\,ds] \tag{5.8}$$

Beweis: Aus (5.7) folgt

$$\frac{d}{dt}\Big(e^{-Kt}\varphi(t)\Big) \le e^{-Kt}g(t) \qquad \text{fast überall in } [t_0,a[\,.$$

Integration von t_0 bis t ergibt

$$e^{-Kt}\varphi(t) - e^{-Kt_0}\varphi(t_0) \le \int\limits_{t_0}^{t} e^{-Ks}g(s)\,ds$$

und damit die Behauptung □

Aus Lemma 5.2 und 5.3 schließen wir

Satz 5.1 *Die Funktion $f : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^n$ erfülle die Carathéodory-Bedingung und sei auf beschränkten Mengen in $\mathbb{R}_+ \times \mathbb{R}^n$ beschränkt. Ferner gelte die Bedingung*

$$f(t,\xi)\cdot\xi \le K|\xi|^2 + g(t)$$

für alle $\xi \in \mathbb{R}^n$ und „fast alle" t, mit einer Konstanten K und einer Funktion g mit $g|_{[0,a]} \in L^1$ für alle $a > 0$. Dann besitzt das Anfangswertproblem

$$\dot{x} = f(.,x) \qquad \text{fast überall in } \mathbb{R}_+\,, \qquad x(0) = x_0 \in \mathbb{R}^n$$

eine globale Lösung $x \in AC([0,\infty[)$. Jede lokale Lösung dieses Anfangswertproblems läßt sich zu einer globalen Lösung fortsetzen.

Eine andere Methode, die Beschränktheit von Lösungen auf Intervallen $[0,a[$ zu sichern, beruht auf sogenannten Energieabschätzungen. Der Name rührt von entsprechenden Anwendungen aus der Physik her. Man betrachtet hier Differentialgleichungssysteme zweiter Ordnung

$$\ddot{x} + f(.,x) = 0\,, \qquad x(0) = x_0\,, \quad \dot{x}(0) = x_1\,, \qquad x_0, x_1 \in \mathbb{R}^n \text{ vorgegeben }. \tag{5.9}$$

f genüge wieder der Carathéodory-Bedingung, der Bedingung der lokalen Beschränktheit, man sucht $x \in C^1$, $\dot{x} \in AC([0,\infty[)$. Wir stellen die Zusatzvoraussetzung, daß f ein „Potential" hat, welches höchstens quadratisch nach $-\infty$ gehen darf, d.h.

$$f(.,x) = \nabla_x F(.,x)\,, \qquad F : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R} \tag{5.10}$$

(beachte: f ist $\mathbb{R}^n$-wertig, F skalar),
und es soll gelten $F(t,\xi) \ge -K|\xi|^2 - g(t)$ mit einer Konstanten K und einer auf beschränkten

Mengen integrablen Funktion g. Weiterhin sei die partielle Ableitung F_t eine Carathéodory-Funktion und erfülle auch $F_t(t,x) \leq \tilde{K} \cdot |x|^2 + \tilde{g}(t)$.

Beispiel Gravitationsgesetz:

$$f(t,\xi) = -\frac{C \cdot \xi}{|\xi|^3}, \qquad \xi \in \mathbb{R}^3, \quad \xi \neq 0$$

(Das Beispiel ist nicht ganz stubenrein wegen der Singularität bei $\xi = 0$.)

Lemma 5.5 *Sei $x \in AC([0,a[)$ eine Lösung von (5.9), und es gelte (5.10) sowie die Carathéodory- und die lokale Beschränktheitsbedingung. Dann ist x auf $[0,a]$ beschränkt.*

Beweis: Der Beweis wurde im Wesentlichen in einer Übungsaufgabe behandelt. Man multipliziert (5.8) mit $\dot{x}$ und erhält die Gleichung

$$\frac{d}{dt}\left(\frac{|\dot{x}|^2}{2}\right) + \frac{d}{dt}F(.,x) - F_t(.,x) = 0\,,$$

und durch Integration

$$\frac{|\dot{x}|^2}{2}\Big|_0^t + F(.,x)\Big|_0^t - \int_0^t F_t(.,x)\,ds = 0\,,$$

und unter Verwendung von (5.9), (5.10)

$$\frac{|\dot{x}|^2}{2} - K|x|^2 - g(.) = K\int_0^t |x|^2\,ds + \int_0^t g(s)\,ds + \frac{|\dot{x}|^2(0)}{2} + F(0,x_0)\,,$$

woraus man durch „Gronwall-ähnliche“ Überlegungen zu der behaupteten Aussage kommt. □

Folgerung:

Satz 5.2 *Die Voraussetzung von Lemma 5.4 gelte für jedes Intervall $[0,a[$, $a \in \mathbb{R}$. Dann gibt es eine globale Lösung $x : \mathbb{R}_+ \to \mathbb{R}^n$ von (5.8).*

6 Variationsrechnung und Randwertprobleme gewöhnlicher Differentialgleichungen

Die Grundaufgabe der („eindimensionalen“) Variationsrechnung lautet: Minimiere das Variationsintegral

$$J(u) = \int_a^b F(t, u(t), u'(t))\, dt$$

unter allen Funktionen $u : [a,b] \to \mathbb{R}^n$, welche Randbedingungen, z. B. $u(a) = u_a$, $u(b) = u_b$, erfüllen und für die $J(u)$ definiert ist. Hierbei ist $F : [a,b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ eine vorgegebene Funktion. Wir setzen voraus

$$F(t,\mu,\eta) \text{ sei stetig in } t \text{ und meßbar in } \mu \text{ und } \eta \text{ (Carathéodory-Bedingungen).} \tag{6.1}$$

Die sybellinische Forderung, daß $J(u)$ definiert sein soll, präzisieren wir dahingehend, daß u absolutstetig sein soll und u' somit fast überall definiert ist. Es ist besser, von Anfang an mit dem Grundraum $AC([a,b])$ anstelle von C^1 zu arbeiten, da für die Existenz von Minima der Raum C^1 nicht die geeigneten Vollständigkeitseigenschaften hat. Er ist nur bezüglich der C^1-Norm vollständig - dies ist zu einschränkend. Damit $J(u)$ für $u \in AC$ definiert ist, benötigt man

$$F \geq -K \text{ mit einer Konstanten } K \in \mathbb{R}\,, \tag{6.2}$$

muß aber dann eventuell für $J(u)$ den Wert $+\infty$ zulassen (was überhaupt nicht schlimm ist). Alternativ können wir mit dem Grundraum

$$H^1 = H^1([a,b]) = H^{1,2}([a,b])$$

oder allgemeiner mit $H^{1,p}$, $1 < p < \infty$, arbeiten. Hierbei ist

$$H^{1,p} = \{u \in AC([a,b]) | u' \in L^p([a,b])\}\,.$$

Die Norm in $H^{1,p}$ ist definiert durch

$$\|u\|_{1,p} := \|u'\|_p + \|u\|_p\,, \qquad \|w\|_p := \Big(\int_a^b |w|^p\, dx\Big)^{1/p}.$$

Damit $J(u)$ für $u \in H^{1,p}$ einen endlichen Wert hat, ist hinreichend, daß für alle $\eta \in \mathbb{R}^n$ und alle $\mu \in \mathbb{R}^n$ mit $|\mu| \leq C$ gilt:

$$F(t,\mu,\eta) \leq K_C|\eta|^p + K_C \qquad \forall t \tag{6.3}$$

mit einer Konstanten K_C. Zum Beweis beachte man, daß in einer Raumdimension $\|u\|_\infty \leq K\|u\|_{1,p}$, gilt. $H^{1,p}$-Funktionen sind sogar hölderstetig um Exponenten $(p-1/p)$, wie man an folgender Abschätzung erkennt:

$$|u(t') - u(t)| = |\int_t^{t'} u'\,dt| \leq (\int_t^{t'} 1\,dt)^{(p-1)/p} (\int_t^{t'} |u'|^p\,dt)^{1/p} \leq |t-t'|^{(p-1)/p} \|u'\|_p \,.$$

Ein klassisches Beispiel, welches die Variationsrechnung für die Neuzeit populär machte, ist das folgende:

Das Bernoullische Problem:

In einer vertikalen Ebene seien zwei Punkte $A = (a, u_a)$, $B = (b, u_b) \in \mathbb{R}^2$ gegeben, $u_a > u_b$. Das Problem besteht nun darin, diejenige Kurve zwischen A und B zu bestimmen, entlang derer sich ein Massepunkt mit Masse m vermöge der Schwerkraft in möglichst kurzer Zeit von A nach B bewegt. Es handelt sich sozusagen um das Problem einer „besten Rutschbahn“. Die Lösung ist keine Gerade!

Wir überlegen uns, wie das Variationsintegral aussieht. Wir stellen das Variationsintegral für C^1-Bahnen $u : [a,b] \to \mathbb{R}$, $u(a) = u_a$, $u(b) = u_b$. (An und für sich ist es beweisdürftig, daß das Minimum in der Klasse der Kurven, welche Funktionsgraphen sind, gesucht wird.)

Sei $v(t)$ die Geschwindigkeit des Massepunktes zur Zeit t und $h(t)$ die Höhe von der x-Achse aus gemessen. Nach dem Energieerhaltungssatz ist die Summe aus kinetischer und potentieller Energie konstant:

$$\frac{1}{2}mv^2(t) + mgh(t) = \frac{1}{2}mv^2(0) + mgu_a = c = \text{const.}\,.$$

Es folgt

$$v(t) = \frac{ds}{dt}(t) = \sqrt{\tfrac{2c}{m} - 2gh(t)}\,.$$

$s(t) =$ Länge der Bahnkurve zwischen der Anfangszeit 0 und t.
Für die Fallzeit T, die der Massepunkt auf der Kurve von A nach B braucht, gilt

$$T = \int_0^{t_B} dt = \int_{s(0)}^{s(t_B)} \frac{d\xi}{v(t)}\,.$$

Hier wurde die Substitutionsregel verwendet (sofern s umkehrbar ist).

$$t = s^{-1}(\xi)\,,\ \Rightarrow \xi = s(t) \Rightarrow d\xi = v(t)\,dt \Rightarrow dt = d\xi/v(t)\,.$$

Zur Zeit t befindet sich der Massepunkt an der Stelle $(x, y(x))$, d.h. es gilt $T = \int\limits_{s(0)}^{s(t_B)} \frac{ds}{\sqrt{\frac{2c}{m} - 2gy(x)}}$, wobei x durch $s(t)$ ausgedrückt ist. Durch die Substitution $s(t) \to x$ ergibt sich

$$ds = \sqrt{1 + (y')^2}\, dx$$

$$T = \int_a^b \frac{\sqrt{1 + (y')^2}}{\sqrt{\frac{2c}{m} - 2gy(x)}}\, dx\,.$$

Aufgrund des Energieprinzips ist $\frac{2c}{m} - 2gy(x) > 0$. T ist nun das gesuchte Funktional, welches es unter allen C^1-Bahnen y zu minimieren gilt.

Die Eulersche Differentialgleichung

Minima von Variationsproblemen erfüllen unter sehr allgemeinen Voraussetzungen die Eulersche Differentialgleichung. In der schwachen Formulierung lautet sie:

$$\sum_{i=1}^{n} \int_a^b \big(F_i(t, u(t), u'(t)) \cdot \varphi_i'(t) + F_{0i}(t, u(t), u'(t))\varphi_i(t)\big)\, dt = 0 \tag{6.4}$$

für alle Funktionen $\varphi = (\varphi_1, \ldots \varphi_n)$, die genügend glatt sind, z.B. $\varphi \in C^1$ mit Nullrandbedingungen erfüllen. (Je nach Art der Randbedingung kann man auf die Forderung, daß φ Nullrandbedingungen erfüllt, verzichten.) Es bedeuten F_i bzw. F_{0i} partielle Ableitungen von F nach dem Argument, wo u_i' bzw. u_i steht. Dies wird im Folgenden noch erläutert.

Unter gewissen Bedingungen läßt sich schließen, daß $F_i(t, u(t), u'(t))$ für die Lösung u der Gleichung (6.4) differenzierbar ist und folgendes Randwertproblem zweiter Ordnung erfüllt:

$$-\frac{d}{dt} F_i(., u, u') + F_{0i}(., u, u') = 0 \text{ a.e. in } [a, b]\,, \quad i = 1, \ldots n\,. \tag{6.5}$$

Wir geben präzise Bedingungen an, wann (6.4) und (6.5) erfüllt sind.

Um möglichst viele Arten von Randbedingungen behandeln zu können, betrachten wir lineare Untermannigfaltigkeiten $V + g \in H^{1,p}([a, b])$, $g \in H^{1,p}$. Man stelle sich z.B. den Fall von inhomogenen Dirichlet-Randbedingungen vor: $u(a) = u_a$, $u(b) = u_b$, und unter V den Teilraum von $H^{1,p}([a, b])$, der aus allen Funktionen v mit $v(a) = v(b) = 0$ besteht. Die Funktion g soll dann eine Funktion aus $H^{1,p}$ mit $g(a) = u_a$, $g(b) = u_b$ sein.
Das Variationsproblem lautet dann: Gesucht ist $u \in V + g$, so daß $J(u)$ minimal ist.

Ein anderes Beispiel für V wäre

$$V = \{v \in H^{1,p}([a, b]) | v(a) = 0\}\,,$$

d.h. es wird keine Bedingung am rechten Intervall gestellt. In jedem Fall wird vorausgesetzt, daß gilt:

$$H^{1,p}([a,b]) \supset V \supset C_0^\infty([a,b])\,. \tag{6.6}$$

Hierbei ist $C_0^\infty([a,b])$ wie üblich die Menge der C^∞-Vektorfunktionen mit kompakten Trägern in $[a,b]$.

Wir behandeln nunmehr allgemein: Gesucht ist $u \in W := V + g$, so daß

$$J(u) \le J(v) \qquad \text{für alle } v \in W\,. \tag{6.7}$$

Wir stellen nun Bedingungen an F_i und F_{0i}, damit die Integrale in (6.4) existieren. Wenn $u' \in L^p$, so gilt $u \in AC$, und wir benötigen die Bedingung

$$|F_i(t,\mu,\eta)| + |F_{0i}(t,\mu,\eta)| \le K_C|\eta|^p + K_C\,,\ \text{für } |\mu| \le C\,,\ \forall \eta \in \mathbb{R}^n \tag{6.8}$$

mit einer Konstanten K_C. Die Bedingung (6.6) sichert, daß $F_{i0}(t,u,u')$ und $F_i(.,u,u')$ in L^1 liegen. In diesem Fall ist der Ausdruck in (6.4) definiert, sofern $\varphi \in C^1([a,b])$.

Satz 6.1 *Es sei $u \in W$ eine Minimalstelle des Variationsintegrals*

$$J(u) = \int_a^b F(t,u,u')\,dt\,,$$

und es gelte (6.6) für V sowie (6.1) bzw. (6.2) oder (6.3) für F mit dem Wachstumsexponenten $p \in [0,\infty[$. Die partiellen Ableitungen F_{i0} und F_i bezüglich des Arguments u_i bzw. u_i', $u = (u_1,\dots u_n)$ mögen für fast alle t existieren, und es gelte die Wachstumsbedingung (6.8) für F_{i0} und F_i. Dann erfüllt u die Eulersche Differentialgleichung (6.4) für alle $\varphi \in V \cap C^1([a,b])$.

Beweis: Es sei $\xi \in V$ und somit $u + s\xi$ für reelles s eine konkurrierende Funktion aus W zur Minimierung von J. Wir betrachten die Funktion $w : \mathbb{R} \to \mathbb{R}$, die definiert ist durch

$$w(s) = J(u + sv)\,, \qquad s \in \mathbb{R}\,.$$

Aufgrund der Minimalität von u gilt:

$$w(0) \le w(s)\,.$$

Falls wir zeigen können, daß w differenzierbar ist, folgt dann $w'(0) = 0$. Es zeigt sich, daß $w'(0)$ gerade der Ausdruck in (6.4) ist. Der Satz ist also bewiesen, wenn wir das folgende Lemma zeigen: $\square$

Lemma 6.1 *Es mögen die Voraussetzungen an F aus Satz (6.1) gelten. Es sei $u \in H^{1,p}([a,b])$ und $\varphi \in V \cap C^1([a,b])$. Dann ist*

$$\frac{d}{ds} J(u+sv)|_{s=0} = \sum_{i=1}^{n} \int_a^b \{F_i(.,u,u')\varphi_i' + F_{i0}(.,u,u')\varphi_i\}\,dt\,. \tag{6.9}$$

Beweis: Wir müssen zeigen, daß die Differenzenquotienten $\frac{1}{h}(J(u+h\varphi) - J(u))$ gegen den Ausdruck (6.9) konvergieren für $h \to 0\,,\ h \neq 0$. Hierzu benötigen wir

$$\int_a^b h^{-1}(F_i(.,u+h\varphi,u'+h\varphi) - F(.,u,u'))\,dt \to \sum_{i=1}^{n} \int_a^b F_i(.,u,u')\varphi_i'\,dt$$
$$+ \sum_{i=1}^{n} \int_a^b F_{i0}(.,u,u')\varphi_i\,dt\,.$$

Wegen der vorausgesetzten partiellen Differenzierbarkeit von $F(t,\mu,\eta)$ bezüglich μ und η gilt die punktweise Konvergenz der Integranden gegen den entsprechenden Limes. Um den Satz von der majorisierten Konvergenz anzuwenden, benötigen wir noch, daß die Integranden gleichmäßig durch eine integrable feste Funktion abgeschätzt werden können. Hierzu stellen wir den Differenzenquotient

$$h^{-1}[F(.,u+h\varphi,u'+h\varphi') - F(.,u,u')] = D_h$$

mit Hilfe des Mittelwertsatzes dar:

$$D_h = \sum_{i=1}^{n} [F_i(.,u+\theta\varphi,u'+\theta\varphi')\varphi_i' + F(.,u+\theta\varphi,u'+\theta\varphi')\varphi_i]$$

mit einer Funktion $\theta = \theta(h,t)\,,\ 0 < \theta < 1$. Wir beachten, daß $|\varphi_i'| + |\varphi_i| \leq K$ und, wegen der Wachstumsbedingungen für F_i und F_{0i} gilt:

$$|F_i(.,i+\theta\varphi\,,\ u'+\theta\varphi')| \leq K_C|u'+\theta\varphi'|^p + K_C \leq 2^{p-1}K_C|u'|^p + \tilde{K}_C$$

analog für F_{i0}. Da $|u'|^p \in L^1$, ist D_h damit durch eine feste L^1-Funktion abgeschätzt. □

Existenz von Minima von Variationsproblemen

Ein Variationsproblem muß nicht notwendigerweise eine Lösung haben. Beispiel:

$$J(u) = \int_0^1 \{(1-(u')^2)^2 + u^2\}\,dt = \min\,,\ u(1) = u(0) = 0\,,\ u \in H^{1,4}([0,1])\,.$$

Man kann zeigen, daß in diesem Beispiel

$$\inf_u J(u) = 0 \,.$$

Es gibt nämlich eine Folge $u_m \in H^{1,\infty}$, $u_m(1) = u_m(0) = 0$ mit der Eigenschaft $J(u_m) \to 0$. Ohnehin gilt $J \geq 0$. Wir geben nur die Graphen der u_m an:

Die u_m seien periodische Hütchenfunktionen der Periode 2^{-m+1} von der Form

$$u_m(t) := \begin{cases} t & ; \quad t \in [0, 2^{-m}) \\ 2^{-m+1} - t & ; \quad t \in [2^{-m}, 2^{-m+1} \end{cases}$$
$$u_m(t + 2^{-m+1}) := u(t) \qquad \forall x \in [0,1]$$

Offensichtlich gilt $u_m^2 \to 0$ gleichmäßig und $(1 - (u_m')^2) = 0$ a.e., woraus $J(u_m) \to 0$ folgt. Andererseits kann das Infinum $J(u) = 0$ nicht angenommen werden. Aus $J(u) = 0$ folgt $u = 0$ und $(1 - (u')^2)^2 = 0$ und somit $O = 1$. Falls der Dozent bei der Einführung der reellen Zahlen das Axiom $0 \neq 1$ nicht vergessen hat, folgt damit ein Widerspruch. Im vorigen Jahrhundert wurde die Annahme des Minimums von Variationsproblemen von bedeutenden Mathematikern wie Riemann fehlerhafterweise als nicht beweisbedürftig angenommen. Das Bemühen der Mathematiker, die Annahme des Minimums zu beweisen, hat zu einer bedeutenden Weiterentwicklung der mathematischen Methoden geführt. Historisch gesehen wurde die Lebesguesche Maßtheorie zu dem Zwecke entwickelt, die Existenz von Minima von Variationsproblemen zu zeigen. Hiermit wollen wir uns im folgenden beschäftigen.

Für die Existenz von Minima benötigen wir zwei wichtige Voraussetzungen, nämlich:

$$J(u) = \int_a^b F(.,u,u')\,dx \to \infty \qquad \text{falls } \|u\|_{1,p} \to \infty\,,\ u \in W\,. \tag{6.10}$$

Man nennt diese Eigenschaft Koerzivität (coerciveness). Eine analoge Voraussetzung benötigt man schon bei der Minimierung reeller Funktionen, z.B. hat $e^{-\xi}$ auf $\mathbb{R}$ kein Minimum, da $e^{-\xi}$ nicht koerzitiv ist.

Die Bedingung (6.10) hat zur Folge, daß Minimalfolgen von $J(u)$, $u \in W$, in $H^{1,p}$ beschränkt ist. Für beschränkte Folgen in $H^{1,p}$ mit $p = 2$ gibt es einen „Ersatz" für den Satz von Bolzano-Weierstraß.

Satz 6.2 (Satz von Banach-Saks) *Sei (u_m) eine beschränkte Folge in einem Hilbertschen Raum H. Dann gibt es eine Teilfolge (u_{m_i}), so daß die arithmetischen Mittel*

$$\frac{1}{N}\sum_{i=1}^{N} u_{m_i}$$

gegen ein Element u aus H konvergieren.

Dies ist ein Satz aus der Funktionalanalysis, der übrigens nicht sonderlich schwer zu beweisen ist. (Für $1 < p < \infty$, also nicht notwendig $p = 2$, gibt es ein ähnliches Resultat, auf das wir hier nicht eingehen.) Ferner benötigen wir ein weiteres Substitut für den Satz von Bolzano-Weierstraß, nämlich:

Satz 6.3 (Satz von Rellich in einer Dimension) . *Sei (u_m) eine beschränkte Folge in $H^{1,p}$ mit einem $p \in]1, \infty[$. Dann gibt es eine gleichmäßig konvergente Teilfolge.*

(In mehreren Dimensionen kann man nur sagen: Es gibt eine in L^p-konvergente Teilfolge.)

Mit diesen Hilfsmitteln läßt sich der folgende Satz beweisen:

Satz 6.4 *Die Funktion $F : [a, b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ sei nach unten beschränkt und genüge den Carathéodory-Bedingungen. Ferner sei $F(t, \mu, \eta)$ bezüglich η konvex und das Variationsintegral $J(u) = \int_a^b F(t, u, u')\,dt$ sei koerzitiv bezüglich der $H^{1,p}$-Norm auf der abgeschlossenen Untermannigfaltigkeit $W = g + V \subset H^{1,p}$. Dann gibt es ein Minimum von J auf W.*

Anmerkung: Neben den Standardbedingungen für F kommt also entscheidend die Bedingung der Konvexität von $F(t, \mu, \eta)$ hinzu. Das zuvor besprochene Gegenbeispiel erfüllt diese Bedingung nicht.

Beweis: Sei $u_m \in W$ eine Minimalfolge, d.h. $J(u_m) \to \inf J(v)$, $v \in W$. Wegen (6.10) ist $\|u_m\| \leq K$ $(m \to \infty)$. Wir dürfen wegen der Sätze von Banach-Saks und Rellich annehmen, daß $u \to n$ gleichmäßig mit einem $u \in H^{1,p}$. Zunächst beweisen wir, daß

$$F(., u_m, u_m') - F(., u, u_m') \Rightarrow 0$$

dem Maß nach. Hierzu beachten wir, daß das Maß der Menge $E_{mL} = \{t | \, |u_m'(t)| > L\}$ klein ist für großes L, nämlich

$$\mu(E_{mL}) \leq \int\limits_{E_{mL}} \frac{|u_m'(t)|^p}{L^p}\,dt \leq K/L^p\,.$$

Wegen der gleichmäßigen Konvergenz der u_m gilt

$$\chi(E_{mL})F(., u_m, u_m') - F(., u, u_m') \to 0 \text{ punktweise fast überall }.$$

(In der Tat, wenn η_m eine beschränkte Folge ist, gilt $F(t, u_m(t), \eta_m) - F(t, u(t), \eta_m) \to 0$, Beweis durch Widerspruch und Anwendung des Satzes von Bolzano-Weierstraß auf (η_{m_i}).)
Da die punktweise Konvergenz die Maßkonvergenz nach sich zieht, gilt

$$w_{mL} := \chi(E_{mL}) F(., u_m, u_m') - F(., u, u_m') \to 0$$

dem Maß nach. Daraus folgt: Für jedes $\sigma > 0$ gilt:

$$\mu(w_{mL} > \sigma) < \frac{\varepsilon}{2}, \qquad m \geq m(\varepsilon, L, \sigma)\,. \tag{6.11}$$

Seien nun σ und ε vorgegeben. Wähle L so groß, daß $\mu(E_{mL}) \leq \frac{\varepsilon}{2}$. Für $m \geq m(\varepsilon, L(\varepsilon), \sigma)$ gilt (6.11). Aus

$$\mu\big(F(., u_m, u_m')) - F(., u, u_m') > \sigma\big) \leq \mu(E_{mL}) + \mu(w_{mL})$$

folgt

$$\mu\big(F(., u_m, u_m') - F(., u, u_m'\big) < \sigma < \frac{\varepsilon}{2}$$

für $m \geq m(\varepsilon, L(\varepsilon), \sigma)$. Die Maßkonvergenz ist damit bewiesen.
Aus dem Satz von Egoroff folgt damit, daß zu jedem $\varepsilon > 0$ eine Menge E_ε mit $\mu(E_\varepsilon) < \varepsilon$ existiert, so daß

$$F(., u_m, u_m') - F(., u, u_m') \to 0 \text{ gleichmäßig auf } ([a,b]) - E_\varepsilon\,,$$

nach Übergang zu einer Teilfolge. Wir erhalten somit

$$\int\limits_{([a,b])-E_\varepsilon} F(., u, u_m')\,dt \leq \int\limits_{([a,b])-E_\varepsilon} F(., u_m, u_m')\,dt + o(1) \quad (m \to \infty)\,,$$

und wegen $F \geq -K$

$$\int\limits_{([a,b])-E_\varepsilon} F(., u, u_m')\,dt \leq \int_a^b F(., u_m, u_m')\,dt + o(1) + K\mu(E_\varepsilon) \tag{6.12}$$

$$\leq \inf J + o(1) + \tilde{K}\varepsilon\,, \tag{6.13}$$

da $\int_a^b F(., u_m, u_L')\,dt \leq \inf L + \varepsilon,\ m \geq m_0$ - wegen der Eigenschaft, daß (u_m) Minimalfolge ist. Wir halten fest, s. (6.13)

$$\int\limits_{([a,b])-E_\varepsilon} F(., u, u_m')\,dt \leq \inf J + K\varepsilon \qquad \text{für } m \geq m_0\,. \tag{6.14}$$

Wir wenden nun den Satz von Banach-Saks erneut an. Danach gibt es eine Teilfolge der bereits „gezogenen“ Folge, so daß

$$\frac{1}{N}\sum_{j=1}^{N} u_{m_j} \to u \qquad \text{in } H^{1,2} \quad (N \to \infty)$$

konvergiert. Aus (6.14) schließen wir

$$\frac{1}{N}\sum_{m=m_0}^{N+m_0} \int\limits_{([a,b])-E_\varepsilon} F(.,u,u_m')\,dt \le \inf J + K\varepsilon\,,$$

und damit, wegen der vorausgesetzten Konvexität,

$$\int\limits_{([a,b])-E_\varepsilon} F(.,u,\frac{1}{N}\sum_{m=m_0}^{N+m_0} u_m')\,dt \le \inf J + K\varepsilon\,.$$

Da $\frac{1}{N}\sum\limits_{m=1}^{N} u_m' \to u'$ in L^2, gilt auch

$$\frac{1}{N}\sum_{m=m_0}^{N+m_0} u_m' \to u' \text{ in } L^2 \qquad (N \to \infty)\,.$$

Da $F \ge -K$, schließen wir aus dem Satz von Fatou

$$\int\limits_{([a,b])-E_\varepsilon} F(.,u,u_m') \le \liminf \int\limits_{([a,b])-E_\varepsilon} F(.,u,\frac{1}{N}\sum_{m=m_0}^{N+m_0} u_m')\,dt$$

und erhalten

$$\int\limits_{([a,b])-E_\varepsilon} F(.,u,u')\,dt \le \inf J + \varepsilon\,.$$

Durch Grenzübergang $\varepsilon \to 0$ folgt

$$\int\limits_{([a,b])} F(.,u,u')\,dt \le \inf J\,,$$

d.h. u ist minimal. $\square$

7 Regularität schwacher Lösungen von Randwertproblemen gewöhnlicher Differentialgleichungen

Die Variationsrechnung gibt uns Kriterien zum Nachweis der Existenz schwacher Lösungen von Randwertproblemen. Wir kümmern uns daher nicht mehr um die Existenzfrage und nehmen an, daß für $i = 1, \dots, n$ Funktionen F_i und F_{0i} von $[a,b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ gegeben sind, die den Carathéodory-Bedingungen und der folgenden Wachstumsbedingung mit einem $p \in]1, \infty[$ genügen:

$$|F_i(t,\mu,\eta)| + |F_{0i}(t,\mu,\eta)| \le K_C|\eta|^p + K_C \tag{7.1}$$

für alle $\eta \in \mathbb{R}^n$ und alle $\mu \in \mathbb{R}^n$ mit $|\mu| \le C$.

Wir setzen hier nicht mehr voraus, daß sich die F_i und F_{0i} als partielle Ableitungen einer $\mathbb{R}$-wertigen Funktion F schreiben lassen. Für $u \in H^{1,p}$ gilt $F_i, F_{0i} \in L^1$. Im Folgenden schließen wir aus der Beziehung

$$\int_a^b F_i\varphi_i'\,dt + \int_a^b F_{0i}\varphi_i\,dt = 0 \qquad \text{für alle } \varphi_i \in C_0^\infty\,, \tag{7.2}$$

(i fest, keine Summationskonvention) daß F_i absolutstetig ist. Man erinnere sich: Ein Element $F_i \in L^1$ heißt absolutstetig, wenn es eine Funktion $\tilde{F}_i \in [F_i]$ gibt, welche absolutstetig ist. $[F_i]$ ist die Menge aller Funktionen, die sich von F_i nur auf einer Menge vom Maß Null unterscheiden. Für den folgenden Satz wird nicht verlangt, daß F_i, F_{0i} die Gestalt $F_i(.,u,u')$ etc. hat.

Satz 7.1 *Seien $F_i, F_{0i} \in L^1[a,b]$, und es gelte die Bedingung (7.2). Dann ist F_i absolutstetig.*

Die Konsequenz ist erstaunlich, wenn F_i die Form $F_i(.,u,u')$ hat und aus einer Eulerschen Differentialgleichung herrührt. Obwohl $F_i(t,\mu,\eta)$ bezüglich t unstetig sein darf, wird dies dann durch das i.A. unstetige u' so korrigiert, daß $F_i(t,u,u')$ absolutstetig ist.

Weitere Konsequenzen:

Korollar 7.1

$$-F_i' + F_{0i} = 0 \qquad \textit{fast überall in } [a,b], \quad i = 1, \dots, n \tag{7.3}$$

Bemerkung: Gleichung (7.3) ist die Eulersche Differentialgleichung in der üblichen Form, wenn F_i, F_{0i} die Gestalt $F_i = F_i(.,u,u')$, F_{0i} analog, haben.

Beweis: Wir nehmen an, daß $\psi \in C_0^\infty([a,b])$ kompakten Träger in $[a,b]$ hat (d.h. $\psi = 0$ in $U(a)$ und $U(b)$) und setzen in (7.2)

$$\varphi_i = \omega_h * \psi = \int_{-\infty}^{\infty} \omega_h(x-t)\psi(t)\,dt\,.$$

Hierbei ist ω_h eine $C_0^\infty(\mathbb{R})$-Mittelfunktion. Daraus erhalten wir, indem wir $\omega_h *$ auf den ersten Faktor wälzen:

$$(\omega_h * F_i, \psi') + (\omega_h * F_{0i}, \psi) = 0\,.$$

Hierbei bedeutet $(y,z) = \int_a^b yz\,dt$. Durch partielle Integration folgt

$$-((\omega_h * F_i)', \psi) + (\omega_h * F_{0i}, \psi) = 0\,.$$

Wir wählen nun

$$\psi_\rho = \omega_\rho * [\operatorname{sign}(\omega_h * F_i)' \cdot \chi(e)]\,..$$

wobei $e \subset [a+\varepsilon, b-\varepsilon]$ und $\chi(e)$ die charakteristische Funktion von e ist. ψ_ρ ist so konstruiert, daß $\psi_\rho \in C_0^\infty$, und bei festem h

$$|\psi_\rho| \le K_h\,,$$

und $\psi_\rho \to \operatorname{sign}(\omega_h * F_i)' \cdot \chi[a-\varepsilon, b-\varepsilon]$ punktweise fast überall.

Die Konvergenzeigenschaften folgen aus den bekannten Eigenschaften der Faltungsoperation. Wegen des Satzes von Lebesgue darf man zur Grenze übergehen und erhält

$$\int_a^b |(\omega_h * F_i)'|\chi(e)\,dt = \int (\omega_h * F_{0i})\chi(e)\,dt\,.$$

Daraus folgt

$$\int_e |(\omega_h * F_i)'|\,dt \le \int_e |\omega_h * F_{0i}|\,dt\,. \tag{7.4}$$

Wir wählen nun

$$e = \bigcup_{i=1}^{N} [t_i, t_i']\,, \qquad t_i < t_i' < t_{i+1}$$

und $\mu(e) < \delta$. Da $\omega_h * F_{0i} \to F_{0i}$ in L^1, ist die Absolutstetigkeit der Integrale über $\omega_h * F_{0i}$ gleichgradig. Daher ist

$$\int_e |\omega_h * F_{0i}|\,dt \le \varepsilon \qquad \text{gleichmäßig für } h \to 0\,,\ \text{falls } \mu(e) < \delta(\varepsilon)\,.$$

Daraus und aus (7.4) folgt

$$\sum_{j=1}^{N} |(\omega_h * F_i)'(t_j) - (\omega_h * F_i')(t_j')| \leq \int_e |(\omega_h * F_i)| \, dt \leq \varepsilon \,, \tag{7.5}$$

d.h. die $\omega_h * F_i$ sind gleichgradig und absolutstetig im Sinne von Funktionen. Es gibt daher eine gleichmäßig konvergente Teilfolge und eine stetige Funktion $\tilde{F}_i$, die mit F_i fast überall übereinstimmt. Geht man in (7.5) zur Grenze $h \to 0$, erkennt man, daß auch $\tilde{F}_i$ der die Absolutstetigkeit definierenden Bedingung genügt. Es gilt daher

$$F_i \in AC \,.$$

Schließlich wollen wir noch die Gleichung

$$-F_i' + F_{0i} = 0 \qquad \text{fast überall}$$

beweisen. Mit $e = [t, t+s]$ folgt aus (7.2) mit $\varphi = \omega_h * \omega_\rho * \chi(e)$

$$-\int (\omega_h * F_i)' \omega_\rho * \chi(e) \, dt + \int \omega_h * F_{0i} \omega_\rho * \chi(e) \, d\xi = 0$$

Grenzübergang $\rho \to 0$ ergibt

$$\int (\omega_h * F_i)' \chi(e) \, dt + \int \omega_h * F_{0i} \chi(e) \, d\xi = 0$$

und

$$\frac{\omega_h * F_i)(t+s) - (\omega_h * F_i)(t)}{s} + \frac{1}{s} \int_t^{t+s} \omega_h * F_{0i} \, d\xi = 0 \,.$$

Grenzübergang $h \to 0$ ergibt

$$\frac{F_i(t+s) - F_i(t)}{s} + \frac{1}{s} \int_t^{t+s} F_{0i} \, d\xi = 0$$

und Grenzübergang $s \to 0$ die Behauptung. □

Stetigkeit von u'

Ist $F_i(t, \mu, \eta)$ eine C^1-Funktion aller Variablen, so läßt sich die Stetigkeit von u' über den Satz über implizite Funktionen beweisen. Man benötigt hierzu, daß die Matrix $(F_{1\eta}, \ldots, F_{n\eta})$ invertierbar ist. Ist $F_i = \frac{\partial}{\partial \eta_i} F$, so wäre etwa die Bedingung der positiven Definitheit der Matrix $\frac{\partial}{\partial \eta_i} \frac{\partial}{\partial \eta_k} F$ eine übliche, mit der Konvexitätsbedingung aus der Existenztheorie kompatible Bedingung.

8 Eigenwertprobleme für gewöhnliche Differentialoperatoren zweiter Ordnung

Wir beschränken uns hier auf das Standardproblem des sogenannten Sturm-Liouville-Operators

$$Lu = -(a \cdot u')' + cu$$

und betrachten die Eigenwertgleichung

$$Lu = \lambda u \text{ in }]\alpha, \beta[\,, \qquad u(\alpha) = u(\beta) = 0 \tag{8.1}$$

In schwacher Formulierung lautet das Problem: Gesucht ist $u \in H^{1,2}(]\alpha, \beta[)$, $u \neq 0$ und $\lambda \in \mathbb{R}$, so daß

$$A(u, v) := (au', \varphi') + (cu, \varphi) = \lambda(u, \varphi) \tag{8.2}$$

für alle $\varphi \in C_0^\infty$ (oder, äquivalent, $\varphi \in H^{1,2}(]\alpha, \beta[)$. Hierbei sind a und c meßbare und beschränkte Funktionen, also

$$a, c \in L^\infty\,. \tag{8.3}$$

Ferner soll noch die sogenannte Regularitätsbedingung

$$a \geq \alpha_0 > 0 \quad \text{ mit einem } \alpha_0 > 0 \quad \text{ gelten }. \tag{8.4}$$

Da das Wort „Regularitätsbedingung“ viele verschiedene Bedeutungen hat, sollte man lieber „Elliptizitätsbedingung“ in Analogie zu mehrdimensionalen Problemen sagen.)
Eigenwerte sucht man in der Analysis i.A. in $\mathbb{C}$, der hier vorliegende Operator L ist jedoch symmetrisch, denn es gilt

$$(au', v') + (cu, v) = (av', u') + (cv, u)\,,$$

so daß sinnvollerweise die Eigenwerte in $\mathbb{R}$ gesucht werden. Die Lösungen u von (8.1) bzw. (8.2) heißen Eigenfunktionen.

Aus den Ergebnissen von Kapitel 7 wissen wir, daß au' absolutstetig ist und

$$-(au')' + c \cdot u = \lambda u$$

punktweise fast überall gilt.

Satz 8.1 (Hauptsatz zum Eigenwertproblem) *Seien $a, c \in L^\infty$, und es gelte die Elliptizitätsbedingung (8.4). Dann hat das Problem (8.2) abzählbar viele Eigenwerte endlicher Vielfachheit; die Eigenwerte gehen gegen Unendlich. Die Gesamtheit der Eigenfunktionen ist vollständig in $L^2(]\alpha,\beta[)$. Eigenfunktionen zu verschiedenen Eigenwerten sind orthogonal bezüglich des L^2-Skalarproduktes.*

Der Beweis dieses Hauptsatzes wird zweckmäßigerweise im Rahmen der Funktionalanalysis behandelt. Wir beweisen hier nur, daß die Eigenfunktionen vollständig sind und über ein interessantes Minimax-Prinzip berechnet werden können. Die Aussage, daß Eigenfunktionen zu verschiedenen Eigenwerten orthogonal sind, beweist man wie in der linearen Algebra. Für den niedrigsten Eigenwert beweisen wir das folgende Minimumprinzip:

Satz 8.2 *Seien $a, c \in L^\infty$, $a \geq \alpha_0 > 0$, mit einem $\alpha_0 \in \mathbb{R}$. Dann existiert eine Funktion $u \in H_0^{1,2}(]\alpha,\beta[)$ mit*

(i) $\|u\|_{L^2} = 1$

(ii) $A(u,u) = \inf\{A(v,v) | \quad v \in H_0^{1,2}(]\alpha,\beta[) | \quad \|v\|_{L^2} = 1\} := \lambda$

Die Zahl λ ist der kleinste Eigenwert des Sturm-Liouville-Operators. Die Funktion u mit den Eigenschaften (i) und (ii) ist Eigenfunktion zum Eigenwert λ, welcher in (ii) definiert ist.

Im Folgenden beschränken wir uns auf den Fall $A(u,v) = a(u',v')$, d.h. $c = 0$.

Beweis: Wir wählen eine Minimalfolge (u_j) - schließlich ist (ii) ein Minimierungsproblem. Es gilt $\|u_j\|_{L^2} = 1$ und

$$\alpha_0 \int_\alpha^\beta |u_j'|^2 \, dt \leq A(u_j, u_j) \leq \lambda + \varepsilon_j \qquad \varepsilon_j \to 0\,.$$

Wegen der Ungleichung

$$\int_\alpha^\beta |u_j|^2 \leq c_o \int_\alpha^\beta |u_j'|^2 \, dt$$

für Funktionen mit Nullrandbedingungen ist $\|u_j\|_{H^{1,p}} \leq K$ gleichmäßig für $j \to \infty$.
(Ein Dauerbrenner, je nach Hingezogenheit zur Volksgruppe x oder y heißt die Ungleichung „Poincaré“ oder „Friedrichsche“ oder „Wirtingersche“ Ungleichung. Sie gilt auch

in n-Dimensionen und kann als Übungsaufgabe, z.B. über Fouriersche Reihen bewiesen werden.)

Nach den im vorigen Kapitel verwendeten Auswahlsätzen gibt es eine Teilfolge, so daß $u_j \to u$ gleichmäßig und $\frac{1}{N}\sum\limits_{i=1}^{n} u_i' \to u'$ in L^2. Hierbei ist $u \in H_0^{1,2}$. Wegen der gleichmäßigen Konvergenz gilt

$$1 = \|u_j\|_{L^2} \to \|u\|\,,$$

d.h. $\|u\| = 1$ und insbesondere $u \neq 0$. Da $A(\eta,\eta)$ konvex in η ist, kann man den aus dem vorigen Kapitel verwendeten Schluß anwenden und erhält

$$\begin{aligned} A(u,u) &\le \liminf_{N\to\infty} A\Big(\frac{1}{N}\sum_{j=1}^{N} u_j, \frac{1}{N}\sum_{j=1}^{N} u_j\Big) \le \\ &\le \frac{1}{N}\sum_{j=1}^{N} A(u_j,u_j) + \varepsilon_j' \le \lambda + \varepsilon_j' + \sum_{j=1}^{N}\frac{\varepsilon_j}{N} \to \lambda\,. \end{aligned}$$

Daraus folgt $A(u,u) \le \lambda$, und da λ bereits das Infimum ist, gilt $A(u,u) = \lambda$. Das Infimum wird somit angenommen, u sei die minimierende Funktion. Sei nun $v \in H_0^{1,2}$, $\|v\|_{L^2} = 1$ und $(v,u) = 0$. Dann ist auch $u(t) = \cos(tu) + \sin(tv) \in H_0^{1,2}$ und $\|u(t)\|_{L^2} = 1$. Die Funktion $A(u(t),u(t))$ hat daher an der Stelle $t = 0$ ein Minimum und es gilt

$$\begin{aligned} 0 &= \frac{d}{dt}A(u(t),u(t))|_{t=0} = 2A(u(0),\frac{d}{dt}u(0)) \\ &= 2A(u(t), -\sin(tu) + \cos(tv))|_{t=0} = 2A(u,v)\,. \end{aligned}$$

Es gilt also

$$A(u,v) = 0 \qquad \text{für alle} \qquad v \in H_0^{1,2} \text{ mit } (u,v)_{L^2} = 0\,.$$

Daher ist

$$\begin{aligned} A(u,\mu u + \nu v) = \mu A(u,u) = \lambda\mu = \lambda(u,\mu u) = \\ = \lambda(u,\mu u + \nu v)\,. \end{aligned}$$

Alle $H_0^{1,2}$-Funktionen lassen sich in der Form $\varphi = \mu u + \nu v$ darstellen. Daraus folgt

$$A(u,\varphi) = \lambda(u,\varphi)\,, \qquad \varphi \in H_0^{1,2}\,.$$

Schließlich bemerken wir, daß für jede normierte Eigenfunktion v zum Eigenwert $\tilde\lambda$ gilt

$$\lambda = \inf(Au,u) \le (Av,v) = \tilde\lambda\,,$$

d.h. λ ist der kleinste Eigenwert. □

Wir berechnen nun den zweitkleinsten Eigenwert.

Satz 8.3 *Mit den Voraussetzungen von Satz 8.1 sei V_1 der Eigenraum zum niedrigsten Eigenwert $\lambda = \lambda_1$. Dann ist*

$$\lambda_2 = \inf\{A(v,v)| \quad v \in H_0^{1,2}\,,\ \|v\|_{L^2} = 1\,,\ v \perp V_1\} \tag{8.5}$$

der zweitkleinste Eigenwert.

Beweis: Die Annahme des Minimums wird wieder wie in Satz 8.1 bewiesen.
Nach Satz 8.1 ist V_1 mindestens eindimensional. Mit ähnlichen Methoden wie in Satz 8.1 zeigt man, daß das Minimumlproblem (8.1) eine Lösung u_2 hat. Wir beweisen, daß u_2 Eigenfunktion zum Eigenwert λ_2 ist.

Sei $u_1 \in V_1$, $v \perp V_1$ und $v \perp u_2$, $\|u_1\|_{L^2} = \|v\|_{L^2} = 1$, $c_1, c_2 \in \mathbb{R}$, $c_1^2 + c_2^2 = 1$. Dann ist $\|u(t)\|_{L^2} = 1$ mit $u(t) = \cos tu_2 + \sin(c_1u_1 + c_2v)$ und die Funktion

$$g(t) = A(u(t), u(t))$$

hat an der Stelle $t = 0$ ein Minimum. Daher gilt

$$0 = g'(0) = A(u(t)\,, -\sin tu_2 + \cos t(c_1u_1 + c_2v))|_{t=0} = A(u_2, c_1u_1 + c_2v)\,,$$

und nach Multiplikation mit α ergibt sich

$$A(u_2, \tilde{c}_1u_1 + \tilde{c}_2v) = 0$$

für alle $\tilde{c}_1, \tilde{c}_2 \in \mathbb{R}$.
Es gilt nun

$$\beta\lambda_2 = \lambda_2(u_2, \beta u_2) = \lambda_2(u_2, \beta u_2 + \tilde{c}_1u_1 + \tilde{c}_2v)\,,$$

andererseits

$$\begin{aligned} 0 + \beta\lambda_2 &= A(u_2, \tilde{c}_1u_1 + \tilde{c}_2v) + \beta A(u_2, u_2) = \\ &= A(u_2, \beta u_2 + \tilde{c}_1u_1 + \tilde{c}_2v)\,, \end{aligned}$$

und damit

$$(\beta\lambda_2 =) \quad A(u_2, \varphi) = \lambda_2(u_2, \varphi)\,, \qquad \varphi = \beta u_2 + \tilde{c}_1u_1 + \tilde{c}_2v\,.$$

Damit sind alle $\varphi \in H_0^{1,2}$ erfaßt und die Eigenwertgleichung für u_2 ist bewiesen. Schließlich ist klar, daß $\lambda_1 < \lambda_2$, da bei der Infimumsbildung zur Definition von λ_1 mehr Elemente als bei λ_2 zugelassen sind. □

Damit der Beweis präzise ist, muß noch gezeigt werden, daß nicht bereits $V_1 = H_0^{1,2}$ ist. Dies wird durch den folgenden Satz ausgeschlossen:

Satz 8.4 *Unter der Voraussetzung von Satz 8.1 haben die Eigenwerte endliche Vielfachheit.*

> **Beweis:** Man verwendet ein einfaches funktionalanalytisches Argument: Wäre der Eigenraum unendlich dimensional, hätte er eine Orthonormalbasis (φ_j). Wegen der Eigenwertgleichung gilt $A(\varphi_j, \varphi) = \lambda_1(\varphi_j, \varphi)\,,\ \ \varphi \in H_0^{1,2}$ und man könnte folgern, daß $\|\varphi_j',\|_{L^2} \leq K$. Dann könnte man eine in L^2 konvergente Teilfolge auswählen. Es gilt dann $\|\varphi_j - \varphi_h\|_{L^2} \to 0$, andererseits ist $\|\varphi_j - \varphi_h\| = \sqrt{2}$. □

Die weiteren Eigenwerte λ_3, λ_4 etc. werden in Analogie zu Satz 8.3 konstruiert.

Satz 8.5 *Die Eigenwerte $\lambda_1, \lambda_2, \lambda_3 \ldots \lambda_{j-1}$ seien schon konstruiert. Es gelte $\lambda_1 < \lambda_2 < \ldots \lambda_{j-1}$. Dann ist*

$$\lambda_j = \inf\{A(u,u) | u \in H_0^{1,2}\,,\ \|u\|_{L^2} = 1\,, \quad u \perp V_1, \ldots V_{j-1}\}$$

der nächstgrößere Eigenwert des Sturm-Liouville-Operators.

Der Beweis verläuft analog zu Satz 8.3 und wird hier nicht ausgeführt.

Satz 8.6 *Die Eigenwerte können sich nicht im Endlichen häufen.*

Man beweist dies ähnlich wie Satz 8.4.

Satz 8.7 (Vollständigkeit) *Die Menge der endlichen Linearkomkbinationen $\sum_{j=1}^{N} \alpha_j u_j\,,\ u_j \in V_j$, ist dicht in $H_0^{1,2}$.*

> **Beweis:** Andernfalls gäbe es einen abgeschlossenen Teilraum $W \neq \{0\}\,,\ W \in H_0^{1,2}$, so daß $W \perp V_j\,,\ j = 1, 2 \ldots$. Dies ist ein einfacher Satz aus der Funktionalanalysis, den wir nicht beweisen wollen. Man konstruiert sich wiederum durch
>
> $$\lambda = \inf\{A(u,u)\big| \quad \|u\| = 1\,, \quad u \in W\}$$
>
> einen Eigenwert, der größer ist als die bereits konstruierten λ_j. Da die λ_j gegen Unendlich gehen nach Satz 8.4, ist dies nicht möglich.

Folgerung (Vergleiche die Ausführungen über verallgemeinerte Fouriersche Reihen): Jede Funktion $\omega \in L^2$ läßt sich als eine verallgemeinerte, in L^2 konvergente Fourierreihe

$$\omega = \sum_{j=1}^{\infty} (\omega, u_j) u_j$$

darstellen, wobei die u_j die Eigenfunktionen des Sturm-Liouvielle-Operators sind. Es gibt sehr viele detaillierte Untersuchungen über Eigenfunktionen, z.B. Aussagen über die Anzahl der Nullstellen. □

9 Anhang 1

Auswahlsätze in der Analysis

Aus den Anfängervorlesungen kennt man den wichtigen Satz:

Satz 9.1 (Satz von Bolzano-Weierstraß) *Jede beschränkte Folge von Vektoren des $\mathbb{R}^n$ besitzt eine konvergente Teilfolge.*

Dieser Satz ist Grundlage für die meisten weiteren Auswahlsätze der Analysis. Auswahlsätze werden vor allem für Existenzbeweise benötigt. Eine Verfeinerung ist der entsprechende Satz über Doppelfolgen:

Satz 9.2 *Sei $a^k \in \mathbb{R}$, $|a_i^k| \leq K$ für $i, k, \in \mathbb{N}$. Dann gibt es eine Teilfolge $(i_j \in \mathbb{N})_{j=1}^{\infty}$, so daß für jedes $k \in \mathbb{N}$ $(a_{i_j}^k)_{j=1}^{\infty}$ konvergiert $(j \to \infty)$.*

In den Anfängervorlesungen lernt man zumeist folgendes grundlegendes Axiom

Auswahlaxiom (der Mengenlehre): Es sei $\{M_\iota, \iota \in \mathcal{F}\}$, eine Familie von Mengen. Dann gibt es eine Funktion f, die jedem $\iota \in \mathcal{F}$ ein Element aus M_ι zuordnet.

Diese Aussage scheint völlig evident zu sein. Sie ist jedoch problematisch, weil sie zu dem sogenannten **Wohlordnungssatz** äquivalent ist.

Das Auswahlaxiom wird eher außerhalb der Analysis angewandt; eine Ausnahme ist die Konstruktion nicht-meßbarer Mengen sowie der sogenannte Satz von Hahn-Banach aus der Funktionalanalysis, der jedoch in konkreten Anwendungen mit schwächeren Versionen des Auswahlaxioms beim Beweis auskommt.

Ein bekannter Satz aus der Funktionalanalysis besagt, daß ein Analogon des Satzes von Bolzano-Weierstraß in unendlich dimensionalen Räumen ohne Zusatzvoraussetzungen nicht möglich ist.

Satz 9.3 *Sei V ein normierter linearer Raum. Jede beschränkte Folge $(u_m \in V)_{m=1}^{\infty}$ besitze eine konvergente Teilfolge. Dann gilt* $\dim V < \infty$.

Im Hilbert-Raum H ist diese Aussage einleuchtend. Ist $\dim H = 0$, so konstruiert man sich ein Orthonormalsystem (φ_j). Da $\|\varphi_j\| = 1$ ist, ist (φ_j) beschränkt und, wenn der Satz von Bolzano-Weierstraß in H richtig wäre, gäbe es eine konvergente Teilfolge $\varphi_{j_k} \to \varphi$ $(k \to \infty)$.

Dann ist $\|\varphi_{j_k} - \varphi_{j_{k+1}}\| \to 0$, andererseits ist wegen der Orthogonalität $\|\varphi_{j_k} - \varphi_{j_{k+1}}\| = \sqrt{2}$. q.e.a.

Da man in unendlich dimensionalen Banachräumen unbedingt einen Ersatz für Bolzano-Weierstraß haben will, sucht man zwei Auswege:
1. Man stellt Zusatzvoraussetzungen an die Folge.
2. Man schwächt die Topologie ab.

Wir befassen uns zunächst mit Auswahlsätzen, in denen Zusatzannahmen an die Folge gestellt werden. Häufig verwendet wird der folgende Satz:

Satz 9.4 (Satz von Arzela-Ascoli) *Es seien $\Omega \subset \mathbb{R}^n$ offen und die Folge $(f_j : \Omega \to \mathbb{R}^m)$ gleichgradig stetig, d.h. zu $\varepsilon > 0$ $\exists \delta_\varepsilon$, so daß $|f_j(x) - f_j(x')| < \varepsilon$ für alle $x, x' \in \Omega$ mit $|x - x'| < \delta$ und alle $j \in \mathbb{N}$. Ferner seien die f_j gleichmäßig beschränkt. Dann gibt es eine Teilfolge (f_{j_k}) und eine stetige Funktion f, so daß*

$$f_{j_k} \to f \quad \textit{gleichmäßig} \quad (k \to \infty)\,.$$

Der Beweis beruht darauf, daß man auf die doppelt indizierte Folge

$$f_j(r_k)\,, \qquad \bigcup_{k=1}^{\infty}\{r_k\} = \mathbb{Q}^n \cap \Omega$$

den Bolzano-Weierstraßschen Doppelfolgensatz anwendet. Es gibt dann eine Teilfolge (j_l), so daß $f_{j_l}(r_k)$ konvergiert $(l \to \infty,\ k \in \mathbb{N}')$. Mit der gleichgradigen Stetigkeit erhält man dann die Konvergenz von $f_{j_l}(x)$ nicht nur für $x = r_k \in Q^n \cap \Omega$, sondern für alle $x \in \Omega$.

Weniger oft verwendet wird der Satz von Helly:

Satz 9.5 (Satz von Helly) *Es seien $[a,b] \in \mathbb{R}$ und $(f_j : [a,b] \to \mathbb{R})$ eine beschränkte Folge monotoner Funktionen. Dann gibt es eine Teilfolge, die fast überall gegen eine monotone Funktion konvergiert.*

Ein Dauerbrenner - ähnlich wie der Satz von Arzela-Ascoli - ist das folgende Resultat:

Satz 9.6 (Satz von Rellich) *(Spezialfall für Funktionen mit kompaktem Träger) Es sei $p \in [1,\infty]$ und $(f_j \in H^{1,p}(\mathbb{R}^n))_{j=1}^{\infty}$ eine in $H^{1,p}$ gleichmäßig beschränkte Folge mit gleichmäßig beschränktem Träger. Dann gibt es eine Teilfolge (f_{j_k}) und eine Funktion $f \in H^{1,p}$, so daß*

$$f_{j_k} \to f \quad \textit{in } L^p\,.$$

(Dies kann man auch als Abschwächung der Topologie verstehen.)

Beweis: Wir führen den Beweis nur in zwei Dimensionen. Dies ist ausreichend für das Verständnis des n-dimensionalen Falles. Ferner beschränken wir uns auf den Fall $p = 2$ und arbeiten mit Fourier-Analysis. Die Funktionen f_j lassen sich als Fourier-Reihen

$$f_k(x) = \sum_{j \in \mathbb{Z}\times\mathbb{Z}} c_j^k l^{i\alpha jx}$$

schreiben. Hierbei benutzen wir die Multiindex-Schreibweise $j = (j_1, j_2)$, $jx = j_1x_1 + j_2x_2$. Aus der Ungleichung $\|f_j\|_{1,2} \leq K$ folgt

$$\sum_{j \in \mathbb{Z}\times\mathbb{Z}} |c_j^k|^2\, |j|^2 \leq K_\alpha \quad \text{gleichmäßig für } k \in \mathbb{Z}\,. \tag{9.1}$$

Nach dem Bolzano-Weierstraßschen Doppelfolgen-Satz gibt es eine Teilfolge $\Lambda \subset \mathbb{N}$, so daß

$$c_j^k \to c_j \quad \text{für } k \in \Lambda\,, \qquad k \to \infty\,,$$

und es gilt

$$\sum |c_j|^2\, |j|^2 < \infty\,.$$

Wir setzen $f = \sum c_j e^{i\alpha jx}$ und erhalten $f \in H^{1,2}$. Es gilt

$$\|f_k - f\|_{L^2}^2 = K_\alpha \sum_{|j|=N} |c_j^k - c_j|^2 + K_\alpha \sum_{|j|\geq N} |c_j^k - c_j|^2\,.$$

Wegen (9.1) können wir N so groß wählen, daß $K_j \sum\limits_{|j|\geq N} |c_j^k - c_j|^2 \leq \frac{\varepsilon^2}{4}$. Wegen der Konvergenz $c_j^k \to c_j$ fällt auch der erste Teil $K_\alpha \sum\limits_{j \leq N} |C_j^k - C_j|^2 < \frac{\varepsilon^2}{4}$ aus $(k \geq k(\varepsilon))$, und wir erhalten

$$\|f_k - f\|_{L^2}^2 < \varepsilon \quad k(\varepsilon)\,, \quad k \in \Lambda\,.$$

□

Wir erwähnen noch ein Lemma von Hausdorff:

Lemma 9.1 (Lemma von Hausdorff) *Sei $(f_j \in L^p)_{j=1}^\infty$ eine in L^p gleichmäßig beschränkte Folge mit gleichmäßig beschränktem Träger. Die Translationsoperatoren E_h, die definiert sind durch*

$$E_h f(x) = f(x+h)\,,$$

seien gleichmäßig stetig bezüglich j für $|h| \to 0$. Dann gibt es eine in L^p-konvergente Teilfolge (f_{j_k}).

Wir besprechen nun einen anderen Typ von Auswahlsätzen, die auf einer Abschwächung der Topologie beruhen. Dies wird im Fall des Hilbert-Raumes erläutert.

Satz 9.7 *Sei H ein reeller oder komplexer Hilbert-Raum mit Skalarprodukt $(.,.)$ und Norm $\|.\|$. Eine Folge $(u_m \in H)$ konvergiert nach Definition genau dann („stark") gegen ein $u \in H$, wenn*

$$\|u_m - u\| \to 0 \qquad (m \to \infty)\,.$$

Es gibt nun einen weiteren Konvergenzbegriff, nämlich den der „schwachen Konvergenz".

Definition 9.1 *Eine Folge $(u_m \in H)$ konvergiert schwach gegen ein $u \in H$, wenn für alle $\varphi \in H$ gilt:*

$$(u_m, \varphi) \to (u, \varphi) \qquad (m \to \infty)\,.$$

Die starke Konvergenz impliziert die schwache. Umgekehrt gilt dies nur in endlich dimensionalen Räumen. Für schwache Konvergenz vereinbaren wir die Schreibweise $u \rightharpoonup u$.

Satz 9.8 (Satz über die schwache Kompaktheit beschränkter Folgen) *Die Folge $(u_m \in H)_{m \in \mathbb{N}}$ sei beschränkt. Dann gibt es eine schwach konvergente Teilfolge (u_{m_i}) mit $u_{m_i} \rightharpoonup u$ $(i \to \infty)$ und $u \in H$.*

Beweis: Sei φ_j ein vollständiges Orthonormalsystem in H. (Ein solches läßt sich konstruieren, wenn H nicht separabel ist, betrachtet man den von (u_m) erzeugten Hilbert-Raum.) Es seien u_j^m die verallgemeinerten Fourier-Koeffizienten der u_m:

$$u_m = \sum_{j=1}^{\infty} \mu_j^m \varphi_j\,.$$

Da $\|u_m\| \le K$, sind die μ_j^m gleichmäßig beschränkt. Nach dem Doppelfolgensatz gibt es eine Teilfolge $\Lambda \subset \mathbb{N}$, so daß

$$\mu_j^m \to \mu_j \qquad (m \to \infty, m \in \Lambda)$$

für alle $j \in \mathbb{N}$. Da $\|u_m\|^2 = \sum\limits_{j=1}^{\infty} |\mu_j^m|^2 \le K$, gilt auch

$$\sum |\mu_j|^2 \le K \qquad \text{und}$$

$$u := \sum_{j=1}^{\infty} \mu_j \varphi_j \in H\,.$$

Wir zeigen, daß $u^m \rightharpoonup u$ $(m \in \Lambda)$. Hierzu sei $\psi \in H$ mit der verallgemeinerten Fourier-entwicklung

$$\psi = \sum_{j=1}^{\infty} \psi_j \varphi_j\,, \qquad \psi_j \in \mathbb{R} \text{ oder } \mathbb{C}.$$

Es gilt

$$(u^m - u, \psi) = \sum_{j=1}^{\infty} (\mu_j^m - \mu^m)\psi_j = \sum_{j=1}^{N} + \sum_{j>N}\,.$$

Wir wählen N so groß, daß

$$\begin{aligned}\sum_{j>N} |(\mu_j^m - \mu_j)|\,|\psi_j| &\le \Big(\sum_{j=1}^{\infty} |\mu_j^m - \mu_j^m|^2\Big)^{1/2}\Big(\sum_{j>N} |\psi_j|^2\Big)^{1/2} \\ &\le 2K\Big(\sum_{j>N} |\psi_j|^2\Big)^{1/2} < \varepsilon/2\,.\end{aligned}$$

Wegen der Konvergenz $\mu_j^m \to \mu_j$ $(m \to \infty,\ m \in \Lambda)$ gilt $\sum\limits_{j=1}^{N} (\mu_j^m - \mu_j)\psi_j \to 0$. Daraus folgt

$$|(u^m - u, \psi)| < \varepsilon/2 + \varepsilon/2 \qquad \text{für } m < m(\varepsilon)\,, \quad m \in \Lambda\,.$$

□

Für die „konvexe Analysis“ ist der Satz von Banach-Saks wichtig:

Satz 9.9 *Es sei (u_m) eine beschränkte Folge in einem Hilbert-Raum H. Dann gibt es eine Teilfolge (m_j) und ein Element u, so daß die arithmetischen Mittel stark in H gegen u konvergieren:*

$$\frac{1}{N}\sum_{j=1}^{N} u_{m_j} \to u \qquad (N \to \infty)\,.$$

Beweis: Nach dem vorigen Satz gibt es eine Teilfolge, die wir nach Umnummerierung ebenfalls mit (u_m) bezeichnen, so daß

$$u_m \rightharpoonup u \quad \text{schwach in } H$$

mit einem $u \in H$. Wir setzen $v_m = u_m - u$. Es gilt dann $v_m \rightharpoonup 0$ schwach in H, und es genügt zu zeigen, daß es eine Teilfolge (m_j) gibt, so daß $\frac{1}{N}\sum\limits_{j=1}^{N} v_{m_j} \to 0$ stark in H. Dies wird durch folgendes Auswahlverfahren erreicht:

(i) Wähle $v_{m_1} = v_1$.

(ii) Die Elemente $v_{m_1}, \ldots, v_{m_{j-1}}$ seien schon konstruiert. Wähle v_{m_j}, so daß

$$|(v_{m_1}, v_{m_j})| < \frac{1}{j^2}, \ldots |(v_{m_{j-1}}, v_{m_j})| < \frac{1}{j^2}\,.$$

Wegen $v_{m_j} \rightharpoonup 0$ ist dies möglich. Wir behaupten, daß (v_{m_j}) die gesuchte Folge ist:

$$\|\frac{1}{N}\sum_{j=1}^{N} v_{m_j}\|^2 = \frac{1}{N^2}\sum_{j=1}^{N}\|v_{m_j}\|^2 + \frac{2}{N^2}\sum_{j>l}(v_{m_j}, v_{m_l})\,.$$

Da $\|v_{m_j}\| \le K$ gilt

$$\frac{1}{N^2}\sum_{j=1}^{N}\|v_{m_j}\|^2 \le \frac{K}{N} \to v \qquad (N \to \infty)\,.$$

Für den obigen zweiten Summanden gilt nach Konstruktion

$$|\frac{2}{N^2}\sum_{j>l}(v_{m_j}, v_{m_l})| \le \frac{2}{N^2}\sum_{j=1}^{N}\frac{j}{j^2} = \frac{2}{N^2}\sum_{j=1}^{N}\frac{1}{j} = \frac{2}{N}\sum_{j=1}^{N}\frac{1}{j^2} \to 0 \qquad (N \to \infty)\,.$$

Damit folgt die behauptete Konvergenz der arithmetischen Mittel der v_{m_j}. □

Für sogenannte reflexive Banach-Räume wie z.B. die L^p-Räume mit $1 < p < \infty$ gibt es ein Analogon des Satzes von der schwachen Kompaktheit beschränkter Folgen in Hilbertschen Räumen und ein Analogon des Satzes von Banach-Saks.

Im Raum $L^1[a,b]$ ist die entsprechende Aussage falsch, d.h., ist $f_m in L^1, \|f_m\| \le K$, so gibt es i.A. kein Element $f \in L^1$ und keine zugehörige Teilfolge f_{m_i} mit

$$\int_a^b f_{m_i}\varphi\,dx \to \int_a^b f\varphi\,dx \qquad (i \to \infty)$$

für alle $\varphi \in L^\infty[a,b]$. Die Auswahlsätze sind jedoch so wichtig, daß man den „Austritt aus dem Raum“ zuläßt, d.h. in einem geeignet zu definierenden Sinne gehen die f_{m_i} schwach gegen ein Element aus dem sogenannten „bidualen Raum“ $(L^1)^{**}$, den man auch als Raum der Maße interpretieren kann. Diese Fragen werden in fortgeschrittenen Vorlesungen behandelt.

Wie man vielleicht erwartet, gibt es einen wunderschönen Auswahlsatz in der Theorie der komplexen Funktionen einer Variablen, nämlich den

Satz 9.10 (Satz von Montel) *Es sei $\Omega \subset \mathbb{C}$ offen und $f_m : \Omega \to \mathbb{C}$ holomorph sowie $|f_m| \leq K$ gleichmäßig auf Ω. Dann gibt es eine Teilfolge (f_{m_i}) und eine holomorphe Funktion $f : \Omega \to \mathbb{C}$, so daß*

$$f_{m_i} \to f \quad \textit{und} \quad (\frac{d}{dz})^k f_{m_i} \to (\frac{d}{dz})^k f$$

gleichmäßig auf allen $\Omega' \subset\subset \Omega$.

Beweis: Aus der Cauchyschen Integralformel entnimmt man, daß alle Ableitungen $f_{m_i}^{(k)}$ gleichmäßig beschränkt sind für jedes $\Omega' \subset\subset \Omega$:

$$(\frac{d}{dz})^k f_{m_i} = \int_\Gamma (\frac{d}{dz})^k \frac{f_{m_i}(\zeta)}{\zeta - z}\, d\zeta\,.$$

Hierbei ist $\Gamma \subset \Omega - \bar{\Omega}'$. Die f_{m_i} und alle Ableitungen erfüllen damit die Voraussetzungen des Satzes von Arzel-Ascoli, d.h. $f_{m_i} \to f$ gleichmäßig mitsamt allen Ableitungen. Die Holomorphie von f liest man dann wieder aus der Cauchyschen Integralformel ab, wenn man in dieser mit dem Indes m_i zur Grenze übergeht. □

10 Anhang 2

Anwendungen des Satzes von Montel

Gewöhnliche Differentialgleichungen im Komplexen:

Wir betrachten das Anfangswertproblem

$$\begin{cases} y(z_0) & = y_0 \in \mathbb{C}^n \\ \frac{d}{dz}(y) & = f(z, y(z)) \quad \text{in } U(z_0) \end{cases}$$

Hierbei ist f eine komplexe Funktion auf $U(z_0) \times U(y_0) \subset \mathbb{C}^{n+1}$, die durch eine konvergente Potenzreihe in mehreren Variablen dargestellt wird. Es gilt

Satz 10.1 *Es gibt eine holomorphe Lösung $y : V(z_0) \to \mathbb{C}^n$ des Anfangswertproblems mit einer Umgebung $V(z_0) \subset U(z_0)$.*

Beweis: Der Beweis geschieht ähnlich wie im Reellen. Man formt das Anfangswertproblem in eine Integralgleichung mit einem (wegunabhängigen!) Kurvenintergral $\int\limits_{z_0}^{z}$ um:

$$\mathfrak{A}y(z) := y(z) = y_0 + \int\limits_{z_0}^{z} f(\zeta, y(\zeta))\, d\zeta \,,$$

und zeigt, daß für genügend kleine Umgebungen $V(z_0)$ die Abbildung $\mathfrak{A}$ die Menge aller vektorwertigen holomorphen Funktionen y mit $|y - y_0| < \delta$ in sich abbildet und daß das Bild aus gleichgradig stetigen Funktionen besteht. Nach dem Satz von Montel ist die Menge aller dieser Funktionen in der Topologie der gleichmäßigen Konvergenz aller Ableitungen auf kompakten Teilgebieten von $V(z_0)$ kompakt - außerdem gilt die Kompaktheit der Bildmenge von $\mathfrak{A}$ im Raum der stetigen Funktionen. Nach dem Satz von Tychonoff - ein Analogon des Satzes von Schauder - für sogenannte topologische Vektorräume - gibt es daher einen Fixpunkt. □

Das besondere an diesem Beweis ist, daß man ohne große Mühe die Holomorphie der Lösung geschenkt bekommt.

Eine weitere, sehr elegante Anwendung des Satzes von Montel ist der Riemannsche Abbildungssatz, dessen Formulierung und Beweis man z.B. im Bändchen von Cartan über Funktionentheorie nachlesen kann.