Mathematisches Institut der Universität Bonn

Vorlesung im Sommersemester 2003

Variationsrechnung II

Vorläufige Korrekturversion

Privatdozent Dr. Heiko von der Mosel

Letzte Änderung: 5. November 2003

Fehler / Verbesserungsvorschläge bitte an heiko@math.uni-bonn.de.

Erstellt und ausgearbeitet von Philipp Reiter.
Korrekturen von Yvonne Deuster, Anna Engels und Ulrich Menne.
Überarbeitet und ergänzt von Heiko von der Mosel.

Inhaltsverzeichnis

Übersicht

In der „Natur“ beobachtet man viele (zeitlich) stetige aber nicht notwendig glatte Erscheinungen, z. B. Auslenkung eines elastischen Bandes, Leistung einer einseitig beschichteten Solarzelle, Überlaufen einer Talsperre und weitere Schwellwertphänomene; auch Produktionsfunktionen aus der Ökonomie sind hier zu nennen. In der Regel kann man jedoch nicht von der Differenzierbarkeit der entsprechenden Funktionen ausgehen. Dies motiviert die Frage nach alternativen Konzepten, um etwa Minimierungsaufgaben lösen zu können.
Die Kapitel 1 – 6 der diesem Skript zugrundeliegenden Vorlesung sind eine Ausarbeitung verschiedener Abschnitte aus [Cla83], in denen der von FRANK H. CLARKE ab 1973 entwickelte Kalkül der verallgemeinerten Gradienten für LIPSCHITZ-stetige Abbildungen vorgestellt wird.[1]

Kapitel 1 Zentrales Konzept der Theorie sind die *verallgemeinerte Richtungsableitung*

$$f^0(x;v) = \limsup_{\substack{y \to x \\ t \searrow 0}} \frac{f(y+tv) - f(y)}{t}$$

und der *verallgemeinerte Gradient*

$$\partial f(x) := \{\zeta \in X^* \mid \langle \zeta, v\rangle_{X^* \times X} \leq f^0(x;v)\ \forall\, v \in X\}$$

für reellwertige Funktionen f auf einem BANACHraum X. Im Falle LIPSCHITZ-stetiger Funktionen ist $f^0(x;v)$ endlich und $\partial f(x)$ nicht leer, so dass es nahe liegt, sich auf diese Funktionenklasse zu beschränken.
Nach einer Darstellung elementarer Eigenschaften werden Ableitungsbegriffe aus der nichtlinearen Funktionalanalysis eingeführt. Ein Exkurs in die konvexe Analysis liefert eine alternative (und historisch dem CLARKEschen Kalkül vorausgehende) Charakterisierung verallgemeinerter Gradienten über *Subdifferentiale*, sofern die betrachteten Funktionen konvex sind.
Den Abschluss bilden Analoga zur Linearität der klassischen Ableitung (Skalarmultiplikation, Summenregel, Linearkombinationen). Im Allgemeinen können wir hier nur *Inklusionen*, keine Identitäten, der entsprechenden verallgemeinerten Gradienten erhalten. *Identitäts*aussagen liefert der wichtige *Regularitäts*begriff. Als besonders bedeutsam erweist sich die Aussage $0 \in \partial f(x)$ für lokale Extremierer x. Auf dem Mittelwertsatz fußen zwei Kettenregeln, im ersten Fall für $X \to \mathbb{R}^n \to \mathbb{R}$, im zweiten für eine BANACHraumwertige und eine reellwertige Funktion, $X \to Y \to \mathbb{R}$. Für die Identität benötigt man hier erneut schärfere Voraussetzungen, ebenso für die nun folgende

[1] Für die biographischen Anmerkungen wurde [GIS90] benutzt; weitere interessante Informationen findet man unter `http://www-gap.dcs.st-and.ac.uk/~history/BiogIndex.html`.

Produkt- und Quotientenregel. Den Abschluss bilden die Betrachtung cartesischer Produkte und ein einfaches Beispiel aus der Ausgleichsrechnung.

Kapitel 2 Ausgehend von der Distanzfunktion d_C werden der *Tangentialkegel*

$$T_C(x) := \left\{ v \in X \mid d_C^0(x;v) = 0 \right\}$$

und seine *Polarisierung*, der *Normalkegel*, vorgestellt. In Proposition 2.5 wird eine Art *Penalty*-Methode angegeben, Satz 2.10 ist eine fundamentale Charakterisierung des Tangentialkegels, speziellere folgen (z. B. Sätze 2.16, 2.20). Der *Kontingent*-Kegel liefert die Grundlage eines Regularitätsbegriffs für Mengen. Hieraus ergibt sich, dass die Regularität der *Funktion* f an der Stelle x zu der einer *Menge*, nämlich des *Epigraphen* von f in $(x, f(x))$, äquivalent ist. Dies ermöglicht dann eine allgemeinere Definition von verallgemeinertem Gradient und Regularität auch für nicht LIPSCHITZ-stetige Funktionen.

[Resultate aus diesem Kapitel werden, abgesehen von den Propositionen 2.4 und 2.5, in den folgenden nicht verwendet.]

Kapitel 3 Der Satz von RADEMACHER für LIPSCHITZ-Abbildungen gestattet eine alternative Charakterisierung des verallgemeinerten Gradienten als Limes strikt differenzierbarer Funktionen mit Ausnahme einer Nullmenge. Die Beschränkung auf den endlichdimensionalen Fall ermöglicht die Betrachtung vektorwertiger Funktionen, was auf die Definition der *verallgemeinerten* JACOBI*schen* mit Eigenschaften analog zu denen des verallgemeinerten Gradienten führt. Wir erhalten eine vektorwertige Version des Mittelwertsatzes und eine dritte Kettenregel.

[Resultate aus diesem Kapitel werden in den folgenden beiden nicht verwendet.]

Kapitel 4 Vereinfacht gesprochen geht es hier um Vertauschbarkeit von Differentiation und Integration. Für Funktionenscharen $(f_t)_{t\in T}$ werden „Integralinklusionen“ der Form

$$\partial \left(\int_T f_t(\cdot)\,\mathrm{d}\mu(t) \right)(x) \subset \int_T \partial f_t(x)\,\mathrm{d}\mu(t)$$

hergeleitet. Auch hier gibt es Bedingungen für Gleichheit der Mengen. Das zentrale Hilfsmittel [IL74] kann im Rahmen dieser Vorlesung nur zitiert werden. In einem Beispiel werden LAGRANGE-Funktionen diskutiert; bei Differenzierbarkeit folgen die EULER-LAGRANGE-Gleichungen als Spezialfall. Weiterhin wird die Situation $\partial\left(\max_{t\in T} f_t(\cdot)\right)(x)$ untersucht.

Kapitel 5 Mit Hilfe von Satz 4.9 überträgt sich auch eine Version der LAGRANGE-Multiplikatorregel auf die Theorie der verallgemeinerten Gradienten. Ein weiteres Hilfsmittel ist das Variationsprinzip von EKELAND, das allgemein (ohne Zitieren anderer Resultate aus diesem Skript) für vollständige metrische Räume bewiesen wird.

Kapitel 6 Ebenfalls für vollständige metrische Räume gilt eine Verallgemeinerung des BANACHschen Fixpunktsatzes. Unter deutlich schwächeren Voraussetzungen erhalten wir immerhin noch die *Existenz* eines Fixpunktes.

Um eine Version des Umkehrsatzes für LIPSCHITZ-stetige Funktionen herzuleiten, kehren wir wieder zur endlich-dimensionalen Situation aus Kapitel 3 zurück. Hier tritt —

wie auch beim Satz über implizite Funktionen — eine sehr enge Parallele zum differenzierbaren Fall zutage: Im Wesentlichen kommt es nur auf die Überprüfung einer Rang-Bedingung an.

[Dieses Kapitel verwendet bis auf den Satz von EKELAND keine Ergebnisse aus den Kap. 4 und 5.]

Kapitel 7 Zum Abschluss soll die Anwendung der bis hierher erarbeiteten Theorie auf ein konkretes Modell demonstriert werden. Dabei geht es um die mathematische Behandlung von *Selbstkontaktphänomenen elastischer Stäbe* im Rahmen der Variationsrechnung. Hierbei tritt mit dem von GONZALEZ und MADDOCKS eingeführten *globalen Krümmungsradius* für Kurven eine nichtglatte Nebenbedingung auf, die mit dem CLARKEschen Kalkül analysiert wird.

1 Verallgemeinerte Gradienten

Soweit nicht anders vermerkt machen wir folgende

Generalvoraussetzung für Kapitel 1
Im Folgenden seien X, Y immer BANACHräume über $\mathbb{R}$ mit den Normen $\|\cdot\|_X$ bzw. $\|\cdot\|_Y$. $U \subset X$ sei eine offene Teilmenge.

Die Norm des *Dualraums* $X^* := L(X, \mathbb{R})$ ist durch $\|l\|_{X^*} := \sup_{\|x\| \leq 1} |l(x)|$ gegeben. Wir schreiben $\langle l, x \rangle := \langle l, x \rangle_{X^* \times X} := l(x)$.

Grundlagen

Definition 1.1 [LIPSCHITZ-Stetigkeit]
Sei f eine reellwertige Funktion auf U.

(i) f heißt LIPSCHITZ[2]*-stetig* auf U, wenn es ein $K > 0$ gibt mit

$$|f(x) - f(y)| \leq K \|x - y\|_X \text{ für alle } x, y \in X.$$

(ii) f heißt LIPSCHITZ-*stetig nahe* x, wenn es $K > 0$ und $\varepsilon > 0$ gibt, so dass

$$|f(y) - f(y')| \leq K \left\|y - y'\right\|_X \text{ für alle } y, y' \in B_\varepsilon(x).$$

K heißt LIPSCHITZ-*Konstante* von f.

Definition 1.2 [Verallgemeinerte Richtungsableitung]
Sei f eine reellwertige Funktion auf U, $x \in U$, $v \in X$. Wir nennen

$$f^0(x; v) := \limsup_{\substack{y \to x \\ t \searrow 0}} \frac{f(y + tv) - f(y)}{t}$$

verallgemeinerte Richtungsableitung von f an der Stelle $x \in X$.

Proposition 1.3
Sei f LIPSCHITZ-*stetig nahe $x \in U$ mit* LIPSCHITZ-*Konstante K und $v, v_1, v_2 \in X$. DANN gilt:*

[2] RUDOLF LIPSCHITZ (1832 – 1903) Deutscher Mathematiker. Schüler von DIRICHLET, ab 1864 Professor in Bonn. LIPSCHITZ arbeitete auf vielen Gebieten der Mathematik, z. B. Zahlentheorie, FOURIER-Reihen, Differentialgleichungen, Elastizitätstheorie. Bedeutend sind seine Studien über Differentialformen, die einen wichtigen Beitrag zur Entwicklung der Differentialgeometrie darstellen.

(i) $f^0(x;X) \subset \mathbb{R}$, *d. h.* $f^0(x;v)$ *ist endlich.*

(ii) $f^0(x;\lambda v) = \lambda f^0(x;v) \ \forall\ \lambda \geq 0.$ (Homogenität)

(iii) $f^0(x;v_1+v_2) \leq f^0(x;v_1) + f^0(x;v_2).$ (Subadditivität)

(iv) $|f^0(x;v)| \leq K \|v\|_X.$

(v) $f^0(.;.) : X \times X \to \mathbb{R}$ *ist* oberhalbstetig, *d. h. für* $x_i \to x$ *und* $v_i \to v$ *gilt*

$$f^0(x;v) \geq \limsup_{i \longrightarrow \infty} f^0(x_i;v_i).$$

(vi) $|f^0(x;v_1) - f^0(x;v_2)| \leq K \|v_1 - v_2\|_X.$

(vii) $f^0(x;-v) = (-f)^0(x;v).$

Die Eigenschaften (ii) und (iii) bezeichnet man als *Sublinearität* von $f^0(x;\cdot)$.

Beweis.

(i) Für alle $v \in X$ gibt es ein $\delta_0 = \delta_0(v)$, so dass für alle $y \in B_{\delta_0}(x)$ und $t < \delta_0$

$$y, y+tv \in B_\varepsilon(x) \qquad \text{und} \qquad \left|f^0(x;v)\right| = \left|\limsup_{\substack{y\to x\\ t\searrow 0}} \frac{f(y+tv)-f(y)}{t}\right| \leq K \|v\|_X.$$

(ii) $$\limsup_{\substack{y\to x\\ t\searrow 0}} \frac{f(y+t(\lambda v))-f(y)}{t}$$
$$= \limsup_{\substack{y\to x\\ t\searrow 0}} \frac{f(y+(t\lambda)v)-f(y)}{t\lambda}\cdot\lambda = \limsup_{\substack{y\to x\\ t'\searrow 0}} \frac{f(y+t'v)-f(y)}{t'}\cdot\lambda = \lambda f^0(x;v)$$

(iii) $$\limsup_{\substack{y\to x\\ t\searrow 0}} \frac{f(y+t(v_1+v_2))-f(y)}{t}$$
$$\leq \limsup_{\substack{y\to x\\ t\searrow 0}} \frac{f(y+t(v_1+v_2))-f(y+tv_2)}{t} + \limsup_{\substack{y\to x\\ t\searrow 0}} \frac{f(y+tv_2)-f(y)}{t}$$
$$= f^0(x;v_1) + f^0(x;v_2)$$

(iv) Siehe (i).

(v) Zu $x_i \to x$, $v_i \to v$ findet man Folgen $(y_i)_{i\in\mathbb{N}} \subset X$, $(t_i)_{i\in\mathbb{N}} \subset \mathbb{R}$, $t_i > 0$ mit

$$\|y_i - x_i\|_X + t_i \leq \frac{1}{i} \qquad \text{und} \qquad f^0(x_i;v_i) - \frac{1}{i} \leq \frac{f(y_i+t_iv_i)-f(y_i)}{t_i}.$$

Nun ist

$$\begin{aligned} \frac{f(y_i+t_iv_i)-f(y_i)}{t_i} &= \frac{f(y_i+t_iv)-f(y_i)}{t_i} + \frac{f(y_i+t_iv_i)-f(y_i+t_iv)}{t_i} \\ &\leq f^0(x;v) + K\|v_i - v\|_X \\ &= f^0(x;v) + o(1) \qquad \text{für } i \to \infty, \end{aligned}$$

also $\limsup_{i\to\infty} f^0(x_i;v_i) \leq f^0(x;v)$.

(vi) Seien $v_1, v_2 \in X$ beliebig, $t > 0$ nahe 0, y nahe x. Dann erhalten wir

$$\begin{aligned} f(y+tv_1) - f(y) &\leq f(y+tv_2) - f(y) + K\,t\,\|v_1 - v_2\|_X \\ \Longrightarrow \qquad f^0(x;v_1) &\leq f^0(x;v_2) + K\,\|v_1 - v_2\|_X \end{aligned}$$

und durch Vertauschen von v_1 und v_2

$$\left|f^0(x;v_1) - f^0(x;v_2)\right| \leq K\,\|v_1 - v_2\|_X\,.$$

(vii) $$f^0(x;-v) = \limsup_{\substack{y\to x\\ t\searrow 0}} \frac{f(y-tv)-f(y)}{t} = \limsup_{\substack{y\to x\\ t\searrow 0}} \frac{(-f)(y)-(-f)(\overbrace{y-tv}^{=:u})}{t}$$
$$= \limsup_{\substack{u\to x\\ t\searrow 0}} \frac{(-f)(u+tv)-(-f)(u)}{t} = (-f)^0(x;v)$$

□

Definition 1.4 [Verallgemeinerter Gradient]
Sei $f : U \to \mathbb{R}$ Lipschitz-stetig nahe x. Dann heißt

$$\partial f(x) := \{\zeta \in X^* \mid \langle \zeta, v\rangle_{X^*\times X} \leq f^0(x;v)\ \forall\, v \in X\}$$

verallgemeinerter Gradient von f an der Stelle x.

Proposition 1.5
Sei $f : U \to \mathbb{R}$ Lipschitz-stetig in der Nähe von $x \in U$. DANN gilt:

(i) $\partial f(x) \neq \emptyset$

(ii) $\partial f(x)$ *ist konvex.*

(iii) $\partial f(x)$ *ist schwach∗-kompakt.*

(iv) $\|\zeta\|_{X^*} \leq K\ \forall\, \zeta \in \partial f(x)$

(v) $f^0(x;v) = \max_{\zeta\in\partial f(x)} \langle \zeta, v\rangle_{X^*\times X}$, $f^0(x;\cdot)$ *ist also eine Stützfunktion von* $\partial f(x)$ *im Sinne von Definition 1.7.*

Beweis. (i) Aufgabe 1.1 (a)

(ii) Seien $\zeta, \xi \in \partial f(x)$. Dann gilt für alle $\lambda \in [0,1]$

$$\begin{aligned} \langle \lambda\zeta + (1-\lambda)\xi, v\rangle_{X^*\times X} &= \lambda \underbrace{\langle \zeta, v\rangle_{X^*\times X}}_{\leq f^0(x;v)} + (1-\lambda)\underbrace{\langle \xi, v\rangle_{X^*\times X}}_{\leq f^0(x;v)} \\ &\leq f^0(x;v). \end{aligned}$$

(iii) Da $\partial f(x)$ in der Menge $\left\{\zeta \in X^* \,\middle|\, \left|\langle \zeta, v\rangle_{X^*\times X}\right| \le 1 \;\forall\, v \in B_{1/K}(0) \subset X\right\}$ liegt, die nach dem Satz von Banach[3]-Alaoglu[4] A.12 schwach$*$-kompakt ist, brauchen wir nur noch die schwach$*$-Abgeschlossenheit von $\partial f(x)$ zu zeigen. $X^* \setminus \partial f(x)$ ist offen, denn

$$X^* \setminus \partial f(x) \overset{!}{=} \bigcup_{v \in X \subset X^{**}} v^{-1}(f^0(x;v), \infty) \in \mathfrak{T}^*,$$

vgl. A.8 (iii) und Definition A.9. In der Tat haben wir einerseits zu jedem $\zeta \in X^* \setminus \partial f(x)$ ein $v_\zeta \in X$ mit $\langle \zeta, v_\zeta\rangle_{X^*\times X} > f^0(x; v_\zeta)$, also $\zeta \in v_\zeta^{-1}(f^0(x;v_\zeta), \infty)$, was „$\subset$" beweist. Andererseits gibt es für jedes $\zeta \in \bigcup_{v\in X} v^{-1}(f^0(x;v), \infty)$ ein $v_\zeta \in X$ mit $\langle \zeta, v_\zeta\rangle_{X^*\times X} > f^0(x; v_\zeta)$, woraus sich „$\supset$" ergibt.

(iv) Aus $\langle \zeta, v\rangle_{X^*\times X} \le f^0(x;v)$ für alle $v \in X$ folgt mit Proposition 1.3 (iv) einerseits $\langle \zeta, v\rangle_{X^*\times X} \le K\,\|v\|_X$, andererseits $-\langle \zeta, v\rangle_{X^*\times X} = \langle \zeta, -v\rangle_{X^*\times X} \le f^0(x;-v) \le K\,\|-v\|_X$, also $\|\zeta\|_{X^*} \le K$.

(v) Angenommen, es gäbe ein $v \in X$ mit

$$f^0(x;v) > \max_{\zeta \in \partial f(x)} \langle \zeta, v\rangle_{X^*\times X}\,. \tag{1.1}$$

Dann ist $v \neq 0$, und nach Aufgabe 1.1 (b) könnten wir ein $\zeta \in X^*$ mit

$$\langle \zeta, v\rangle_{X^*\times X} = f^0(x;v) \qquad \text{und} \qquad \langle \zeta, w\rangle_{X^*\times X} \le f^0(x;w) \text{ für alle } w \in X$$

finden, also $\zeta \in \partial f(x)$ im Widerspruch zu (1.1).

□

Beispiel 1.6 [Betragsfunktion] Betrachte $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto |x|$. Für $x \neq 0$ ist

$$\limsup_{\substack{y\to x\\ t\searrow 0}} \frac{f(y+tv) - f(y)}{t} = (\operatorname{sign} x)\, v,$$

mithin

$$\partial f(x) = \left\{\zeta \in \mathbb{R} \,\middle|\, f^0(x;v) \ge \zeta v \;\forall\, v \in \mathbb{R}\right\} = \{\operatorname{sign} x\}.$$

Betrachten wir nun den Fall $x = 0$.

$v > 0$ Wir wählen Folgen $(y_i)_{i\in\mathbb{N}}, (t_i)_{i\in\mathbb{N}} \subset \mathbb{R}$, $y_i, t_i > 0$, also auch $y_i + t_i v > 0$, so dass $\frac{f(y_i+t_iv)-f(y_i)}{t_i} = v$ die Ungleichung $f^0(0;v) \ge v$ liefert. Da f die Lipschitz-Konstante $K = 1$ hat, impliziert Proposition 1.3 (iv) $\left|f^0(0;v)\right| \le |v| = v$, und wir erhalten insgesamt $f^0(0;v) = v$.

[3] Stefan Banach (1892 – 1945) Polnischer Mathematiker, Oberhaupt der polnischen Mathematikerschule, die in den 1920er und 30er Jahren Mengenlehre und Funktionalanalysis stark beeinflussten. Banach hatte sich sein mathematisches Wissen weitgehend autodidaktisch erworben und wurde ab 1916 von Hugo Steinhaus gefördert. Seine Hauptleistung besteht in der Theorie *linearer Operatoren* auf den nach ihm benannten *vollständigen normierten linearen Räumen*, in die er Ergebnisse aus Topologie und Kategorientheorie mit einfließen ließ. Seine bedeutendste Arbeit ist der *Satz von der offenen Abbildung* (1929), Satz A.14 im Anhang; im Gegensatz zu Hahn entwirft Banach hier einen Ansatz einer allgemeinen *Dualitätstheorie*, die dann die Theorie der *lokal-konvexen Räume* beeinflusste.

[4] Leonidas Alaoglu (1. Hälfte 20. Jh. (?)) Griechischer (?) Mathematiker.

$v = 0$ $\quad f^0(0;0) = 0.$

$v < 0$ $\quad$ Wir wählen Folgen $(y_i)_{i\in\mathbb{N}}, (t_i)_{i\in\mathbb{N}} \subset \mathbb{R}$, $y_i < 0, t_i > 0$, also $y_i + t_i v < 0$, so dass $\frac{f(y_i+t_iv)-f(y_i)}{t_i} = -v$ die Ungleichung $f^0(0;v) \geq -v$ liefert. Proposition 1.3 (iv) ergibt $\left|f^0(0;v)\right| \leq |v| = -v$, und schließlich $f^0(0;v) = -v$.

Für beliebiges $v \in \mathbb{R}$ haben wir jetzt $f^0(0;v) = |v|$ und

$$\partial f(0) = \left\{\zeta \in \mathbb{R} \middle| \underbrace{f^0(x;v)}_{=|v|} \geq \zeta v \;\forall\, v \in \mathbb{R}\right\} = [-1,1].$$

◇

Definition 1.7 [Stützfunktionen]

- Für ein nicht-leeres $C \subset X$ nennen wir

$$\sigma_C : X^* \to \mathbb{R} \cup \{+\infty\}, \qquad \zeta \mapsto \sup_{x\in C} \langle \zeta, x\rangle_{X^*\times X}$$

die *Stützfunktion von C*.

- Für ein nicht-leeres $\Sigma \subset X^*$ ist

$$\sigma_\Sigma : X \to \mathbb{R} \cup \{+\infty\}, \qquad x \mapsto \sup_{\eta\in\Sigma} \underbrace{\langle \eta, x\rangle_{X^*\times X}}_{=\langle x,\eta\rangle_{X^{**}\times X^*}}$$

die *Stützfunktion von Σ*.

Proposition 1.8 [Hörmander[5]]
Seien $C, D \subset X$ nicht-leer, konvex und abgeschlossen, $\Sigma, \Delta \subset X^*$ nicht-leer, konvex und schwach$*$-abgeschlossen. DANN gilt:

(i) $C \subset D \iff \sigma_C(\zeta) \leq \sigma_D(\zeta) \quad \forall\, \zeta \in X^*.$

(ii) $\Sigma \subset \Delta \iff \sigma_\Sigma(x) \leq \sigma_\Delta(x) \quad \forall\, x \in X.$

(iii) *Σ ist genau dann schwach$*$-kompakt, wenn $\sigma_\Sigma(X) \subset \mathbb{R}$.*

(iv) *$\sigma : X \to \mathbb{R} \cup \{+\infty\}$ sei positiv homogen erster Ordnung, subadditiv, unterhalbstetig und nicht konstant $+\infty$. DANN existiert ein nicht-leeres, konvexes und schwach$*$-abgeschlossenes $\Sigma \subset X^*$ mit $\sigma(x) = \sigma_\Sigma(x)$ für alle $x \in X$.*

Den Beweis findet man in [Hör55].

Definition 1.9 [Mengenwertige Funktionen]
Seien X, Y topologische Räume.

[5] Lars Hörmander (geb. 1931) Schwedischer Mathematiker. Hörmander studierte in Lund, wo er Schüler von M. Riesz war, lehrte später in den U. S. A. und war ab 1968 Professor in Lund. Seine Arbeiten befassen sich überwiegend mit der Theorie partieller Differentialgleichungen und komplexer Analysis. 1962 wurde ihm die Fields-Medaille verliehen.

(i) Wir bezeichnen eine Funktion $\Gamma : X \to \mathfrak{P}(Y)$ als *mengenwertige Funktion* und $\mathrm{Graph}(\Gamma) := \{(x,y) \mid x \in X, y \in \Gamma(x)\}$ als ihren *Graphen.*

(ii) Γ heißt *abgeschlossen*, wenn $\mathrm{Graph}(\Gamma)$ in der Topologie von $X \times Y$ abgeschlossen ist.

(iii) Falls X und Y topologische Vektorräume sind (vgl. Definition A.5), so nennen wir Γ *oberhalbstetig* an der Stelle $x \in X$, wenn es für jede Umgebung V von $\Gamma(x)$ eine Umgebung U von x mit $\Gamma(U) \subset V$ gibt. Im BANACHraumfall ist die Bedingung äquivalent zu

$$\forall\, \varepsilon > 0\ \exists\, \delta > 0 : \Gamma(B_\delta(x)) \subset B_\varepsilon(\Gamma(x)).$$

Γ heißt *oberhalbstetig auf* $X' \subset X$, wenn Γ in allen Punkten aus X' oberhalbstetig ist.

Bemerkung. Besteht $\Gamma(x)$ für jedes $x \in X$ aus genau einem Element $\gamma(x) \in Y$, so ist Γ *oberhalbstetig* an der Stelle x genau dann, wenn $\gamma : X \to Y$ *stetig* an der Stelle x ist.

Proposition 1.10
Sei $f : U \to \mathbb{R}$ LIPSCHITZ-stetig nahe $x \in U$.

(i) Sei $(x_i)_{i\in\mathbb{N}}$ eine Folge in X mit $x_i \to x$, $(\zeta_i)_{i\in\mathbb{N}}$ eine Folge in X^* mit $\zeta_i \in \partial f(x_i)$ für alle $i \in \mathbb{N}$ und ζ ein Häufungspunkt von (ζ_i) bezüglich der schwach$*$-Folgenkonvergenz. DANN gilt $\zeta \in \partial f(x)$.

(ii) $\partial f(x) = \bigcap_{\delta>0} \bigcup_{y\in B_\delta(x)} \partial f(y)$

(iii) Für endlich-dimensionale X ist $\partial f(.)$ stets oberhalbstetig an der Stelle x.

Beweis.

(i) Sei $v \in X$, dann findet man eine Teilfolge $(\zeta_{i_k})_{k\in\mathbb{N}}$ mit $\langle \zeta_{i_k}, v\rangle_{X^*\times X} \to \langle \zeta, v\rangle_{X^*\times X}$ für $k \to \infty$. Nach Definition 1.4 folgt $f^0(x_{i_k}; v) \geq \langle \zeta_{i_k}, v\rangle_{X^*\times X} \to \langle \zeta, v\rangle_{X^*\times X}$. Mit Proposition 1.3 (v) ergibt sich $f^0(x; v) \geq \langle \zeta, v\rangle_{X^*\times X}$, also $\zeta \in \partial f(x)$.

(ii) „$\subset$" klar; „$\supset$": Sei $\zeta \in \bigcap_{\delta>0} \bigcup_{y\in B_\delta(x)} \partial f(y)$, dann gibt es für jedes $\delta > 0$ ein $y \in B_\delta(x)$ mit $\zeta \in \partial f(y)$. Für eine Nullfolge $\delta_i \searrow 0$ können wir daher eine Folge $(y_i)_{i\in\mathbb{N}} \subset X$, $y_i \to x$, $\zeta \in \partial f(y_i)$ konstruieren. Aus $\langle \zeta, v\rangle_{X^*\times X} \leq f^0(y_i; v)$ folgt mit Proposition 1.3 (v) gerade $\langle \zeta, v\rangle_{X^*\times X} \leq \limsup_{i\to\infty} f^0(y_i; v) \leq f^0(x; v)$.

(iii) Falls $\partial f(.)$ nicht oberhalbstetig bei x ist, finden wir ein $\varepsilon > 0$, so dass es zu jedem $\delta > 0$ ein $x' \in B_\delta(x) \subset X$ mit $\partial f(x') \not\subset B_\varepsilon(\partial f(x)) \subset X^*$ gibt, d. h. es existiert ein $\zeta \in \partial f(x')$ mit $\zeta \notin B_\varepsilon(\partial f(x))$. Für eine Folge $\delta_i \searrow 0$ gibt es $(x_n)_{n\in\mathbb{N}} \subset X$ mit $x_i \to x$ und eine Folge

$$(\zeta_i)_{i\in\mathbb{N}} \subset X^*, \qquad \zeta_i \in \partial f(x_i), \qquad \zeta_i \notin B_\varepsilon(\partial f(x)).$$

Wähle i so groß, dass x_i in der Umgebung von x enthalten ist, auf der f LIPSCHITZ-stetig mit LIPSCHITZ-Konstante K ist. Dann liefert Proposition 1.5 (iv) $\|\zeta_i\|_{X^*} \leq K$. Wegen $\dim X < \infty$ ist $\|\zeta_i\|_X = \|\zeta_i\|_{X^*}$ beschränkt (siehe [Alt02, Ü 4.3]), also gibt es eine Teilfolge $(\zeta_{i_k})_{k\in\mathbb{N}} \subset X$ und ein $\zeta \in X$ mit $\zeta_{i_k} \xrightarrow{k\to\infty} \zeta$ in X und damit $\zeta_{i_k} \rightharpoonup \zeta$ in X^*. (i) liefert nun $\zeta \in \partial f(x)$ im Widerspruch zu $\|\zeta_i - \eta\|_{X^*} \geq \varepsilon$ für alle $\eta \in \partial f(x)$.

□

Ableitungen in Banachräumen

Definition 1.11 [Richtungs-, Gâteaux-, Hadamard-, Fréchet- und strikte Ableitung]
Seien X, Y Banachräume und $F : X \to Y$.

(i) $F'(x; v) := \lim_{t \searrow 0} \frac{F(x+tv)-F(x)}{t}$ heißt, sofern der Limes existiert, *einseitige Richtungsableitung von F an der Stelle x in Richtung v.*

(ii) Wir sagen, F besitze eine Gâteaux[6]*-Ableitung an der Stelle x*, wenn es ein $DF(x) \in L(X, Y)$ mit $F'(x; v) = \langle DF(x), v\rangle_{L(X,Y)\times X}$ für alle $v \in X$ gibt.

F heißt *stetig (*Gâteaux-*)differenzierbar an der Stelle x*, wenn es eine Umgebung $N \subset X$ von x gibt, so dass $DF(y)$ für alle $y \in N$ existiert und die Abbildung $DF(\cdot) : N \to L(X, Y)$ stetig in der Operatornorm ist.

(iii) Sofern das Argument des Limes $t \searrow 0$ in der Definition von $F'(x; v)$ auf jedem Kompaktum V *gleichmäßig* für alle $v \in V$ konvergiert, wird $DF(x)$ Hadamard[7]*-Ableitung* genannt.

Gilt dies sogar für alle beschränkten Mengen V, so heißt $DF(x)$ Fréchet[8]*-Ableitung.*

(iv) F besitzt eine *strikte Ableitung* $D_sF(x) \in L(X, Y)$ *an der Stelle x*, wenn

$$\lim_{\substack{x' \to x \\ t \searrow 0}} \frac{F(x' + tv) - F(x')}{t} = \langle D_sF(x), v\rangle_{L(X,Y)\times X} \quad \text{für alle } v \in X$$

und die Konvergenz $(x' \to x, t \searrow 0)$ gleichmäßig in v auf jedem Kompaktum $V \subset X$ ist.

Die genannten Objekte sind im Existenzfall eindeutig, weil Grenzwerte in Hausdorff[9]-Räumen eindeutig sind und Abbildungen durch die Werte auf allen Urbildern eindeutig bestimmt sind.

Bemerkung. Wenn $F'(x; v)$ existiert, dann ist $F'(x; v) \le F^0(x; v)$; eine hinreichende Bedingung für Gleichheit ist die Existenz der strikten Ableitung; vgl. Definition 1.30 und Proposition 1.32.

[6] René Gâteaux (1880 – 1914) Französischer Mathematiker, Schüler von Hadamard. Gâteaux arbeitete auf dem Gebiet der Funktionalanalysis und befasste sich vor allem mit der Übertragung allgemeiner funktionentheoretischer Begriffe auf Operatoren in topologischen Räumen. Gâteaux fiel zu Beginn des Ersten Weltkriegs; ein Teil seiner Arbeiten wurde posthum herausgegeben.

[7] Jaques Hadamard (1865 – 1963) Französischer Mathematiker. Ausgehend von funktionentheoretischen Arbeiten, vor allem über Taylor-Reihen, beeinflusste Hadamard erfolgreich viele Bereiche der Mathematik, z. B. die Theorie der Differentialgleichungen, Differentialgeometrie und vor allem die Funktionalanalysis. Bei der Besetzung Frankreichs emigrierte Hadamard in die U. S. A., später nach England, kehrte aber 1945 nach Paris zurück.

[8] Maurice Fréchet (1878 – 1963) Französischer Mathematiker, Schüler von Hadamard. Fréchet arbeitete auf verschiedenen Gebieten, v. a. Topologie, Integrationstheorie, Funktionalanalysis und Wahrscheinlichkeitstheorie. Mit seiner Habilitation (1906) leistete er einen bedeutenden Beitrag zur Theorie der metrischen Räume. Auf ihn gehen die Begriffe *separabel* und *relativ folgenkompakt* zurück. Sein Differentialbegriff spielt in der nichtlinearen Funktionalanalysis eine große Rolle.

[9] Felix Hausdorff (1868 – 1942) Deutscher Mathematiker, Professor in Leipzig, Greifswald und Bonn. Statt einer Kurzbiographie ein Literaturhinweis: Mehrtens, Herbert: Felix Hausdorff — *Ein Mathematiker seiner Zeit*, Mathematisches Institut der Universität Bonn, 1980.

Proposition 1.12
Seien U eine Umgebung von x und $F : U \to Y$ eine Funktion. DANN sind äquivalent:

(i) F ist strikt differenzierbar bei x mit $D_sF(x) = \zeta$.

(ii) F ist LIPSCHITZ*-stetig nahe x und $\lim_{\substack{x' \to x \\ t \searrow 0}} \frac{F(x'+tv)-F(x')}{t} = \langle \zeta, v \rangle_{X^* \times X}$ für alle $v \in X$.*

Beweis. „(i) $\Rightarrow$ (ii)“: Hier ist nur die LIPSCHITZ-Stetigkeit zu zeigen. Wäre F nicht LIPSCHITZ-stetig nahe x, so gäbe es für alle $i \in \mathbb{N}$ Punkte $x_i, x_i' \in B_{1/i}(x) \subset X$, die

$$\left\|F(x_i) - F(x_i')\right\|_Y > i \left\|x_i - x_i'\right\|_X$$

erfüllen. Wir definieren zwei Nullfolgen $v_i := \frac{x_i - x_i'}{\sqrt{i}\left\|x_i - x_i'\right\|_X} \in X, t_i := \sqrt{i}\,\|x_i - x_i'\|_X \in (0, \frac{2}{\sqrt{i}})$ und setzen $V := \{v_i \mid i \in \mathbb{N}\} \cup \{0\}$. V ist offenbar kompakt, also finden wir für jedes $\varepsilon > 0$ ein $n_\varepsilon \in \mathbb{N}$, so dass für alle $i \geq n_\varepsilon$

$$\left\| \frac{F(x_i' + t_i v) - F(x_i')}{t_i} - \langle D_sF(x), v \rangle \right\|_Y < \varepsilon \quad \forall\, v \in V \tag{1.2}$$

gilt. Nun haben wir einerseits wegen $x_i' = x_i + t_i v_i$

$$\left\| \frac{F(x_i + t_i v_i) - F(x_i)}{t_i} \right\|_Y > \sqrt{i} \xrightarrow{i \to \infty} \infty,$$

andererseits

$$\|\langle D_sF(x), v_i \rangle\|_Y \leq \|\zeta\|_{L(X,Y)} \|v_i\|_X \leq \|\zeta\|_{L(X,Y)} \frac{1}{\sqrt{i}} \xrightarrow{i \to \infty} 0,$$

was ein Widerspruch zu (1.2) ist.
„(ii) $\Rightarrow$ (i)“: Sei $V \subset X$ ein Kompaktum und $\varepsilon > 0$. Dann finden wir nach Voraussetzung für jedes $v \in V$ ein $\delta = \delta(v) > 0$, so dass für alle $x' \in B_\delta(x)$ und $t \in (0, \delta)$

$$\left\| \frac{F(x' + tv) - F(x')}{t} - \langle \zeta, v \rangle \right\|_Y < \varepsilon.$$

Wir fordern außerdem $\delta(v) < \frac{\varepsilon}{K + \|\zeta\|_{X^*}}$, wobei K die LIPSCHITZ-Konstante von F ist. Damit folgt

$$\left\| \frac{F(x' + tv') - F(x')}{t} - \frac{F(x' + tv) - F(x')}{t} \right\|_Y \leq K \left\|v' - v\right\|_X,$$

also ergibt sich zusammen für alle $x' \in B_\delta(x), v' \in B_\delta(v)$ und $t \in (0, \delta)$

$$\left\| \frac{F(x' + tv') - F(x')}{t} - \left\langle \zeta, v' \right\rangle \right\|_Y < K \left\|v' - v\right\|_X + \varepsilon + \|\zeta\|_{X^*} \left\|v' - v\right\|_X < 3\varepsilon.$$

Die Mengen $B_{\delta(v)}(v)$ bilden eine offene Überdeckung, aus der man wegen der Kompaktheit endlich viele auswählen kann, die immer noch V ganz überdecken. Man nehme nun den Radius des kleinsten Balls als δ für die gleichmäßige Konvergenz. $\square$

Corollar 1.13
Wenn $F : X \to Y$ stetig (Gâteaux-)differenzierbar bei x ist, dann ist F strikt differenzierbar bei x.

Beweisskizze. Um $DF(x) = D_sF(x)$ zu zeigen, genügt es nach Proposition 1.12 nachzuweisen, dass F Lipschitz-stetig und $\lim_{\substack{x' \to x \\ t \searrow 0}} \frac{F(x'+tv)-F(x')}{t} = \langle DF(x), v\rangle_{X^* \times X}$ für alle $v \in X$ ist. Nach einem Mittelwertsatz für Vektoren [McL65] folgt ersteres aus dem zweiten. Wir zeigen also wegen A.8 (i) für Folgen $x_i \to x$, $t_i \searrow 0$, $v_i \to v$

$$\sup_{\|\theta\|_{Y^*} \leq 1} \left\langle \theta, \frac{F(x_i + t_i v_i) - F(x_i)}{t_i} - \langle DF(x), v\rangle_{L(X,Y)\times X} \right\rangle_{Y^* \times Y} \xrightarrow{i\to\infty} 0.$$

Nach dem eben zitierten Vektor-Mittelwertsatz gibt es ein Folge $(x_i^*)_{i\in\mathbb{N}}$ mit $x_i^* \in \{x_i + \lambda t_i v_i \mid \lambda \in [0,1]\}$ und

$$\left\langle \theta, \frac{F(x_i + t_i v_i) - F(x_i)}{t_i} \right\rangle_{Y^* \times Y} = \left\langle \theta, \langle DF(x_i^*), v_i\rangle_{L(X,Y)\times X} \right\rangle_{Y^* \times Y}.$$

Damit ergibt sich

$$\begin{aligned} &\left| \left\langle \theta, \frac{F(x_i + t_i v_i) - F(x_i)}{t_i} - \langle DF(x), v\rangle_{L(X,Y)\times X} \right\rangle_{Y^* \times Y} \right| \\ \leq\ & \|\theta\|_{Y^*} \left(\left| \langle DF(x_i^*), v_i - v\rangle_{L(X,Y)\times X} \right| + \left| \langle DF(x_i^*) - DF(x), v\rangle_{L(X,Y)\times X} \right| \right) \xrightarrow{i\to\infty} 0. \end{aligned}$$

□

Proposition 1.14
Sei $f : U \to \mathbb{R}$ Lipschitz-stetig nahe $x \in U$, und f besitze eine Gâteaux- (oder Hadamard- oder Fréchet-) Ableitung $Df(x)$ bei x. DANN gilt $Df(x) \in \partial f(x)$.

Beweis. Es gilt

$$f^0(x; v) = \limsup_{\substack{y \to x \\ t \searrow 0}} \frac{f(y + tv) - f(y)}{t} \geq \lim_{t \searrow 0} \frac{f(x + tv) - f(x)}{t} = \langle Df(x), v\rangle$$

für alle $v \in X$. □

Bemerkung. Es gibt Beispiele, für die $\{Df(x)\} \subsetneqq \partial f(x)$ gilt und die Voraussetzungen der Proposition 1.14 erfüllt sind, vgl. Aufgabe 1.4.

Proposition 1.15
Sei f eine reellwertige Funktion auf U, $x \in U$.

(i) Falls f strikt differenzierbar bei x ist, DANN ist f Lipschitz-stetig nahe x und $\partial f(x) = \{D_s f(x)\}$.

(ii) Falls f Lipschitz-stetig nahe x und $\partial f(x) = \{\zeta\}$ für genau ein $\zeta \in X^*$, DANN ist f strikt differenzierbar bei x und $D_s f(x) = \zeta$.

Beweis.

(i) Nach Proposition 1.12 folgt aus der strikten Differenzierbarkeit bei x die LIPSCHITZ-Stetigkeit nahe x. Aus

$$f^0(x;v) = \lim_{\substack{x'\to x\\ t\searrow 0}} \frac{f(x'+tv)-f(x')}{t} = \langle D_s f(x), v\rangle \quad \text{für alle } v \in X \tag{1.3}$$

folgt $\partial f(x) \supset \{D_s f(x)\}$. Um „$\subset$" zu sehen, betrachten wir ein beliebiges $\zeta \in \partial f(x)$. Dann gilt

$$\langle \zeta, v\rangle = -\langle \zeta, -v\rangle \geq -f^0(x;-v) \overset{(1.3)}{=} -\langle D_s f(x), -v\rangle = \langle D_s f(x), v\rangle \overset{(1.3)}{=} f^0(x;v) \geq \langle \zeta, v\rangle$$

für alle $v \in X$, mithin $\zeta = D_s f(x)$.

(ii) Nach Proposition 1.12 brauchen wir nur $\lim_{\substack{x'\to x\\ t\searrow 0}} \frac{f(x'+tv)-f(x')}{t} = \langle \zeta, v\rangle$ für alle $v \in X$ zu zeigen. $\zeta \in \partial f(x)$ bedeutet $f^0(x;v) \geq \langle \zeta, v\rangle$ für alle $v \in X$. Mit dem Satz von HAHN-BANACH A.11 findet man (vgl. Aufgabe 1.1 (b)) zu jedem $v \in X \setminus \{0\}$ ein lineares $\zeta_v : X \to \mathbb{R}$ mit

$$\langle \zeta_v, v\rangle = f^0(x;v) \quad \text{und} \quad \langle \zeta_v, w\rangle \leq f^0(x;w)\ \forall\, w \in X,$$

also $\zeta_v \in X^*$ und wegen Proposition 1.3 (iv) $\zeta_v \in \partial f(x) = \{\zeta\}$. Das bedeutet aber $\zeta_v = \zeta$, mithin

$$\langle \zeta, v\rangle = f^0(x;v) \text{ für alle } v \in X. \tag{1.4}$$

Wir zeigen nun die Existenz der strikten Ableitung:

$$\begin{aligned}
\liminf_{\substack{x'\to x\\ t\searrow 0}} \frac{f(x'+tv)-f(x')}{t} &= -\limsup_{\substack{x'\to x\\ t\searrow 0}} \frac{f(x')-f(x'+tv)}{t}\\
&= -\limsup_{\substack{x'\to x\\ t\searrow 0}} \frac{f(x'+tv-tv)-f(x'+tv)}{t}\\
&= -\limsup_{\substack{y\to x\\ t\searrow 0}} \frac{f(y-tv)-f(y)}{t}\\
&= -f^0(x;-v) \overset{(1.4)}{=} -\langle \zeta, -v\rangle = \langle \zeta, v\rangle\\
&\overset{(1.4)}{=} f^0(x;v) = \limsup_{\substack{x'\to x\\ t\searrow 0}} \frac{f(x'+tv)-f(x')}{t}.
\end{aligned}$$

Folglich existiert $\lim_{\substack{x'\to x\\ t\searrow 0}} \frac{f(x'+tv)-f(x')}{t}$ und ist gleich $\langle \zeta, v\rangle$.

$\square$

Corollar 1.16
Sei X endlich-dimensional, $f : U \to \mathbb{R}$ LIPSCHITZ-stetig nahe $x \in U$, $\varepsilon > 0$. DANN sind äquivalent:

(i) Für alle $x' \in B_\varepsilon(x)$ gibt es ein $\zeta_{x'} \in X^*$ mit $\partial f(x') = \{\zeta_{x'}\}$.

(ii) f ist stetig differenzierbar in $B_\varepsilon(x)$.

Beweis. „(ii) $\Rightarrow$ (i)“: f ist nach Corollar 1.13 strikt differenzierbar in $x' \in B_\varepsilon(x)$. Mit Proposition 1.15 (i) folgt $\partial f(x') = \{D_s f(x')\}$.
„(i) $\Rightarrow$ (ii)“: Nach Proposition 1.15 (ii) existiert $\{D_s f(x')\}$ für alle $x' \in B_\varepsilon(x)$. Aus Proposition 1.10 (iii) folgt wegen $\dim X < \infty$ die Oberhalbstetigkeit von $\partial f(.)$. Nach der Bemerkung zu Definition 1.9 ist $D_s f(.)$ stetig in $B_\epsilon(x)$. □

Beispiel 1.17 Seien $X = \mathbb{R}$, $U = (0,1)$, $\phi \in L^\infty(U)$, $f(x) := \int_0^x \phi(s)\,\mathrm{d}s$. Wegen

$$|f(x) - f(y)| \leq \left| \int_y^x \phi(s)\,\mathrm{d}s \right| \leq \|\phi\|_{L^\infty(0,1)}\, |x-y|$$

ist f Lipschitz-stetig, und nach Proposition 1.5 (i) ist $\partial f(x) \neq \emptyset$ für alle $x \in (0,1)$. Ferner gilt nach [Rud87, 7.10] $f'(x) = \phi(x)$ für fast alle $x \in (0,1)$, also ergibt sich $\phi(x) \in \partial f(x)$ für fast alle $x \in (0,1)$ mit Proposition 1.14.
Wir betrachten nun eine Folge $x_i \xrightarrow{i\to\infty} x \in (0,1)$ mit $\phi(x_i) \in \partial f(x_i)$ für alle $i \in \mathbb{N}$. Ist ζ Häufungspunkt von $(\phi(x_i))_{i\in\mathbb{N}}$, so folgt $\zeta \in \partial f(x)$ mit Proposition 1.10 (i). Wir setzen nun

$$\phi^+(x) := \inf_{\substack{N\subset(0,1)\\|N|=0}} \sup_{\substack{(x_i)_{i\in\mathbb{N}}\\ \subset(0,1)\setminus N}} \limsup_{x_i\to x} \phi(x_i) \quad \text{und} \quad \phi^-(x) := \sup_{\substack{N\subset(0,1)\\|N|=0}} \inf_{\substack{(x_i)_{i\in\mathbb{N}}\\ \subset(0,1)\setminus N}} \liminf_{x_i\to x} \phi(x_i).$$

Nach obiger Überlegung liegen $\phi^+(x)$ und $\phi^-(x)$ in $\partial f(x)$, und mit Proposition 1.5 (ii) folgt $[\phi^-(x), \phi^+(x)] \subset \partial f(x)$.
Andererseits gilt $\frac{f(y+t)-f(y)}{t} = \frac{1}{t}\int_y^{y+t} \phi(s)\,\mathrm{d}s$. In Aufgabe 1.5 (b) wird $f^0(x;1) = \limsup_{\substack{y\to x\\ t\searrow 0}} \frac{f(y+t)-f(y)}{t} \overset{!}{\leq} \phi^+(x)$ gezeigt, womit sich $\zeta \cdot 1 \leq f^0(x;1) \leq \phi^+(x)$ und ebenso $\zeta \geq \phi^-(x)$ für alle $\zeta \in \partial f$ ergibt. Dies liefert

$$[\phi^-(x), \phi^+(x)] = \partial f(x).$$

◇

Konvexität

Für detailliertere Informationen zur *konvexen Analysis* verweisen wir auf [Roc70] und [BF71].

Definition 1.18 [Konvexe Funktionen]
Eine auf einer konvexen Menge $U \subset X$ definierte reellwertige Funktion f heißt *konvex*, falls für je zwei Punkte $x, y \in U$ und jedes $\lambda \in (0,1)$

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$$

gilt.

Proposition 1.19
Sei $U \subset X$ (offen und) konvex und $f : U \to \mathbb{R}$ eine konvexe Funktion; weiterhin gebe es ein $p \in U$ und eine offene Umgebung von p in U, auf der f beschränkt sei. DANN ist f Lipschitz-stetig nahe jeden Punktes $x \in U$.

Der folgende Beweis dieser Aussage geht auf ROBERTS und VARBERG zurück [RV74]. Dort findet man auch ein Beispiel einer konvexen Funktion auf einem normierten Raum, die nirgendwo beschränkt ist: Man betrachte den Raum der Polynome $p = p(x)$ in einer Variablen über den reellen Zahlen normiert durch die Supremumsnorm auf $[-1, 1]$. Die konvexe Funktion f ist gegeben durch $f(p) := |p'(0)|$, tatsächlich gilt

$$f(\lambda p + (1-\lambda)q) = |\lambda p'(0) + (1-\lambda)q'(0)| \leq \lambda f(p) + (1-\lambda)f(q) \quad \forall\, \lambda \in (0,1).$$

Andererseits kann man aber ein festes Polynom p in der Supremumsnorm auf $[-1, 1]$ beliebig gut durch Polynome p_j mit vielen Oszillationen und $|p_j'(0)| \to \infty$ für $j \to \infty$ approximieren, so dass man keine Umgebung von p finden kann, auf der f beschränkt ist.

Beweis von Proposition 1.19. Sei $x \in U$ und ohne Einschränkung $p = 0$; anderenfalls betrachte $\tilde{f}(x) := f(x + p)$. Nach Voraussetzung gibt es $\varepsilon > 0$, $M > 0$, so dass $|f(\tilde{p})| \leq M$ für alle $\tilde{p} \in B_\varepsilon(0)$. Wähle nun $\varrho > 1$ mit $y := \varrho x \in U$. Dann gilt für $\lambda := \frac{1}{\varrho} \in (0, 1)$

$$V := \left\{ v \in X \mid \exists\, x' \in B_\varepsilon(0) : v = (1-\lambda)x' + \lambda y \right\} \stackrel{!}{=} B_{\varepsilon(1-\lambda)}(x),$$

also ist V eine Umgebung von x und $V \subset U$, da U konvex. Für $v \in V$ erhalten wir nun

$$\begin{aligned} f(v) = f((1-\lambda)x' + \lambda y) &\leq (1-\lambda)f(x') + \lambda f(y) \\ &\leq (1-\lambda)M + \lambda \underbrace{f(y)}_{=:c<\infty}, \end{aligned}$$

also ist $f|_V$ nach oben beschränkt.
Zu jedem $z \in V$ findet man ein $z' \in V$ mit $x := \frac{z+z'}{2}$, und es gilt $f(x) = f(\frac{z+z'}{2}) \leq \frac{1}{2}f(z) + \frac{1}{2}f(z')$, also

$$\begin{aligned} f(z) &\geq 2f(x) - f(z') \\ &\geq \underbrace{2f(x)}_{=:c'>-\infty} - \underbrace{((1-\lambda)M + \lambda f(y))}_{\leq(1-\lambda)M+\lambda c<\infty}, \end{aligned}$$

also ist $f|_V$ nach unten beschränkt.
Mit Corollar 1.20 folgt die Behauptung, wenn man dort $U := V$, $\delta := \frac{\varepsilon}{2}(1-\lambda)$, $V := B_\delta(x)$ setzt. ☐

Corollar 1.20
Sei $U \subset X$ (offen und) konvex und $f : U \to \mathbb{R}$ eine konvexe Funktion mit $|f| \leq M$; weiterhin enthalte U die δ-Umgebung einer Menge V. Dann gilt

$$|f(x) - f(y)| \leq \frac{2M}{\delta} \|x - y\|_X \quad \textit{für alle } x, y \in V.$$

Der Beweis wird als Aufgabe 1.7 gestellt.

Definition 1.21 [Subdifferential]
Sei $U \subset X$ (offen und) konvex, $f : U \to \mathbb{R}$ eine konvexe Funktion und $x \in U$. Dann heißt

$$\text{Sub}\, f(x) := \{\zeta \in X^* \mid f(x') - f(x) \geq \left\langle \zeta, x' - x \right\rangle_{X^* \times X} \ \forall\, x' \in U\}$$

das *Subdifferential von f an der Stelle x*.

Proposition 1.22
Sei $U \subset X$ (offen und) konvex und $f : U \to \mathbb{R}$ eine konvexe und nahe $x \in U$ LIPSCHITZ-stetige Funktion mit LIPSCHITZ-Konstante K. DANN gilt

$$\operatorname{Sub} f(x) = \partial f(x) \quad \textit{und} \quad f'(x;v) = f^0(x;v) \quad \textit{für alle } v \in X.$$

Beweis. Wir zitieren aus der konvexen Analysis die Existenz der Richtungsableitung $f'(x;\cdot)$ und die Identität $f'(x;\cdot) = \sigma_{\operatorname{Sub} f(x)}(.)$ für jeden Punkt $x \in U$. Zusammen mit $f^0(x;\cdot) = \sigma_{\partial f(x)}(\cdot)$ brauchen wir nach Proposition 1.8 (ii) nur die zweite Aussage zu zeigen.
Sei $\delta > 0$ vorgegeben. Dann gilt

$$\begin{aligned} f^0(x;v) &= \lim_{\varepsilon \searrow 0} \sup_{x' \in B_{\varepsilon\delta}(x)} \sup_{t \in (0,\varepsilon)} \underbrace{\frac{f(x'+tv) - f(x')}{t}}_{\nearrow \text{ in } t, \text{ da konvex}} \\ &= \lim_{\varepsilon \searrow 0} \sup_{x' \in B_{\varepsilon\delta}(x)} \frac{f(x'+\varepsilon v) - f(x')}{\varepsilon}. \end{aligned}$$

Außerdem haben wir für $x' \in B_{\epsilon\delta}(x)$

$$\left| \frac{f(x+\varepsilon v) - f(x)}{\varepsilon} - \frac{f(x'+\varepsilon v) - f(x')}{\varepsilon} \right| \leq \frac{2K}{\varepsilon} \left\| x - x' \right\|_X \leq 2\delta K,$$

so dass

$$f^0(x;v) \leq \lim_{\varepsilon \searrow 0} \frac{f(x+\varepsilon v) - f(x)}{\varepsilon} + 2\delta K = f'(x;v) + 2\delta K,$$

also mit $\delta \to 0$ $f'(x;y) \leq f^0(x;v) \leq f'(x;v)$ wegen der Bemerkung nach Definition 1.11. □

Beispiel 1.23 Betrachte die Funktion $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) := \max_{i \in \{1,\dots,n\}} x_i$, wobei $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. f ist als Maximum linearer Funktionen konvex und LIPSCHITZ-stetig. Wir setzen $I(x) := \{i \in \{1,\dots,n\} \mid f(x) = x_i\}$ und sehen

$$\begin{aligned} f'(x;v) &= \lim_{t \searrow 0} \frac{\max_{i \in \{1,\dots,n\}}(x_i + tv_i) - f(x)}{t} \\ &= \lim_{t \searrow 0} \frac{\max_{i \in I(x)}(x_i + tv_i) - f(x)}{t} \qquad (1.5) \\ &= \lim_{t \searrow 0} \frac{\max_{i \in I(x)}(x_i + tv_i - x_i)}{t} = \max_{i \in I(x)} v_i, \end{aligned}$$

wobei (1.5) wegen $x_i < f(x) \Rightarrow x_i + tv_i < \max_{i \in \{1,\dots,n\}}(x_i + tv_i)$ für $i \notin I(x)$ und $t \ll 1$ richtig ist. Also

$$\begin{aligned} \partial f(x) &= \left\{ \zeta \in \mathbb{R}^n \,\middle|\, \zeta \cdot v \leq \max_{i \in I(x)} v_i \quad \forall\, v \in \mathbb{R}^n \right\} \\ &= \left\{ \zeta \in \mathbb{R}^n \mid \textstyle\sum_{i=1}^n \zeta_i = 1, \zeta_i \geq 0 \;\forall\, i = 1,\dots,n,\; \zeta_j = 0 \;\forall\, j \notin I(x) \right\}. \end{aligned}$$

Für die TSCHEBYSCHEW[10]-Norm $f(x) := \max_{i\in\{1,\dots,n\}} |x_i|$ ergibt sich ein anderer verallgemeinerter Gradient $\partial f(x)$, vgl. Aufgabe 1.9. $\diamond$

Proposition 1.24
Sei $U \subset X$ (offen und) konvex, $f : U \to \mathbb{R}$ eine LIPSCHITZ-stetige Funktion. DANN sind äquivalent:

(i) f ist konvex auf U.

(ii) $\langle \zeta - \zeta', x - x'\rangle_{X^*\times X} \geq 0$ für alle $x, x' \in U, \zeta \in \partial f(x), \zeta' \in \partial f(x')$.

Beweisskizze. „(i) $\Rightarrow$ (ii)“: Seien $\zeta \in \partial f(x) = \text{Sub } f(x), \zeta' \in \partial f(x') = \text{Sub } f(x')$. Dann gilt

$$\begin{aligned} f(x') &\geq f(x) + \langle \zeta, x' - x\rangle_{X^*\times X}, \\ f(x) &\geq f(x') + \langle \zeta', x - x'\rangle_{X^*\times X}, \end{aligned}$$

und durch Addition dieser beiden Gleichungen ergibt sich (ii).
„(ii) $\Rightarrow$ (i)“: Der Beweis beruht auf dem Mittelwertsatz 1.33 für CLARKE-Gradienten. $\square$

Elementare Rechenregeln

Proposition 1.25 [Skalarmultiplikation]
Sei $f : U \to \mathbb{R}$ LIPSCHITZ-stetig nahe $x \in U$. DANN gilt $\partial(sf)(x) = s\partial f(x)$ für alle $s \in \mathbb{R}$.

Beweis. sf ist offenbar LIPSCHITZ-stetig nahe x. Sei zunächst $s \geq 0$. Dann ist

$$\begin{aligned} (sf)^0(x; v) &= \limsup_{\substack{y\to x\\ t\searrow 0}} \frac{(sf)(y + tv) - (sf)(y)}{t} \\ &= s \limsup_{\substack{y\to x\\ t\searrow 0}} \frac{f(y + tv) - f(y)}{t} = sf^0(x; v), \end{aligned}$$

mithin

$$\begin{aligned} \partial(sf)(x) &= \{\zeta \in X^* \mid \langle \zeta, v\rangle_{X^*\times X} \leq (sf)^0(x; v)\ \forall\, v \in X\} \\ &= \{\zeta \in X^* \mid \exists\, \bar{\zeta} \in \partial f(x) : \zeta = s\bar{\zeta}\} = s\partial f(x). \end{aligned}$$

[10] PAFNUTI LWOWITSCH TSCHEBYSCHEW, *auch* CHEBYSHEV (1821 – 1894) Russischer Mathematiker; Professor in St. Petersburg. TSCHEBYSCHEW entstammte einer adeligen Familie. Seine 1837 an der Moskauer Universität begonnenen Studien schloss er 1841 mit einer Arbeit über numerische Auflösbarkeit algebraischer Gleichungen höheren Grades ab. 1846 promovierte er über Wahrscheinlichkeitstheorie und habilitierte sich anschließend in St. Petersburg, wo er 1850 eine Professur erhielt. In der nachfolgenden Zeit begründete er den Ruf dieser Stadt als geistiges Zentrum einer ganzen Generation von Mathematikern. TSCHEBYSCHEW betrieb sowohl Grundlagenforschung als auch angewandte Mathematik. Neben bedeutenden Arbeiten zu Zahlentheorie (anknüpfend an Vorarbeiten EULERS), Primzahlverteilung, Konvergenz von Reihen, quadratischen Formen, Approximationstheorie, elliptischen Formen und vor allem zur Wahrscheinlichkeitstheorie beschäftigte er sich auch mit dem Modell einer Rechenmaschine, arbeitete in verschiedenen technischen Kommissionen (z. B. für die Artillerie) und erwarb sich Verdienste um die Kartographierung Russlands (hier ebenfalls auf Vorarbeiten EULERS aufbauend).

Sei nun $s < 0$. Wir brauchen den Beweis nur für $s = -1$ zu führen.

$$\begin{aligned} \zeta \in \partial(-f)(x) \quad &\iff \quad (-f)^0(x;v) \geq \langle \zeta, v\rangle_{X^*\times X} \ \forall\, v \in X \\ &\overset{\text{Prop. 1.3 (vii)}}{\iff} \quad f^0(x;-v) \geq \underbrace{\langle \zeta, v\rangle_{X^*\times X}}_{=\langle -\zeta,-v\rangle_{X^*\times X}} \ \forall\, v \in X \\ &\iff \quad -\zeta \in \partial f(x) \iff \zeta \in -\partial f(x) \end{aligned}$$

□

Proposition 1.26 [Lokale Extrema]
Falls $f : U \to \mathbb{R}$ bei $x \in U$ lokal minimal (oder maximal) ist, so gilt $0 \in \partial f(x)$.

Beweis. Sei f minimal in x; anderenfalls betrachte $-f$ mit $\partial(-f)(x) = -\partial f(x)$. Nun gilt

$$f^0(x;v) = \limsup_{\substack{y\to x\\ t\searrow 0}} \frac{f(y+tv)-f(y)}{t} \geq \limsup_{t\searrow 0} \frac{f(x+tv)-f(x)}{t} \geq 0 \text{ für alle } v \in X,$$

da x Minimierer ist. Da $\langle 0, v\rangle_{X^*\times X} = 0 \leq f^0(x;v)$, folgt die Behauptung direkt aus Definition 1.4. □

Proposition 1.27 [Summenregel]
Gegeben seien nahe $x \in U$ LIPSCHITZ-stetige Funktionen $f_1, \ldots, f_n : U \to \mathbb{R}$. DANN gilt

$$\partial\left(\textstyle\sum_{i=1}^n f_i\right)(x) \subset \textstyle\sum_{i=1}^n \partial f_i(x) := \left\{\textstyle\sum_{i=1}^n \zeta_i \,\middle|\, \zeta_i \in \partial f_i(x)\right\}.$$

Beweis. Nach Proposition 1.8 genügt es, $\sigma_{\partial(\sum_{i=1}^n f_i)(x)}(v) \leq \sigma_{\sum_{i=1}^n \partial f_i(x)}(v)$ für alle $v \in X$ zu zeigen, denn $\partial\left(\sum_{i=1}^n f_i\right)(x)$ und $\sum_{i=1}^n \partial f_i(x)$ sind schwach$*$-abgeschlossen (vgl. dazu die Argumentation am Anfang des Beweises von Satz 1.36, Seite 23) und konvex. Wir beschränken uns auf $n = 2$; der allgemeine Fall folgt dann induktiv.

$$\begin{aligned} \sigma_{\partial(f_1+f_2)(x)}(v) &= \sup_{\zeta\in\partial(f_1+f_2)(x)} \langle \zeta, v\rangle \overset{\text{Prop. 1.5 (v)}}{=} (f_1+f_2)^0(x;v) \\ &= \limsup_{\substack{y\to x\\ t\searrow 0}} \frac{(f_1+f_2)(y+tv)-(f_1+f_2)(y)}{t} \\ &\leq \limsup_{\substack{y\to x\\ t\searrow 0}} \frac{f_1(y+tv)-f_1(y)}{t} + \limsup_{\substack{y\to x\\ t\searrow 0}} \frac{f_2(y+tv)-f_2(y)}{t} \\ &= f_1^0(x,v) + f_2^0(x,v) \overset{\text{Prop. 1.5 (v)}}{=} \max_{\zeta_1\in\partial f_1(x)} \langle \zeta_1, v\rangle + \max_{\zeta_2\in\partial f_2(x)} \langle \zeta_2, v\rangle \\ &= \max_{\zeta\in\partial f_1(x)+\partial f_2(x)} \langle \zeta, v\rangle \overset{\text{Prop. 1.5 (v)}}{=} \sigma_{\sum_{i=1}^n \partial f_i(x)}(v) \end{aligned}$$

□

Corollar 1.28 [Gleichheit in Summenregel]
Gegeben seien nahe $x \in U$ LIPSCHITZ-stetige Funktionen $f_1, \ldots, f_n : U \to \mathbb{R}$, von denen $f_1, \ldots, f_{n-1}$ strikt differenzierbar bei x seien. DANN gilt

$$\partial\left(\textstyle\sum_{i=1}^n f_i\right)(x) = \textstyle\sum_{i=1}^n \partial f_i(x).$$

Beweis. $F := \sum_{i=1}^{n-1} f_i$ ist strikt differenzierbar bei x. Wenn wir $F^0(x;v) + f_n^0(x;v) = (F + f_n)^0(x;v)$ zeigen, folgt die Behauptung mit der Argumentation aus dem Beweis von Proposition 1.27. Man überlege sich zunächst für reelle Folgen $(a_n), (b_n)_{n\in\mathbb{N}}, a_n \to a$, dass $\limsup(a_n + b_n) = a + \limsup b_n$. Nun erhalten wir

$$\begin{aligned}
(F + f_n)^0(x;v) &= \limsup_{\substack{y\to x \\ t\searrow 0}} \frac{(F + f_n)(y + tv) - (F + f_n)(y)}{t} \\
&= \lim_{\substack{y\to x \\ t\searrow 0}} \frac{F(y + tv) - F(y)}{t} + \limsup_{\substack{y\to x \\ t\searrow 0}} \frac{f_n(y + tv) - f_n(y)}{t} \\
&= F'(x;v) + f_n^0(x;v) = F^0(x;v) + f_n^0(x;v),
\end{aligned}$$

wobei sich die letzte Umformung aus der Bemerkung nach Definition 1.11 ergibt. □

Corollar 1.29 [Linearkombinationen]
Gegeben seien nahe $x \in U$ LIPSCHITZ-stetige Funktionen $f_1, \ldots, f_n : U \to \mathbb{R}$ und $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$. DANN gilt

$$\partial\left(\sum_{i=1}^n \alpha_i f_i\right)(x) \subset \sum_{i=1}^n \alpha_i \partial f_i(x).$$

Gleichheit der Mengen besteht, wenn $f_1, \ldots, f_{n-1}$ strikt differenzierbar bei x sind.

Beweis. Die Inklusion ergibt sich aus den Propositionen 1.25 und 1.27; mit den f_i sind auch die $\alpha_i f_i$ strikt differenzierbar für $i = 1, \ldots, n-1$, so dass Corollar 1.28 die Gleichheit beweist. □

Definition 1.30 [Reguläre Funktionen]
$f : U \to \mathbb{R}$ heißt *regulär* an der Stelle $x \in U$, falls die Richtungsableitung $f'(x;v)$ für jedes $v \in X$ existiert und mit der verallgemeinerten Richtungsableitung $f^0(x;v)$ übereinstimmt.

Bemerkung. Ist f regulär, so folgt im Allgemeinen *nicht* die Regularität von $-f$.

Corollar 1.31
In der Summenregel (Proposition 1.27) gilt Gleichheit, wenn $f_1, \ldots, f_n$ regulär.
In der Regel für Linearkombinationen (Corollar 1.29) gilt Gleichheit, wenn $f_1, \ldots, f_n$ regulär und $\alpha_1, \ldots, \alpha_n \geq 0$.

Beweis. Wie wir noch sehen werden (Proposition 1.32 (iii)), ist Regularität stabil unter nicht-negativen Linearkombinationen. Wir beschränken uns wieder auf $n = 2$.

$$(f_1 + f_2)^0(x;v) \overset{\text{Def. 1.30}}{=} (f_1 + f_2)'(x;v) = f_1'(x;v) + f_2'(x;v) \overset{\text{Def. 1.30}}{=} f_1^0(x;v) + f_2^0(x;v)$$

□

Proposition 1.32 [Reguläre Funktionen]
Sei $f : U \to \mathbb{R}$ LIPSCHITZ-stetig nahe $x \in U$. DANN gilt:

(i) f *strikt differenzierbar bei* $x \Longrightarrow f$ *regulär bei* x

(ii) f *konvex* $\Longrightarrow f$ *regulär*

(iii) $f_1, \ldots, f_n$ *regulär,* $s_1, \ldots, s_n \geq 0 \Longrightarrow \sum_{i=1}^n s_i f_i$ *regulär*

(iv) f GÂTEAUX-*differenzierbar bei* x, f *regulär* $\Longrightarrow \partial f(x) = \{Df(x)\}$

Beweis. (i) vgl. Definition 1.11 (iv) und die sich daran anschließende Bemerkung.

(ii) Ist f konvex, so folgt nach einem Satz aus der konvexen Analysis die Existenz der Richtungsableitung $f'(x;v)$.

Mit Proposition 1.22 erhalten wir $f^0(x;v) = f'(x;v)$.

(iii) Wir betrachten wieder den Fall $n = 2$. Für $s \geq 0$ gilt

$$(sf)'(x;v) = sf'(x;v) \qquad \text{und} \qquad (sf)^0(x;v) = sf^0(x;v),$$

wobei wir letztere Identität im Beweis zu Proposition 1.25 gesehen haben. Ferner

$$(f_1 + f_2)' = f_1' + f_2' \overset{\text{Def. 1.30}}{=} f_1^0 + f_2^0 \ \geq \ (f_1 + f_2)^0 \geq (f_1 + f_2)',$$

wobei sich die letzte Abschätzung aus der Bemerkung nach Definition 1.11 ergibt.

(iv) Wir haben $f^0(x;v) = f'(x;v) = \langle Df(x), v\rangle$ für alle $v \in X$, mithin $Df(x) \in \partial f(x)$. Für jedes $\zeta \in \partial f(x)$ gilt nun $\langle \zeta, v\rangle \leq f^0(x;v) = \langle Df(x), v\rangle$. Mit $v \leftrightarrow -v$ erhalten wir $\zeta = Df(x)$, vgl. Beweis zu Proposition 1.15 (i).

□

Im Folgenden setzen wir für $x, y \in X$

$$\begin{aligned} [x,y] &:= \{\lambda x + (1-\lambda)y \mid \lambda \in [0,1]\}\,, \\ (x,y) &:= \{\lambda x + (1-\lambda)y \mid \lambda \in (0,1)\}\,. \end{aligned}$$

Satz 1.33 [Mittelwertsatz von LEBOURG [Leb75]]
Seien $x, y \in X$, $[x,y] \subset U$ *und* $f : U \to \mathbb{R}$ LIPSCHITZ-*stetig mit* LIPSCHITZ-*Konstante* K. *DANN gibt es ein* $u \in (x,y)$ *mit*

$$f(y) - f(x) \in \langle \partial f(u), y - x\rangle_{X^* \times X}\,. \tag{1.6}$$

Um den Mittelwertsatz zu beweisen, formulieren wir zunächst das folgende

Lemma 1.34
Unter den Voraussetzungen von Satz 1.33 gilt für $g : [0,1] \to \mathbb{R}$, $g(t) := f(x_t)$ *mit* $x_t := x + t(y-x)$, *dass* g LIPSCHITZ-*stetig ist und*

$$\partial g(t) \subset \langle \partial f(x_t), y - x\rangle_{X^* \times X} \tag{1.7}$$

erfüllt.

Beweis. Die LIPSCHITZ-Stetigkeit folgt aus

$$|g(t_1) - g(t_2)| \leq K\|x + t_1(y-x) - x - t_2(y-x)\|_X \leq K\|y - x\|_X\,|t_1 - t_2|$$

mit LIPSCHITZ-Konstante $K_{x,y} := K\|x - y\|_X$. Ferner ist für g konstant auf $B_\varepsilon(0,1)$ fortgesetzt

$$\partial g(t) = \big\{\zeta \in \mathbb{R} \mid \zeta v \leq g^0(t;v)\ \forall\, v \in \mathbb{R}\big\}$$

eine nicht-leere, konvexe, abgeschlossene und endliche Teilmenge des $\mathbb{R}$, also ein nicht-leeres, kompaktes Intervall. Ebenso ist wegen Proposition 1.5 (iii)

$$\langle \partial f(x_t), y-x\rangle = \{\langle \eta, y-x\rangle \mid \eta \in \partial f(x_t)\}$$

ein nicht-leeres, abgeschlossenes Intervall. Damit sind die Voraussetzungen von Proposition 1.8 erfüllt, und die Behauptung folgt jetzt aus

$$\begin{aligned}
\sigma_{\partial g(t)}(v) &= g^0(t;v) = \limsup_{\substack{s\to t\\ \lambda\searrow 0}} \frac{g(s+\lambda v)-g(s)}{\lambda} \\
&= \limsup_{\substack{s\to t\\ \lambda\searrow 0}} \frac{f(x+(s+\lambda v)(y-x)) - f(x+s(y-x))}{\lambda} \\
&\leq \limsup_{\substack{y'\to x_t\\ \lambda\searrow 0}} \frac{f(y'+\lambda(y-x)v)-f(y')}{\lambda} = f^0(x_t;(y-x)v) \\
&= \max_{\zeta\in\partial f(x_t)} \langle \zeta, (y-x)v\rangle_{X^*\times X} = \max_{w\in\langle\partial f(x_t), y-x\rangle} wv \\
&= \sigma_{\langle \partial f(x_t), y-x\rangle}(v).
\end{aligned}$$

□

Beweis von Satz 1.33. Wir definieren $\theta : [0,1] \to \mathbb{R}$ durch $\theta(t) := f(x_t) + t[f(x) - f(y)]$. Offenbar ist $\theta(0) = \theta(1) = f(x)$ und θ Lipschitz-stetig. Nach dem Satz von Rolle[11] gibt es ein $\tilde{t} \in (0,1)$ mit $\theta(\tilde{t}) = \max_{[0,1]} \theta$ bzw. $\min_{[0,1]} \theta$, und es gilt

$$0 \overset{\text{Prop. 1.26}}{\in} \partial\theta(\tilde{t}) \overset{\substack{\text{Prop. 1.15(i)}\\ \text{Prop. 1.27}}}{\subset} \underbrace{\partial(f(x_{\tilde{t}}))}_{=\partial g(\tilde{t})} + [f(x)-f(y)] \overset{\text{Lemma 1.34}}{\subset} \langle \partial f(x_{\tilde{t}}), y-x\rangle + [f(x)-f(y)].$$

Das heißt, es gibt ein $\zeta \in \partial f(x_{\tilde{t}})$ mit $0 = \langle \zeta, y-x\rangle + [f(x)-f(y)]$. Mit $u := x_{\tilde{t}}$ folgt die Behauptung. □

Beispiel 1.35 Wir betrachten die Funktion $f(x) = \int_0^x \phi(s)\,\mathrm{d}s$ mit $\phi \in L^\infty(0,1)$ aus Beispiel 1.17. Wir wissen bereits $\partial f(x) = [\phi^-(x), \phi^+(x)]$. Nach dem Mittelwertsatz 1.33 gibt es ein $u \in (0,1)$ und ein $\zeta \in [\phi^-(u), \phi^+(u)]$ mit

$$\int_0^1 \phi(s)\,\mathrm{d}s = f(1) - f(0) \overset{!}{=} \langle \zeta, 1-0\rangle = \zeta.$$

Für jedes $\varepsilon > 0$ gibt es nun $x, y \in B_{\varepsilon/2}(u)$ mit

$$\phi(x) < \phi^-(u) + \varepsilon \qquad \text{und} \qquad \phi(y) > \phi^+(u) - \varepsilon,$$

also

$$\phi(x) - \varepsilon < \int_0^1 \phi(s)\,\mathrm{d}s < \phi(y) + \varepsilon.$$

◇

[11] Michel Rolle (1652 – 1719) Französischer Mathematiker. Autodidakt; ab 1685 Mitglied der *Académie Royale des Sciences*. Er befasste sich v. a. mit Algebra; den hier genannten Satz bewies er für Polynome.

Die *konvexe Hülle* einer Menge $M \subset X$ (bzw. $M \subset X^*$) ist gerade die Menge

$$\operatorname{conv} M := \{ \sum_{i=1}^n \alpha_i x_i \,|\, x_i \in M, \alpha_i \in [0,1], i = 1,\dots,n, \sum_{i=1}^n \alpha_i = 1 \}\,.$$

Mit $\overline{\operatorname{conv}}\, M$ bezeichnen wir den Abschluss der konvexen Hülle von M bzgl. der von der Norm induzierten Topologie in X (bzw. bzgl. der schwach$*$-Topologie in X^*).

Satz 1.36 [Kettenregel I]
Seien $h : X \to \mathbb{R}^n$, $g : \mathbb{R}^n \to \mathbb{R}$, $f := g \circ h$, h LIPSCHITZ-stetig nahe x, g LIPSCHITZ-stetig nahe $h(x)$ mit gemeinsamer LIPSCHITZ-Konstante K. DANN ist f LIPSCHITZ-stetig nahe x mit

$$\partial f(x) \subset \overline{\operatorname{conv}}\, \{ \sum_{i=1}^n \alpha_i \zeta_i \,|\, \zeta_i \in \partial h_i(x), \alpha = (\alpha_1,\dots,\alpha_n) \in \partial g(h(x)) \}\,. \tag{1.8}$$

Dabei gilt Gleichheit, sofern eine der folgenden Bedingungen erfüllt ist.

(i) g ist regulär bei $h(x)$, h_i ist regulär bei x, und jedes Element $\alpha \in \partial g(h(x))$ hat nicht-negative Komponenten.

(ii) $n = 1$, und g ist strikt differenzierbar bei $h(x)$.

(iii) g ist regulär bei $h(x)$ und h strikt differenzierbar bei x.

In den Fällen (ii) und (iii) ist „$\overline{\operatorname{conv}}$“ überflüssig.

Beweis von (1.8). Wir möchten Proposition 1.8 (ii) anwenden. Hierfür zeigen wir zunächst, dass die Mengen nicht-leer, schwach$*$-abgeschlossen und konvex sind. Für $\partial f(x)$ und $\partial h_i(x)$, $i = 1,\dots,n$, wissen wir dies aus Proposition 1.5. Die rechte Seite betreffend bemerken wir weiter, dass die Summe $\partial h_1(x) + \partial h_2(x)$ schwach$*$-abgeschlossen ist[12]: Wir identifizieren $v \in X$ wie üblich mit der durch $\zeta \mapsto \langle \zeta, v\rangle_{X^* \times X}$ definierten Abbildung $v \in X^{**}$ und sehen, dass $X^* \setminus (\partial h_1(x) + \partial h_2(x)) = \bigcup_{v \in X} v^{-1}(h_1^0(x;v) + h_2^0(x;v), \infty)$ offen bezüglich der schwach$*$-Topologie ist. Ferner ist $\partial h_1(x) + \partial h_2(x)$ nach dem Satz von BANACH-ALAOGLU, Satz A.12, sogar schwach$*$-kompakt, denn diese Menge ist in $B_{K_1+K_2}(0) \subset X^*$ enthalten, wobei K_1, K_2 die LIPSCHITZ-Konstanten von h_1,h_2 sein mögen. Nach Proposition 1.25 gilt $s\partial h_i(x) = \partial(sh_i)(x)$ für jedes $s \in \mathbb{R}$. Und schließlich ist mit $\{\dots\}$ auch $\overline{\operatorname{conv}}\,\{\dots\}$ schwach$*$-kompakt. Nach A.6 (iv) ist der schwach$*$-Abschluss einer konvexen Menge wieder konvex. Nun können wir Proposition 1.8 (ii) anwenden und erhalten die Inklusion (1.8) aus der Ungleichung

$$\sigma(v) := \sigma_{\{\dots\}}(v) \overset{\text{Aufg. 1.11}}{=} \sigma_{\overline{\operatorname{conv}}\,\{\dots\}}(v) \overset{!}{\geq} \sigma_{\partial f(x)}(v) = f^0(x;v).$$

Wegen Lemma 1.37 genügt es dazu,

$$\begin{aligned} f^0(x;v) - \varepsilon \quad \leq \quad & \sigma^\varepsilon(v) := \sup \big\{ \textstyle\sum_{i=1}^n \tilde{\alpha}_i \langle \zeta_i, v\rangle_{X^* \times X} \,\big| \\ & \zeta_i \in \partial h_i(x_i), \tilde{\alpha} = (\tilde{\alpha}_1,\dots,\tilde{\alpha}_n) \in \partial g(u), x_i \in B_\varepsilon(x), u \in B_\varepsilon(h(x)) \big\} \end{aligned} \tag{1.9}$$

zu zeigen. Sei nun $v \in X$ gegeben. Wir wählen $x' \in B_\varepsilon(x)$, $0 < t \ll 1$, so dass

$$x' + tv \in B_\varepsilon(x), \qquad h(x'), h(x' + tv) \in B_\varepsilon(h(x)), \qquad f^0(x;v) \leq \frac{f(x' + tv) - f(x')}{t} + \varepsilon.$$

[12]Diese Tatsache wurde auch schon im Beweis der Summenregel, Proposition 1.27 benutzt.

Nun folgern wir

$$f(x'+tv)-f(x') = g(h(x'+tv)) - g(h(x')) \overset{\text{Satz 1.33}}{=} \textstyle\sum_{i=1}^n \tilde{\alpha}_i \left(h_i(x'+tv)-h_i(x')\right)$$

mit $\tilde{\alpha} \in \partial g(u) \subset \mathbb{R}^n$ für ein $u \in (h(x'), h(x'+tv))$, was sich durch erneute Anwendung von Satz 1.33 in $\sum_{i=1}^n \tilde{\alpha}_i \left\langle \tilde{\zeta}_i, tv \right\rangle_{X^* \times X}$ umschreiben lässt, wobei $\tilde{\zeta}_i \in \partial h_i(x_i)$, $x_i \in (x', x'+tv)$. Dies beweist (1.9). □

Lemma 1.37
Unter den Voraussetzungen von Satz 1.36 erfüllt die in (1.9) *definierte Funktion* σ^ε

$$\lim_{\varepsilon \to 0} \sigma^\varepsilon(v) = \sigma(v)$$

für jedes $v \in X$.

Beweis. Fixiere ein $v \in X$. Zunächst bemerken wir $\sigma(v) \leq \sigma^\varepsilon(v)$ für alle $v \in X$. Die Behauptung folgt mithin, wenn wir zeigen:

$$\forall\, \delta > 0\ \exists\, \varepsilon_0 = \varepsilon_0(\delta, K)\ \forall\, \varepsilon \leq \varepsilon_0 : \quad \sigma^\varepsilon(v) \leq \sigma(v) + n\delta(1 + K\,\|v\|_X)$$

Wegen der Lipschitz-Stetigkeit von h können wir ε so klein wählen, dass

$$|h_i(\bar{x}) - h_i(y)| \leq K\|\bar{x} - y\|_X \quad \forall\, \bar{x}, y \in B_\varepsilon(x),$$

und wegen der Oberhalbstetigkeit der verallgemeinerten Richtungsableitung (Proposition 1.3 (v))

$$h_i^0(y; \pm v) \leq h_i^0(x; \pm v) + \frac{\delta}{K} \quad \forall\, y \in B_\varepsilon(x).$$

Multiplikation mit $|\tilde{\alpha}_i|$ liefert

$$h_i^0(y_i; \tilde{\alpha}_i v) \leq h_i^0(x; \tilde{\alpha}_i v) + \delta \quad \forall\, \tilde{\alpha} \in \partial g(u), u \in B_\varepsilon(h(x)),$$

denn $|\tilde{\alpha}_i|_{\mathbb{R}^n} \leq K$. Die Oberhalbstetigkeit des verallgemeinerten Gradienten (Proposition 1.10 (iii)) ergibt (bei eventuell nötiger Verkleinerung von ε)

$$\partial g(u) \subset B_\delta(\partial g(h(x))).$$

Damit erhalten wir

$$\begin{aligned}
&\sigma^\varepsilon(v) \\
&\overset{(1.9)}{=} \sup\Big\{ \textstyle\sum_{i=1}^n \tilde{\alpha}_i \langle \zeta_i, v\rangle_{X^* \times X} \ \Big| \\
&\qquad\qquad \zeta_i \in \partial h_i(x_i), \tilde{\alpha} = (\tilde{\alpha}_1, \dots, \tilde{\alpha}_n) \in \partial g(u), x_i \in B_\varepsilon(x), u \in B_\varepsilon(h(x)) \Big\} \\
&\leq \sup\Big\{ \textstyle\sum_{i=1}^n \underbrace{\max_{\zeta_i \in \partial h(x_i)} \tilde{\alpha}_i \langle \zeta_i, v\rangle_{X^* \times X}}_{\overset{\text{Prop. 1.5}}{=} h^0(x_i; \tilde{\alpha}_i v) \leq h^0(x; \tilde{\alpha}_i v) + \delta} \ \Big|\ x_i \in B_\varepsilon(x), \tilde{\alpha} \in \partial g(u), u \in B_\varepsilon(h(x)) \Big\} \\
&\leq \sup\Big\{ \textstyle\sum_{i=1}^n h^0(x; \tilde{\tilde{\alpha}}_i v) \ \Big|\ \tilde{\tilde{\alpha}} \in B_\delta(\partial g(h(x))) \Big\} + n\delta
\end{aligned}$$

Für die Abschätzung von $\tilde{\tilde{\alpha}} \in B_\delta(\partial g(h(x)))$ können wir jetzt die starke Topologie im $\mathbb{R}^n$ verwenden. Hierzu schreiben wir $\tilde{\tilde{\alpha}} = \alpha + \beta$ mit $\alpha \in \partial g(h(x))$, $|\beta|_{\mathbb{R}^n} \leq \delta$, also $\left|\beta_i \langle \zeta_i, v\rangle_{X^*\times X}\right| \leq |\beta| \, \|\zeta_i\|_{X^*} \|v\|_X \leq \delta K \|v\|_X$. Damit folgt

$$\begin{aligned}
\sigma^\varepsilon(v) &\leq \sup\left\{ \textstyle\sum_{i=1}^n \max_{\zeta_i \in \partial h(x)} \left((\alpha_i + \beta_i)\langle \zeta_i, v\rangle_{X^*\times X}\right) \,\middle|\, \alpha \in \partial g(h(x)), |\beta| \leq \delta\right\} + n\delta \\
&\leq \textstyle\sum_{i=1}^n \max_{\zeta_i \in \partial h(x), \alpha \in \partial g(h(x))} \alpha_i \langle \zeta_i, v\rangle_{X^*\times X} + n\delta K \|v\|_X + n\delta \\
&= \sigma(v) + n\delta(1 + K\|v\|_X)
\end{aligned}$$

□

Beweis von Satz 1.36 (i). Wir haben

$$\sigma(v) := \sigma_{\{\dots\}}(v) \stackrel{\text{Aufg. 1.11}}{=} \sigma_{\overline{\operatorname{conv}}\{\dots\}}(v) \stackrel{!}{\leq} \sigma_{\partial f(x)}(v) = f^0(x;v)$$

zu zeigen. Nun gilt

$$\begin{aligned}
\sigma(v) &= \max_{\zeta_i \in \partial h_i(x), \alpha \in \partial g(h(x))} \textstyle\sum_{i=1}^n \alpha_i \langle \zeta_i, v\rangle_{X^*\times X} \\
&\stackrel{\alpha_i \geq 0}{=} \max_{\alpha \in \partial g(h(x))} \textstyle\sum_{i=1}^n \alpha_i \underbrace{\max_{\zeta_i \in \partial h_i(x)} \langle \zeta_i, v\rangle_{X^*\times X}}_{= h_i^0(x;v) \stackrel{h_i \text{ reg.}}{=} h_i'(x;v)} \\
&= \max_{\alpha \in \partial g(h(x))} \textstyle\sum_{i=1}^n \alpha_i h_i'(x;v) \\
&\stackrel{\text{Prop. 1.5 (v)}}{=} g^0(h(x); h'(x;v)) \\
&\stackrel{g \text{ reg.}}{=} g'(h(x); h'(x;v)) \\
&= \lim_{t \searrow 0} \frac{g(h(x) + th'(x;v)) - g(h(x))}{t} \\
&= \lim_{t \searrow 0} \Big(\frac{g(h(x+tv)) - g(h(x))}{t} + \underbrace{\frac{g(h(x) + th'(x;v)) - g(h(x+tv))}{t}}_{= o(t)} \Big) \\
&= f'(x;v) \stackrel{\text{Bem. nach Def. 1.11}}{\leq} f^0(x;v) \quad \Big(\leq \sigma(v)\Big),
\end{aligned}$$

wobei wir $\left|\frac{g(h(x)+th'(x;v))-g(h(x+tv))}{t}\right| \leq K \left|\frac{h(x)+th'(x;v)-h(x+tv)}{t}\right|$, K LIPSCHITZ-Konstante von g, und $\frac{h(x+tv)-h(x)}{t} \xrightarrow{t \searrow 0} h'(x;v)$ verwandt haben. □

Beweis von Satz 1.36 (ii). g ist strikt differenzierbar bei $h(x)$ und $n = 1$, also haben wir $\partial g(h((x)) = \{D_s g(h(x))\} \subset \mathbb{R}$. Wir können ohne Einschränkung annehmen, dass $D_s g(h(x)) \geq 0$, anderenfalls betrachte man $-g$. Wie in (i) haben wir $\sigma(v) \leq f^0(x;v)$ zu zeigen.

$$\begin{aligned}
\sigma(v) &\overset{\text{Aufg. 1.11}}{=} \max\left\{ \langle \alpha\zeta, v\rangle_{X^*\times X} \,\middle|\, \zeta \in \partial h(x), \alpha \in \partial g(h(x))\right\} \\
&= \max_{\zeta\in\partial h(x)} D_s g(h(x))\, \langle \zeta, v\rangle_{X^*\times X} \\
&= D_s g(h(x)) h^0(x;v) \\
&= \limsup_{\substack{y\to x \\ t\searrow 0}} D_s g(h(x)) \frac{h(y+tv)-h(y)}{t} \\
&\overset{!}{\leq} f^0(x;v)
\end{aligned}$$

Es bleibt

$$\limsup_{\substack{y\to x \\ t\searrow 0}} \frac{(g\circ h)(y+tv)-(g\circ h)(y)}{t} \geq \limsup_{\substack{y\to x \\ t\searrow 0}} D_s g(h(x)) \frac{h(y+tv)-h(y)}{t}$$

nachzuweisen. Wähle Folgen $y_i \to x$ und $t_i \to 0$ mit

$$\lim_{i\to\infty} \frac{h(y_i+t_i v)-h(y_i)}{t_i} = \limsup_{\substack{y\to x \\ t\searrow 0}} \frac{h(y+tv)-h(y)}{t}.$$

Wenn h in einer Umgebung von x konstant ist, dann ist nichts zu zeigen. Anderenfalls schreiben wir

$$\begin{aligned}
f^0(x;v) &\geq \limsup_{i\to\infty} \frac{g(h(y_i+t_i v))-g(h(y_i))}{t_i} \\
&= \limsup_{i\to\infty} \frac{g(h(y_i+t_i v))-g(h(y_i))}{h(y_i+t_i v)-h(y_i)} \cdot \frac{h(y_i+t_i v)-h(y_i)}{t_i}.
\end{aligned}$$

Für $i\to\infty$ konvergiert der erste Faktor wegen der strikten Differenzierbarkeit von g bei $h(x)$ gegen $D_s g(h(x))$, der zweite gegen $\limsup_{\substack{y\to x \\ t\searrow 0}} \frac{h(y+tv)-h(y)}{t}$. Das war zu zeigen. □

Der Beweis von Satz 1.36 (iii) wird als Aufgabe 1.12 gestellt.

Satz 1.38 [Kettenregel II]
Seien X, Y BANACHräume, $F : X \to Y$ strikt differenzierbar bei $x \in X$, $g : Y \to \mathbb{R}$ LIPSCHITZ-stetig nahe $F(x)$ mit gemeinsamer LIPSCHITZ-Konstante K, $f := g \circ F$. DANN gilt

$$\partial f(x) \subset \partial g(F(x)) \circ D_s F(x),$$

und es herrscht Gleichheit, sofern eine der folgenden Bedingungen erfüllt ist.

(i) g ist regulär. ($\Rightarrow f$ regulär)

(ii) Für jede Umgebung U von x gibt es eine Umgebung V von $F(x)$, so dass $F(U) \cap V$ dicht in V liegt.

Bemerkung. (ii) ist z. B. erfüllt, wenn $D_s F(x) \in L(X,Y)$ surjektiv ist, vgl. Aufgabe 1.14.

Beweisskizze. Wie im Fall von Satz 1.36 zeigt man zunächst, dass $\partial f(x)$ und $\partial g(F(x)) \circ D_sF(x)$ schwach$*$-kompakt und konvex sind, um dann unter Verwendung von Proposition 1.8 (ii)

$$\begin{aligned} f^0(x;v) &\overset{!}{\leq} \max\left\{ \left\langle \zeta, \langle D_sF(x), v\rangle_{L(X,Y)\times X} \right\rangle_{Y^*\times Y} \,\middle|\, \zeta \in \partial g(F(x)) \right\} \\ &= g^0(F(x); \langle D_sF(x), v\rangle_{L(X,Y)\times X}) \end{aligned}$$

nachzuweisen. Nach dem Mittelwertsatz 1.33 ist hierfür

$$f^0(x;v) - \varepsilon \leq \sup\left\{ \left\langle \zeta, \langle D_sF(x), v\rangle_{L(X,Y)\times X} \right\rangle_{Y^*\times Y} \,\middle|\, \zeta \in \partial g(u), u \in B_\varepsilon(F(x)) \right\} \quad \forall\, \varepsilon > 0$$

hinreichend, denn die rechte Seite konvergiert mit $\varepsilon \searrow 0$ wie in Lemma 1.37 gegen $g^0(F(x); \langle D_sF(x), v\rangle_{L(X,Y)\times X})$.

(i) Zu zeigen ist $g^0(F(x); \langle D_sF(x), v\rangle_{L(X,Y)\times X}) \leq f^0(x;v)$. Wenn g regulär ist, erhalten wir

$$\begin{aligned} & g^0(F(x); \langle D_sF(x), v\rangle_{L(X,Y)\times X}) \\ = \;& g'(F(x); \langle D_sF(x), v\rangle_{L(X,Y)\times X}) = \lim_{t\searrow 0} \frac{g(F(x) + t\,\langle D_sF(x), v\rangle_{L(X,Y)\times X}) - g(F(x))}{t} \\ = \;& \lim_{t\searrow 0} \Big(\frac{g(F(x+tv)) - g(F(x))}{t} + \underbrace{\frac{g(F(x) + t\,\langle D_sF(x), v\rangle_{L(X,Y)\times X}) - g(F(x+tv))}{t}}_{=o(t)} \Big) \\ = \;& \lim_{t\searrow 0} \frac{f(x+tv) - f(x)}{t} \leq f^0(x;v), \end{aligned}$$

wobei wir

$$\begin{aligned} & \left| \frac{g(F(x) + t\,\langle D_sF(x), v\rangle_{L(X,Y)\times X}) - g(F(x+tv))}{t} \right| \\ \leq \;& K \, \Big\| \frac{F(x) - F(x+tv)}{t} + \langle D_sF(x), v\rangle_{L(X,Y)\times X} \Big\|_Y \xrightarrow{t\searrow 0} 0, \end{aligned} \tag{1.10}$$

K die Lipschitz-Konstante von g, ausgenutzt haben.

(ii) Zur Vorbereitung wählen wir $\varepsilon > 0$ und eine Umgebung V von $F(x)$, so dass $F(B_\varepsilon(x)) \cap V$ dicht in V liegt, sowie Folgen $y_i \to F(x)$, $t_i \to 0$ mit

$$\lim_{i\to\infty} \frac{g(y_i + t_i\,\langle D_sF(x), v\rangle_{L(X,Y)\times X}) - g(y_i)}{t_i} = \limsup_{\substack{y\to F(x)\\ t\searrow 0}} \frac{g(y + t\,\langle D_sF(x), v\rangle_{L(X,Y)\times X}) - g(y)}{t}.$$

Für $i \gg 1$ haben wir $y_i \in V$ und können eine Folge $(x_i')_{i\in\mathbb{N}} \subset B_\varepsilon(x)$ mit $\|F(x_i') - y_i\|_Y \leq \frac{t_i}{i}$ konstruieren. Damit erhalten wir

$$\begin{aligned} & \left| \frac{g(y_i + t_i\,\langle D_sF(x), v\rangle_{L(X,Y)\times X}) - g(y_i)}{t_i} - \frac{g(F(x_i') + t_i\,\langle D_sF(x), v\rangle_{L(X,Y)\times X}) - g(F(x_i'))}{t_i} \right| \\ \leq \;& \frac{2K}{t_i} \|F(x_i') - y_i\|_Y \leq \frac{2K}{i} \xrightarrow{i\to\infty} 0. \end{aligned} \tag{1.11}$$

Nun erhalten wir

$$\begin{aligned}
g^0(F(x); \langle D_s F(x), v\rangle_{L(X,Y)\times X}) &= \limsup_{\substack{y\to F(x)\\ t\searrow 0}} \frac{g(y + t\,\langle D_s F(x), v\rangle_{L(X,Y)\times X}) - g(y)}{t} \\
&\overset{(1.11)}{\leq} \limsup_{\substack{x'\to x\\ t\searrow 0}} \frac{g(F(x') + t\,\langle D_s F(x), v\rangle_{L(X,Y)\times X}) - g(F(x'))}{t} \\
&\overset{(1.10)}{\leq} \limsup_{\substack{x'\to x\\ t\searrow 0}} \frac{g(F(x'+tv)) - g(F(x'))}{t} + o(t) \\
&= f^0(x; v),
\end{aligned}$$

was zu zeigen war. □

Eine direkte Anwendung von Satz 1.38 (ii) ist das folgende

Corollar 1.39
Sei Y eine dichte Teilmenge von U, $g : U \to \mathbb{R}$ LIPSCHITZ-stetig nahe $x \in Y$.
DANN ist $f := g|_Y$ LIPSCHITZ-stetig nahe x und $\partial f(x) = \partial g(x)$, d.h.

$$\forall\, \zeta \in \partial f(x)\ \exists\, \bar\zeta \in \partial g(x)\ \forall\, v \in Y\ :\quad \langle \zeta, v\rangle_{Y^*\times Y} = \left\langle \bar\zeta|_Y, v\right\rangle_{Y^*\times Y}.$$

Proposition 1.40 [Maximalfunktionen über endlicher Indexmenge]
Sei $x \in U$, $f_i : U \to \mathbb{R}$, $i = 1,\dots,n$, LIPSCHITZ-stetig nahe x. DANN ist $f(x') := \max_{i\in\{1,\dots,n\}} f_i(x')$ LIPSCHITZ-stetig nahe x und

$$\partial f(x) \subset \operatorname{conv}\ \{\partial f_i(x) \,|\, i \in I(x)\},$$

wobei $I(x) := \{i \in \{1,\dots,n\} \,|\, f(x) = f_i(x)\}$. Gleichheit tritt ein, falls für alle $i \in I(x)$ die Funktionen f_i regulär bei x sind.

Beweis. Wir erinnern an Beispiel 1.23 und definieren $g : \mathbb{R}^n \to \mathbb{R}$, $h : U \to \mathbb{R}^n$ durch

$$g(y) := \max_{i\in\{1,\dots,n\}} y_i \qquad h(x) := \begin{pmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{pmatrix}.$$

Satz 1.36 liefert für $f = g \circ h$

$$\begin{aligned}
\partial f(x) &\subset \overline{\operatorname{conv}}\ \{\textstyle\sum_{i=1}^n \alpha_i\zeta_i \,|\, \zeta_i \in \partial h_i(x), \alpha = (\alpha_1,\dots,\alpha_n) \in \partial g(h(x))\} \\
&\overset{\text{Bsp. 1.23}}{=} \Big\{ \textstyle\sum_{i=1}^n \alpha_i\zeta_i \Big| \zeta_i \in \underbrace{\partial h_i(x)}_{=\partial f_i(x)}, \sum_{i=1}^n \alpha_i = 1, \alpha_i \geq 0, \alpha_j = 0 \text{ für } j \notin I(x) \Big\} \\
&= \operatorname{conv}\ \{\partial f_i(x) \,|\, i \in I(x)\};
\end{aligned}$$

denn Linearkombinationen schwach∗-kompakter Mengen sind wieder schwach∗-kompakt, vgl. die Argumentation am Anfang des Beweises von Satz 1.36, Seite 23. Zusätzlich ist die Menge {...} konvex.
Die Funktion g ist nach Konstruktion konvex und nach Proposition 1.32 (ii) regulär, und die h_i sind regulär, sofern die f_i regulär sind. Wegen $\alpha_i \geq 0$ ist Satz 1.36 (i) anwendbar. □

Proposition 1.41 [Produktregel]
Sei $x \in U$, $f_1, f_2 : U \to \mathbb{R}$ LIPSCHITZ-stetig nahe x.
DANN ist $f_1 f_2$ LIPSCHITZ-stetig nahe x und

$$\partial(f_1 f_2) \subset f_2(x)\partial f_1(x) + f_1(x)\partial f_2(x). \tag{1.12}$$

Falls $f_1(x), f_2(x) \geq 0$ und f_1, f_2 regulär sind, so ist $f_1 f_2$ regulär, und es gilt Gleichheit in (1.12).

Beweis. Definiere $g : \mathbb{R}^2 \to \mathbb{R}$ und $h : U \to \mathbb{R}^2$ durch $g(u) := u_1 u_2$ und $h(x) := \binom{f_1(x)}{f_2(x)}$, mithin $f := f_1 f_2 = g \circ h$. Wir haben

$$\begin{gathered} g'((u_1, u_2); v) = \nabla g(u_1, u_2) \cdot v = u_2 v_1 + u_1 v_2, \\ \partial g(u_1, u_2) = \{\nabla g(u_1, u_2)\}, \qquad \partial g(h(x)) = \left\{ \binom{f_2(x)}{f_1(x)} \right\}. \end{gathered}$$

Anwenden von Satz 1.36 liefert

$$\begin{aligned} \partial f(x) &\subset \overline{\text{conv}}\, \{\alpha_1 \zeta_1 + \alpha_2 \zeta_2 \,|\, \zeta_i \in \partial f_i(x), \alpha = (\alpha_1, \alpha_2) \in \partial g(h(x))\} \\ &= \overline{\text{conv}}\, \{ f_2(x)\zeta_1 + f_1(x)\zeta_2 \,|\, \zeta_i \in \partial f_i(x)\}, \end{aligned}$$

und da $\{ f_2(x)\zeta_1 + f_1(x)\zeta_2 \,|\, \zeta_i \in \partial f_i(x)\}$ bereits schwach$*$-abgeschlossen und konvex ist, können wir „$\overline{\text{conv}}$" streichen. Dies zeigt „$\subset$", die Gleichheit folgt mit Satz 1.36 (i). ☐

Proposition 1.42 [Quotientenregel]
Sei $x \in U$, $f_1, f_2 : U \to \mathbb{R}$ LIPSCHITZ-stetig nahe x, $f_2(x) \neq 0$.
DANN ist $\frac{f_1}{f_2}$ LIPSCHITZ-stetig nahe x und

$$\partial\left(\frac{f_1}{f_2}\right)(x) \subset \frac{f_2(x)\partial f_1(x) - f_1(x)\partial f_2(x)}{f_2^2(x)}. \tag{1.13}$$

Falls zusätzlich $f_1(x), f_2(x) \geq 0$ und $f_1, -f_2$ regulär bei x sind, so ist $\frac{f_1}{f_2}$ regulär bei x, und es gilt Gleichheit in (1.13).

Der Beweis wird als Aufgabe 1.16 gestellt.

Wir betrachten nun die Situation, in der X *cartesisches*[13] *Produkt* der BANACHräume X_1 und X_2 ist, also $X = X_1 \times X_2$. Dann gilt $X^* = X_1^* \times X_2^*$, vgl. [Alt02, Ü 4.5]. Zu einer gegebenen, nahe $x = (x_1, x_2) \in X$ LIPSCHITZ-stetigen Funktion $f : X \to \mathbb{R}$ definieren wir

$$\begin{aligned} f_1'((x_1, x_2); v) &:= (f(\cdot, x_2))'(x_1; v), \\ f_1^0((x_1, x_2); v) &:= (f(\cdot, x_2))^0(x_1; v), \\ \partial_1 f(x_1, x_2) &:= \partial(f(\cdot, x_2))(x_1), \\ \pi_1 \partial f(x_1, x_2) &:= \{\zeta \in X_1^* \,|\, \exists\, \eta \in X_2^* : (\zeta, \eta) \in \partial f(x_1, x_2)\}, \end{aligned}$$

[13] nach RENÉ DESCARTES (1596 – 1650) Französischer Philosoph, Mathematiker und Naturforscher. Seine rationalistische Philosophie prägte die wissenschaftliche Revolution des 17. Jahrhunderts maßgeblich. Weiterhin leistete DESCARTES Beiträge zur Optik und Mechanik und legte die Grundlagen der analytischen Geometrie, die er in seinem richtungsweisenden Werk *La Géometrie* zusammenfasste. Es enthält unter anderem auch den Hauptsatz der Algebra (ohne Beweis); in seinem Nachlass findet sich eine Vorform des EULERschen Polyedersatzes.

und analog $f_2'((x_1,x_2);v)$, $f_2^0((x_1,x_2);v), \partial_2 f(x_1,x_2)$ und $\pi_2\partial f(x_1,x_2)$.
Warnung: Im Allgemeinen ist keine der Mengen $\partial f(x_1,x_2)$ und $\partial_1 f(x_1,x_2)\times\partial_2 f(x_1,x_2)$ in der anderen enthalten, vgl. Beispiel 3.5 in Kapitel 3. Es gilt aber

Proposition 1.43
Sei $f : U \to \mathbb{R}$ regulär bei $x \in U$. DANN gilt

$$\partial f(x_1,x_2) \subset \partial_1 f(x_1,x_2) \times \partial_2 f(x_1,x_2).$$

Bemerkung. Für $\zeta = (\zeta_1,\zeta_2) \in X^*$, $w = (w_1,w_2) \in X$ gilt

$$\langle \zeta, w\rangle_{X^*\times X} = \langle \zeta,(w_1,0)\rangle_{X^*\times X} + \langle \zeta,(0,w_2)\rangle_{X^*\times X} = \langle \zeta_1,w_1\rangle_{X_1^*\times X_1} + \langle \zeta_2,w_2\rangle_{X_2^*\times X_2}.$$

Beweis von Proposition 1.43. Sei $\zeta = (\zeta_1,\zeta_2) \in \partial f(x_1,x_2)$. Dann folgt unter Beachtung voranstehender Bemerkung

$$\begin{aligned} f_1^0((x_1,x_2);v) &\geq f_1'((x_1,x_2);v) = f'((x_1,x_2);(v,0)) \overset{f\text{ reg.}}{=} f^0((x_1,x_2);(v,0)) \\ &\geq \big\langle \underbrace{(\zeta_1,\zeta_2)}_{=\zeta},(v,0)\big\rangle = \langle \zeta_1, v\rangle_{X_1^*\times X_1} \qquad \forall\, v \in X_1, \end{aligned}$$

und analog zeigt man $f_2^0((x_1,x_2);w) \geq \langle \zeta_2,w\rangle_{X_2^*\times X_2}$ für alle $w \in X_2$, womit $\zeta_i \in \partial_i f(x_1,x_2)$ für $i=1,2$ gezeigt ist. □

Ohne Regularitätsannahmen an f kann man immerhin das folgende Resultat zeigen:

Proposition 1.44
Sei $f : U \to \mathbb{R}$ LIPSCHITZ-stetig bei $x = (x_1,x_2)$. DANN gilt

$$\partial_1 f(x_1,x_2) \subset \pi_1 \partial f(x_1,x_2)$$

und folglich

$$\partial_1 f(x_1,x_2) \times \partial_2 f(x_1,x_2) \subset \pi_1\partial f(x_1,x_2) \times \pi_2 \partial f(x_1,x_2).$$

Beweis. Definiere $f_1 : X_1 \to \mathbb{R}$ durch $\xi \mapsto f(\xi,x_2)$ und $F : X_1 \to X_1\times X_2$ durch $\xi \mapsto (\xi,x_2)$. Nun ist $\partial_1 f(x_1,x_2) = \partial f_1(x_1)$ und $\langle D_sF(x_1),v\rangle_{L(X_1,X_1\times X_2)\times X_1} = (v,0)$. Nach Konstruktion haben wir $f_1 = f\circ F$, so dass wir mit Satz 1.38

$$\begin{aligned} \partial f_1(x_1) \quad &\subset \quad \Big\{\zeta \in X_1^* \,\Big|\, \exists\, \eta \in \partial f(x_1,x_2) : \\ &\qquad \langle \zeta, v\rangle_{X_1^*\times X_1} = \Big\langle \eta, \langle DsF(x_1),v\rangle_{L(X_1,X_1\times X_2)\times X_1}\Big\rangle_{X^*\times X} \ \forall\, v\in X_1\Big\} \\ &= \quad \Big\{\zeta \in X_1^* \,\Big|\, \exists\, \eta \in \partial f(x_1,x_2) : \langle \zeta,v\rangle_{X_1^*\times X_1} = \langle \eta,(v,0)\rangle_{X^*\times X} \ \forall\, v \in X_1\Big\} \end{aligned}$$

erhalten. Mit $\eta = (\eta_1,\eta_2)$ gilt $\langle \zeta,v\rangle_{X_1^*\times X_1} = \langle \eta_1,v\rangle_{X_1^*\times X_1}$, also $\zeta \in \pi_1\partial f(x_1,x_2)$. □

Beispiel 1.45 Gegeben seien Messwerte $(i,i) \in \mathbb{R}^2$ für $i = 0,1,\ldots,N-1$ und $(N,0)$. Wir wollen zeigen, dass $y = x$ die beste Ausgleichsgerade bezüglich der euklidischen Norm liefert, sofern $N \geq 5$.

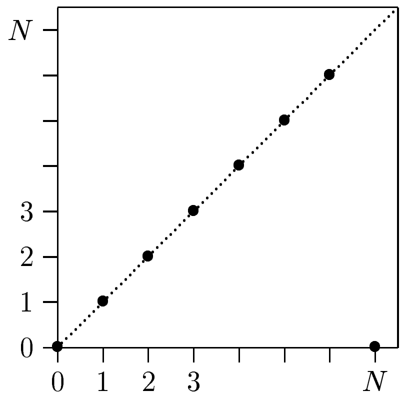

Abbildung 1.1: Messwerte und Messfehler

Wir untersuchen also das Minimierungsproblem

$$f(\alpha,\beta) := |\alpha N + \beta| + \textstyle\sum_{i=0}^{N-1} |\alpha i + \beta - i| \longrightarrow \text{min!} \quad \text{in } \mathbb{R}^2. \tag{1.14}$$

Mit den Bezeichnungen

$$F_{c,k}(\alpha,\beta) := \alpha c + \beta - k, \qquad g(x) = |x|$$

haben wir $f_{c,k}(\alpha,\beta) := g \circ F_{c,k} = |\alpha c + \beta - k|$. Da g konvex, also nach Proposition 1.32 (ii) regulär ist, folgt mit der Linearität der $F_{c,k}$ die Regularität von $f_{c,k}$.
$D_s F_{c,k}(\alpha,\beta) = \binom{c}{1}$ impliziert via Satz 1.38 (i) und Beispiel 1.6

$$\partial f_{c,k}(\alpha,\beta) = \begin{cases} \binom{c}{1} & \text{falls} \quad \alpha c + \beta - k > 0 \\ \lambda\binom{c}{1}, \lambda \in [-1,1], & \text{falls} \quad \alpha c + \beta - k = 0 \\ -\binom{c}{1} & \text{falls} \quad \alpha c + \beta - k < 0 \end{cases}$$

Wegen der Konvexität sind alle kritischen Punkte Minima. Für jeden kritischen Punkt (α,β) liefert Proposition 1.26, dass $0 \in \partial f(\alpha,\beta)$, und nach (1.14) erhalten wir mit Proposition 1.27 und Corollar 1.31

$$\begin{aligned} 0 &\in \partial f_{N,0}(\alpha,\beta) + \textstyle\sum_{i=0}^{N-1} \partial f_{i,i}(\alpha,\beta) \\ &= \lambda_N \binom{N}{1} + \textstyle\sum_{i=0}^{N-1} \lambda_i \binom{i}{1}, \qquad\qquad \lambda_j \in [-1,1], j = 0, \ldots, N \end{aligned}$$

Für $(\alpha,\beta) = (1,0)$ ist $\alpha c + \beta - k = c - k$, also können wir *a priori* keine Aussage über $\lambda_1, \ldots, \lambda_{N-1}$ treffen, allerdings ist $\lambda_N = 1$. In Aufgabe 1.17 wird gezeigt, dass das System

$$\left| \begin{array}{ccccc} 0 & = & N & + & \sum_{i=0}^{N-1} i\lambda_i \\ 0 & = & 1 & + & \sum_{i=0}^{N-1} \lambda_i \end{array} \right| \qquad \lambda_i \in [-1,1]$$

für $N \geq 3$ lösbar und

$$\lambda_N \binom{N}{1} + \textstyle\sum_{i=0}^{N-1} \lambda_i \binom{i}{1}$$

für $(\alpha,\beta) \neq (1,0)$ und $N \geq 5$ *nicht* lösbar ist, was zeigt, dass die Gerade $y = x$ (also $(\alpha,\beta) = (1,0)$) eine Lösung für alle $N \geq 3$ und für $N \geq 5$ die einzige Lösung ist. $\diamond$

Aufgaben

1.1 Sei $f : X \to \mathbb{R}$ LIPSCHITZ-stetig nahe $x \in X$, X ein BANACHraum. Beweisen Sie mit dem Satz A.11 von HAHN-BANACH,

(a) dass $\partial f(x) \neq \emptyset$.

(b) dass es zu $v \in X, v \neq 0$ ein $\xi \in X^*$ gibt, so dass

$$\langle \xi, v\rangle = f^0(x;v) \qquad \text{und} \qquad \langle \xi, w\rangle \leq f^0(x;w) \text{ für alle } w \in X.$$

1.2 Berechnen Sie $\partial f(x)$ für alle $x \in X$ für:

(a) $X = \mathbb{R}, \quad f(x) := 2x + |x| + x^2$

(b) $X = \mathbb{R}^2, \quad f(x) := |x| + |x|^2$

1.3 (i) Für $\dim X < \infty$ stimmen HADAMARD- und FRÉCHET-Ableitung überein.

(ii) Ist eine Funktion LIPSCHITZ-stetig nahe eines Punktes x, so ist in x jede HADAMARD- auch eine GÂTEAUX-Ableitung.

1.4 Berechnen Sie für alle $x \in \mathbb{R}$ den verallgemeinerten Gradienten $\partial f(x)$ der Funktion $f(x) := x^2 \sin \frac{1}{x}$, $x \in \mathbb{R}\backslash\{0\}$, $f(0) := 0$, und zeigen Sie, dass $\zeta \in \partial f(0)$ existieren mit $\zeta \neq f'(0)$. Warum ist das kein Widerspruch zu Proposition 1.15 (i) aus der Vorlesung?

1.5 [Beispiel 1.17] (a) Nennen Sie ein Beispiel für eine Funktion $\phi \in L^\infty([0,1])$ mit $\phi^-(x) < \phi^+(x)$ für ein $x \in (0,1)$.

(b) Zeigen Sie für $f(x) := \int_0^x \phi(t)\,dt$, wobei $\phi \in L^\infty([0,1])$ beliebig ist, dass $f^0(x;1) \leq \phi^+(x)$ und damit $\zeta \leq \phi^+(x)$ für alle $\zeta \in \partial f(x)$, $x \in (0,1)$. Zeigen Sie analog, dass auch $\zeta \geq \phi^-(x)$ für alle $\zeta \in \partial f(x)$, $x \in (0,1)$.

1.6 [Beispiel 1.17] (von U. Menne) Sei $\phi \in L^1(\mathbb{R}^n)$.

(a) Zeigen Sie $\phi^+(x) = \lim_{\varepsilon \searrow 0} \operatorname{ess\,sup} \phi|_{B_\varepsilon(x)}$ und $\phi^-(x) = \lim_{\varepsilon \searrow 0} \operatorname{ess\,inf} \phi|_{B_\varepsilon(x)}$.

(b) Zeigen Sie $\phi^- \leq \phi \leq \phi^+$ fast überall.

(c) Gilt $\phi^- = \phi^+$ fast überall?

Lösung. Zur Definition des essentiellen Supremums vgl. [Alt02], p. 47.

(a) Wir zeigen die Behauptung für ϕ^+, der Fall ϕ^- verläuft analog.
„$\leq$": Zu jedem $t_i := \frac{1}{i}$ gibt es eine Nullmenge N_i mit $\phi \leq \operatorname{ess\,sup} \phi|_{B_{t_i}(x)}$ für alle $x \in B_{t_i}(x) \setminus N_i$. Setze $N := \bigcup_{i\in\mathbb{N}} N_i$, dann ist $\phi^+(x) \leq \sup_{\substack{(x_i)_{i\in\mathbb{N}}\\ \subset \mathbb{R}^n\setminus N}} \limsup_{x_i \to x} \phi(x_i) \leq \operatorname{ess\,sup} \phi|_{B_{t_i}(x)}$.
„$\geq$": Sei N eine Nullmenge. Dann haben wir Folgen $x_i \to x$, $x_i \in \mathbb{R}^n \setminus N$ mit $\limsup_{i\to\infty} \phi(t_i) = \lim_{t\searrow 0} \operatorname{ess\,sup} \phi_{B_t(x)} =: \psi^+(x)$. Nun ist $\sup_{\substack{(x_i)_{i\in\mathbb{N}}\\ \subset \mathbb{R}^n\setminus N}} \limsup_{x_i \to x} \phi(x_i) \geq \psi^+(x)$. Bilden von $\inf_{\substack{N\subset\mathbb{R}^n\\ |N|=0}}$ liefert $\phi^+ \geq \psi^+$.

(b) Sei $\mathscr{D} = \{x_1, x_2, x_3, \ldots\}$ dicht in $\mathbb{R}^n$ und seien $g_{i,k} : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ definiert durch

$$g_{i,k}(x) := \begin{cases} \operatorname{ess\,sup} \phi|_{B_{1/k}(x_l)} & x \in B_{1/k}(x_l) \\ +\infty & \text{sonst} \end{cases} \qquad \text{für alle } x \in \mathbb{R}^n, k, l \in \mathbb{N}.$$

Mit $g := \inf_{k,l\in\mathbb{N}} g_{k,l}$ gilt dann $g \geq \phi$ fast überall und $\phi^+ = g$ (überall).

(c) Wir konstruieren ein Gegenbeispiel: Seien $I := [0,1]$, $\mathscr{D} = \{x_1, x_2, x_3, \ldots\}$ dicht in I, $I_k := B_{2^{-k-2}}(x_k)$, $k \in \mathbb{N}$, $J := \bigcup_{k\in\mathbb{N}} I_k$, $K := I \setminus J$. Es gilt $L^1(J) \leq \frac{1}{2}$. Sei nun $x \in K$. Nach Konstruktion haben wir $L^1(J \cap B_\varepsilon(x)) > 0$ für alle $\varepsilon > 0$. Wenn auch $L^1(K \cap B_\varepsilon(x)) > 0$ für alle $\varepsilon > 0$, so folgt für $\phi := \chi_K$ die Ungleichung $\phi^+(x) = 1 > 0 = \phi^-(x)$. Zu zeigen bleibt, dass $L^1(K \cap B_\varepsilon(x)) > 0$ für alle $\varepsilon > 0$ und fast alle $x \in K$ gilt. In der Tat, sei $x \in K$ ein L^1-LEBESGUE-Punkt von χ_K, dann ist $L^1(K \cap B_\varepsilon(x)) = \frac{2\varepsilon}{\mathscr{L}^1(B_\varepsilon(x))} \int_{B_\varepsilon(x)} \chi_K \, d\mathscr{L}^1 \to 2\varepsilon\chi_K(x) = 2\varepsilon > 0$.

□

1.7 Beweisen Sie Corollar 1.20.
Hinweis: Nutzen Sie die Methode aus der Vorlesung Variationsrechnung I [VdM03], Kapitel 7.2, wo wir bei der Behandlung eines Hindernisproblems Ähnliches für eine konkave Funktion bewiesen haben.

1.8 Seien U, X BANACHräume, $\phi : U \times X \to \mathbb{R} \cup \{+\infty\}$ konvex, und sei $f(u) := \inf_{x \in X} \phi(u, x)$ für $u \in U$. Zeigen Sie: Falls f auf einer konvexen Menge S endlich ist, dann ist f konvex auf S.

1.9 Berechnen Sie $\partial f(x)$ der TSCHEBYSCHEW-Norm

$$f(x) := \max_{i \in \{1,\dots,n\}} |x_i| \quad \text{für alle} \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

1.10 Seien $x, y \in X$, X BANACHraum, und f LIPSCHITZ-stetig auf einer offenen Menge $N \subset X$ mit $[x, y] \subset N$. Beweisen Sie, dass für ein beliebiges $t \in [0, 1]$ die Menge

$$\langle \partial f(x + t(y - x)), y - x \rangle_{X^* \times X} \subset \mathbb{R}$$

ein abgeschlossenes endliches Intervall ist.

1.11 Sei $K \subset X^*$, $K \neq \emptyset$, X^* der Dualraum eines BANACHraumes X. Zeigen Sie für die Stützfunktionen die Identität

$$\sigma_K(v) = \sigma_{\overline{\operatorname{conv}} K}(v) \quad \text{für alle} \quad v \in X.$$

1.12 [Kettenregel 1.36 (iii)] Sei X ein BANACHraum und $h : X \to \mathbb{R}^n$ strikt differenzierbar bei $x \in X$, und $g : \mathbb{R}^n \to \mathbb{R}$ regulär bei $h(x)$. Zeigen Sie, dass dann $f := g \circ h$ regulär bei x ist, und dass

$$\partial f(x) = \left\{ \textstyle\sum_{i=1}^n \alpha_i \zeta_i : \zeta_i \in \partial h_i(x), i = 1, \dots, n, \alpha = (\alpha_1, \dots, \alpha_n) \in \partial g(h(x)) \right\}.$$

(Beachten Sie, dass auf der rechten Seite der Kettenregel hier nicht der schwach∗-Abschluss genommen werden muss.)
Hinweis: Nutzen Sie die Inklusion der allgemeinen Kettenregel, Satz 1.36 der Vorlesung, und gehen Sie dann so vor wie bei der Gleichheitsdiskussion zum Teil (i) dieses Satzes.

1.13 [Kettenregel 1.38] Sei $F : X \to Y$ strikt differenzierbar in $x \in X$ und $g : Y \to \mathbb{R}$ LIPSCHITZ-stetig nahe $F(x) \in Y$, wobei X, Y BANACHräume sind. Zeigen Sie, dass $f := g \circ F : X \to \mathbb{R}$ LIPSCHITZ-stetig nahe x ist, und dass

$$\partial f(x) \subset \partial g(F(x)) \circ D_s F(x)$$

gilt.
Hinweis: Vervollständigen Sie die in der Vorlesung angegebene Beweisskizze von Satz 1.38 anhand des Beweises von Satz 1.36.

1.14 Sei $F : X \to Y$ strikt differenzierbar bei x mit einer *surjektiven* strikten Ableitung $D_s F(x) \in L(X, Y)$; sei $X = A \times B$ mit A BANACHraum, $B = \operatorname{Kern} F'(x)$. Zeigen Sie, dass das Bild jeder offenen Umgebung von x unter F dicht in einer Umgebung von $F(x)$ liegt. Damit erfüllt F eine der Bedingungen, die in Satz 1.38 (Kettenregel II) zur Gleichheit in der Mengeninklusion führen.
Hinweis: Zeigen Sie mit dem Satz über implizite Funktionen in BANACH*räumen, dass f lokal surjektiv ist.*

1.15 Seien X ein BANACHraum und $f_i : X \to \mathbb{R}$ LIPSCHITZ-stetig nahe $x \in X$ für alle $i = 1, \dots, n$ und

$$f(x') := \max\{f_i(x') : i = 1, \dots, n\}, \qquad x' \in X.$$

Zeigen Sie: f ist LIPSCHITZ-stetig nahe x.

1.16 [Quotientenregel] Beweisen Sie Proposition 1.42!

1.17 Beweisen Sie:

(a) Für $N \geq 3$ gibt es Zahlen $\lambda_0, \lambda_1, \dots, \lambda_{N-1} \in [-1, 1]$, so dass

$$0 = N + \sum_{i=0}^{N-1} i\lambda_i, \qquad \text{und} \qquad 0 = 1 + \sum_{i=0}^{N-1} \lambda_i \tag{1.15}$$

gilt.
Bemerkung: Das beweist, dass $0 \in \partial f(1, 0)$ in Beispiel 1.45 aus der Vorlesung, d.h., dass die Gerade $y = x$ tatsächlich eine fehlerminimierende Ausgleichsgerade für die gegebenen Datenpunkte ist.

(b) Sei $N > 4$, $f_{c,k} : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch

$$f_{c,k}(\alpha, \beta) := |\alpha c + \beta - k|.$$

Zeigen Sie die Äquivalenz

$$0 \in \partial f_{N,0}(\alpha, \beta) + \sum_{i=1}^{N-1} \partial f_{i,i}(\alpha, \beta) \qquad \Longleftrightarrow \qquad \alpha = 1, \beta = 0.$$

2 Geometrie

Generalvoraussetzung für Kapitel 2

Im Folgenden sei X immer ein BANACHraum über $\mathbb{R}$ mit der Norm $\|\cdot\|_X$.
$C \subset X$ sei eine nicht-leere Teilmenge.

Definition 2.1 [Distanz- und Indikatorfunktion, Polarisierung]

- Die *Distanzfunktion* $d_C : X \to \mathbb{R}^+$ ist definiert durch $d_C(x) := \inf_{c\in C} \|x - c\|_X$.
- Für eine Menge M sei die *Indikatorfunktion* durch

  $$i_M(v) := \begin{cases} 0 & : \ v \in M \\ +\infty & : \ v \notin M \end{cases}$$

  erklärt. Dann wird für $M \subset X$

  $$\begin{aligned} \mathrm{Polar}(M) \ &:= \ \{\zeta \in X^* \mid \langle \zeta, v\rangle \le i_M(v) \ \forall\, v \in X\} \\ &= \ \{\zeta \in X^* \mid \langle \zeta, v\rangle \le 0 \ \forall\, v \in M\} \end{aligned}$$

  Polarisierung von M genannt.

Proposition 2.2 *(i) Falls C abgeschlossen ist, dann gilt $d_C(x) = 0 \iff x \in C$.*

(ii) d_C ist LIPSCHITZ-*stetig mit* LIPSCHITZ-*Konstante* 1.

(iii) Falls C konvex ist, dann ist d_C eine konvexe Funktion.

Der Beweis dieser Proposition wird als Aufgabe 2.1 gestellt.

Definition 2.3 [Tangenten und Normalen]
Sei $x \in C$.

(i) $v \in X$ heißt *tangential an C in x*, wenn $d_C^0(x; v) = 0$. Wir setzen

$$T_C(x) := \left\{ v \in X \mid d_C^0(x; v) = 0 \right\}$$

und nennen $T_C(x)$ den *Tangentialkegel an C in x.*

(ii) $\zeta \in X^*$ heißt *normal an C in x*, wenn $\langle \zeta, v\rangle \le 0$ für alle $v \in T_C(x)$. Wir setzen

$$N_C(x) := \{\zeta \in X^* \mid \langle \zeta, v\rangle \le 0 \ \forall\, v \in T_C(x)\}$$

und nennen $N_C(x)$ den *Normalkegel an C in x.*

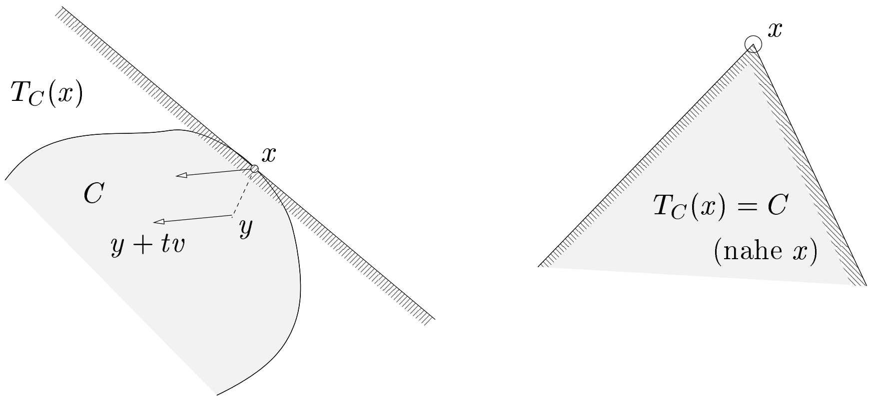

Abbildung 2.1: Tangentialkegel

Bemerkung.

(i) In $x \in C$ gilt für jedes $v \in X$

$$d_C^0(x;v) = \limsup_{\substack{y \to x \\ t \searrow 0}} \frac{d_C(y+tv) - d_C(y)}{t} \geq \limsup_{t \searrow 0} \frac{d_C(x+tv) - \overbrace{d_C(x)}^{=0}}{t} \geq 0,$$

d. h. in der Definition von $T_C(x)$ ist nur „$\leq$“ wesentlich.

(ii) Es gilt $A \subset B \Longrightarrow \mathrm{Polar}(A) \supset \mathrm{Polar}(B)$. $N_C(x)$ ist die Polarisierung von $T_C(x)$.

(iii) $T_C(x)$ ist ein konvexer, abgeschlossener (positiver) Kegel, $N_C(x)$ ein konvexer, schwach$*$-abgeschlossener (positiver) Kegel, vgl. Aufgabe 2.2 (a).

Proposition 2.4
Für $x \in C$ gilt $N_C(x) = \overline{\bigcup_{\lambda \geq 0} \lambda \partial d_C(x)}^*$.

Beweis. „$\subset$“ wird in Aufgabe 2.3 gezeigt.
„$\supset$“: Sei $\zeta \in \partial d_C(x) = \{\eta \in X^* \mid \langle \eta, v \rangle \leq d_C^0(x;v) \; \forall \, v \in X\}$, dann gilt $\langle \zeta, v \rangle \leq 0$ für alle $v \in T_C(x)$, also $\partial d_C(x) \subset N_C(x)$. Aufgabe 2.2 (a) liefert $\overline{\bigcup_{\lambda \geq 0} \lambda \partial d_C(x)}^* \subset N_C(x)$. □

Proposition 2.5
Für $C \subset S \subset X$ sei $f : S \to \mathbb{R}$ Lipschitz-stetig mit Lipschitz-Konstante K, und es gebe ein $x \in C$ mit $f(x) = \min_C f$.
DANN gilt für alle $L \geq K$ und die Funktion $g_L := f + L d_C$

$$g_L(x) = \min_S g_L.$$

Ist C zusätzlich noch abgeschlossen und $L > K$, so liegt jedes $y \in S$, das g_L auf S minimiert, sogar in C.

Beweis. Sei $L \geq K$. Wir nehmen an, es gäbe $y \in S$ und $\varepsilon > 0$ mit $g_L(y) < g_L(x) - L\varepsilon$. Wir wählen ein $c \in C$ mit $\|y - c\|_X \leq d_C(y) + \varepsilon$. Jetzt gilt

$$\begin{aligned} f(c) &\leq f(y) + K\,\|y - c\|_X \leq f(y) + L(d_C(y) + \varepsilon) = g_L(y) + L\varepsilon \\ &< g_L(x) = f(x) + L\,\underbrace{d_C(x)}_{=0}. \quad \lightning \end{aligned}$$

Sei nun $y \in S$ Minimierer und $L > K$. Jetzt haben wir

$$\begin{aligned} g_L(y) &= f(y) + Ld_C(y) \leq f(x) + L\,\underbrace{d_C(x)}_{=0} = f(x) + Kd_C(x) \\ &= g_K(x) \overset{\text{s.o.}}{\leq} g_K(y) = f(y) + Kd_C(y) \end{aligned}$$

Daraus folgt $d_C(y) = 0$ und mit Proposition 2.2 (i) schließlich $y \in C$. □

Corollar 2.6
Sei $C \subset S \subset X$, $f : S \to \mathbb{R}$ Lipschitz-stetig nahe $x \in C \cap \mathring{S}$, und es gelte $f(x) = \min_C f$. DANN ist $0_{X^*} \in \partial f(x) + N_C(x)$.

Beweis. Proposition 2.5 liefert $g_K(x) = \min_S g_K$, wobei K die Lipschitz-Konstante von f nahe x bezeichnet. Propositionen 1.25, 1.26, 1.27 zusammen mit Proposition 2.4 implizieren

$$0 \in \partial g_K(x) \subset \partial f(x) + \underbrace{K(\partial d_C(x))}_{\subset \bigcup_{\lambda \geq 0} \lambda \partial d_C(x) \subset N_C(x)} \subset \partial f(x) + N_C(x).$$

□

Definition 2.7
Sei C konvex, $x \in C$. $\zeta \in X^*$ heißt *normal zu C in x*, wenn $\langle \zeta, x - c \rangle \geq 0$ für alle $c \in C$ ist. Die Menge

$$\tilde{N}_C(x) := \{\zeta \in X^* \mid \langle \zeta, x - c \rangle \geq 0 \; \forall\, c \in C\}$$

nennt man den *Normalkegel von C in x* (im Sinne der konvexen Analysis).

Proposition 2.8
Sei C konvex, $x \in C$. DANN ist $\tilde{N}_C(x) = N_C(x)$.

Beweis. „$\subset$“: Sei $\zeta \in \tilde{N}_C(x)$, betrachte $f(y) := \langle \zeta, x - y \rangle$. Wir haben $f|_C \geq 0$, und $f(x) = 0$ liefert $f(x) = \min_C f$. Nach Corollar 2.6 erhalten wir $0 \in \partial f(x) + N_C(x) = \{-\zeta\} + N_C(x)$ und hieraus $\zeta \in N_C(x)$.
„$\supset$“: Wir betrachten zunächst ein $\zeta \in \partial d_C(x)$. Da d_C nach Proposition 2.2 (iii) konvex ist, gilt $\partial d_C(x) = \operatorname{Sub} d_C(x)$ nach Proposition 1.22, d.h. nach Definition 1.21

$$\begin{aligned} d_C(y) - d_C(x) &\geq \langle \zeta, y - x \rangle \;\; \forall\, y \in X \\ 0 = \underbrace{d_C(c)}_{=0} - \underbrace{d_C(x)}_{=0} &\geq \langle \zeta, c - x \rangle \;\; \forall\, c \in C, \end{aligned}$$

mithin $\zeta \in \tilde{N}_C(x)$, also $\partial d_C(x) \subset \tilde{N}_C(x)$. Für beliebige $\zeta \in N_C(x)$ folgt die Behauptung mit Proposition 2.4 und Aufgabe 2.2 (b). □

Aus Proposition 1.22 folgt direkt

Corollar 2.9
Sei C konvex, $x \in C$. DANN gilt

$$v \in T_C(x) \iff d_C^0(x;v) = d_C'(x;v) = 0.$$

Satz 2.10
Sei $x \in C$.

$$v \in T_C(x) \iff \begin{cases} \forall\, (x_i)_{i\in\mathbb{N}} \subset C, x_i \to x, t_i \searrow 0 \;\exists\; (v_i)_{i\in\mathbb{N}} \subset X, v_i \to v, \\ \text{so dass } x_i + t_i v_i \in C \end{cases}$$

Beweis. „$\Rightarrow$“: Es gilt $d_C^0(x;v) = 0$. Sei $(x_i)_{i\in\mathbb{N}}$ eine Folge in C mit $x_i \to x$, $t_i \searrow 0$. Wir haben

$$0 = d_C^0(x;v) \geq \limsup_{i\to\infty} \frac{\overbrace{d_C(x_i+t_iv)}^{\geq 0} - \overbrace{d_C(x_i)}^{=0}}{t_i},$$

also $\lim_{i\to\infty} \frac{d_C(x_i+t_iv)}{t_i} = 0$. Wir wählen nun eine Folge $(c_i)_{i\in\mathbb{N}} \subset C$ mit

$$\|x_i + t_iv - c_i\|_X \leq d_C(x_i+t_iv) + \frac{t_i}{i}.$$

Für $v_i := \frac{c_i - x_i}{t_i}$ ergibt sich dann $x_i + t_iv_i = c_i \in C$. Wir haben also

$$\|v_i - v\|_X = \left\|\frac{c_i - x_i - t_iv}{t_i}\right\|_X \leq \frac{d_C(x_i+t_iv)}{t_i} + \frac{1}{i} \to 0.$$

„$\Leftarrow$“: Wir betrachten Folgen $(y_i)_{i\in\mathbb{N}} \subset X$, $y_i \to x$, und $t_i \searrow 0$, die

$$d_C^0(x;v) = \lim_{i\to\infty} \frac{d_C(y_i+t_iv) - d_C(y_i)}{t_i}$$

realisieren. Wegen $d_C \geq 0$, $d_C(x) = 0$ brauchen wir nur „≤ 0“ zu zeigen. Hierzu wählen wir eine Folge $(x_i)_{i\in\mathbb{N}} \subset C$ mit $\|x_i - y_i\|_X \leq d_C(y_i) + \frac{t_i}{i}$. Nach Voraussetzung erhalten wir dann eine Folge $(v_i)_{i\in\mathbb{N}}$ mit $v_i \to v$ und $x_i + t_iv_i \in C$, so dass

$$\begin{aligned} d_C(y_i+t_iv) &\leq d_C(y_i+t_iv) - d_C(x_i+t_iv_i) + \underbrace{d_C(x_i+t_iv_i)}_{=0} \\ &\leq \|y_i - x_i\|_X + t_i\|v-v_i\|_X \\ &\leq d_C(y_i) + \frac{t_i}{i} + t_i\|v-v_i\|_X, \end{aligned}$$

woraus sich

$$\frac{d_C(y_i+t_iv) - d_C(y_i)}{t_i} \leq \frac{1}{i} + \|v-v_i\|_X \xrightarrow{i\to\infty} 0$$

ergibt. $\square$

Corollar 2.11
Für $X = X_1 \times X_2$, $C = C_1 \times C_2$, $x = (x_1,x_2) \in C$ ist

$$T_C(x) = T_{C_1}(x_1) \times T_{C_2}(x_2), \qquad N_C(x) = N_{C_1}(x_1) \times N_{C_2}(x_2).$$

Der Beweis besteht in der Anwendung von Satz 2.10 und Polarisierung, vgl. Aufgabe 2.8.

Definition 2.12 [Kontingent-Kegel]
Der *Kontingent-Kegel an C in $x \in C$* ist als

$$K_C(x) := \{v \in X \mid \forall\, \varepsilon > 0\ \exists\, t \in (0,\varepsilon), w \in B_\varepsilon(v) : x + tw \in C\}$$

definiert.

Bemerkung.

(i) Wir haben

$$T_C(x) \subset K_C(x), \tag{2.1}$$

denn zu jedem $v \in T_C(x)$ lässt sich bei Vorgabe der konstanten Folge $x_i \equiv x$ und einer Nullfolge $t_i \searrow 0$ nach Satz 2.10 eine Folge $v_i \to v$ mit $x + t_i v_i \in C$ finden. Wenn wir $i \gg 1$ zu vorgegebenem $\varepsilon > 0$ wählen, liegt $t_i \in (0,\varepsilon)$ und $v_i \in B_\varepsilon(v)$, mithin $v \in K_C(x)$.

(ii) $K_C(x)$ ist nicht notwendig konvex.

(iii) $K_{C_1 \times C_2}(x_1, x_2) \subset K_{C_1}(x_1) \times K_{C_2}(x_2)$, vgl. Aufgabe 2.8.

Definition 2.13 [Reguläre Mengen]
C heißt *regulär in $x \in C$*, falls $K_C(x) = T_C(x)$.

Beispiel 2.14 Für den durch $C := \{c \in \mathbb{R}^2 \mid |c_1| \le |c_2|\}$ definierten Kegel ist $T_C(0) = \{0\}$ (und damit $N_C(0) = \mathbb{R}^2$), aber $K_C(0) = C$, also $T_C(0) \subsetneq K_C(0)$. $\diamond$

Lemma 2.15
C konvex $\Longrightarrow$ C regulär.

Beweis. In der Tat folgt aus $v \in K_C(x)$, dass $d_C^0(x;v) = 0$. Sei nämlich eine Nullfolge $\varepsilon_i \searrow 0$ vorgegeben. Die Voraussetzung liefert uns zu jedem $i \in \mathbb{N}$ ein $t_i \in (0, \varepsilon_i)$ und ein $w_i \in B_{\varepsilon_i}(v)$ mit $x + t_i w_i \in C$. Nach Propositionen 1.32 (ii), 2.2 (iii) ist d_C regulär, also

$$\begin{aligned}
d_C^0(x;v) = d_C'(x;v) &= \lim_{i\to\infty} \frac{d_C(x + t_i v) - \overbrace{d_C(x)}^{=0}}{t_i} \\
&= \lim_{i\to\infty} \frac{d_C(x + t_i v) - \overbrace{d_C(x + t_i w_i)}^{=0}}{t_i} \\
&\le \lim_{i\to\infty} \|v - w_i\|_X \le \lim_{i\to\infty} \varepsilon_i = 0,
\end{aligned}$$

wobei wir die LIPSCHITZ-Stetigkeit der Distanzfunktion benutzt haben. $\square$

Die im soeben bewiesenen Resultat zutage getretene Parallele zwischen Mengen und Funktionen lässt sich folgendermaßen formulieren.

Satz 2.16
Sei $f : X \to \mathbb{R}$ LIPSCHITZ-*stetig nahe* x, $0_{X^*} \notin \partial f(x)$ *und* $C := \{y \in X \mid f(y) \leq f(x)\}$. *DANN gilt*

$$\left\{v \in X \mid f^0(x;v) \leq 0\right\} \subset T_C(x). \tag{2.2}$$

Gleichheit in (2.2) *tritt ein, falls* f *regulär. In diesem Fall ist auch* C *regulär.*

Beweis. Wir können ohne Einschränkung annehmen, dass es ein $\hat{v} \in X$ mit $f^0(x;\hat{v}) < 0$ gibt, denn anderenfalls hätten wir $f^0(x;v) \geq 0$ für alle $v \in X$, also $0 = \langle 0_{X^*}, v\rangle \leq f^0(x;v)$ für alle $v \in X$, so dass $0_{X^*} \in \partial f(x)$ im Gegensatz zur Voraussetzung. Es genügt,

$$f^0(x;\hat{v}) < 0 \Longrightarrow d_C^0(x;\hat{v}) = 0 \tag{2.3}$$

zu zeigen, denn ein beliebiges $\bar{v} \in X$ mit $f^0(x;\bar{v}) \leq 0$ kann mittels $v_\varepsilon := \bar{v} + \varepsilon\hat{v}$ approximiert werden, so dass nach Proposition 1.3 (ii), (iii)

$$f^0(x;v_\varepsilon) \leq \underbrace{f^0(x;\bar{v})}_{\leq 0} + \varepsilon f^0(x;\hat{v}) < 0 \qquad \forall\, \varepsilon > 0.$$

Mit (2.3) ergibt sich $v_\varepsilon \in T_C(x)$, und die Abgeschlossenheit von $T_C(x)$ (Aufgabe 2.2 (a)) liefert schließlich $v \in T_C(x)$.
Um (2.3) zu beweisen, betrachten wir ein $\hat{v} \in X$ mit $f^0(x;\hat{v}) < 0$. Wegen

$$f^0(x;\hat{v}) = \lim_{\varepsilon \searrow 0} \sup_{\substack{y \in B_\varepsilon(x)\\ t \in (0,\varepsilon)}} \frac{f(y + t\hat{v}) - f(y)}{t}$$

gibt es $\varepsilon, \delta > 0$ mit

$$f(y + t\hat{v}) - f(y) \leq -\delta t \quad \forall\, y \in B_\varepsilon(x), t \in (0,\varepsilon).$$

Zu beliebigen Folgen $(x_i)_{i\in\mathbb{N}} \subset C \cap B_\varepsilon(x)$, $x_i \to x$, und $t_i \searrow 0$ haben wir wegen $f(x_i) \leq f(x)$ die Abschätzung

$$f(x_i + t_i\hat{v}) \leq -\delta t_i + f(x_i) \leq f(x),$$

also $x_i + t_i\hat{v} \in C$. Satz 2.10 liefert (2.3).
Nun ist noch die Gleichheitsdiskussion zu führen. Wir zeigen

$$K_C(x) \subset \left\{v \in X \;\middle|\; f^0(x;v) \leq 0\right\},$$

dann folgt die Behauptung mit (2.1) und (2.2). Sei also $v \in K_C(x)$, dann ist $\liminf_{t\searrow 0} \frac{d_C(x+tv) - d_C(x)}{t} = 0$, denn es gibt für jedes $\varepsilon > 0$ Elemente $t_\varepsilon \in (0,\varepsilon)$ und $w \in B_\varepsilon(v)$ mit $x + t_\varepsilon w_\varepsilon \in C$, so dass

$$\begin{aligned} 0 = \liminf_{t\searrow 0} \frac{d_C(x+tv) - \overbrace{d_C(x)}^{=0}}{t} &\leq \lim_{\varepsilon\searrow 0} \frac{d_C(x + t_\varepsilon v)}{t_\varepsilon} = \lim_{\varepsilon\searrow 0} \frac{d_C(x + t_\varepsilon v) - d_C(x + t_\varepsilon w_\varepsilon)}{t_\varepsilon} \\ &\leq \lim_{\varepsilon\searrow 0} \frac{t_\varepsilon \varepsilon}{t_\varepsilon} = 0. \end{aligned}$$

Wir wählen nun eine Folge $t_i \searrow 0$, die

$$\lim_{i\to\infty} \frac{d_C(x+t_iv)}{t_i} = \liminf_{t\searrow 0} \frac{d_C(x+tv) - \overbrace{d_C(x)}^{=0}}{t} = 0$$

realisiert, d. h. zu jedem $\varepsilon > 0$ gibt es ein $i_0 \in \mathbb{N}$, so dass $d_C(x+t_iv) \leq \varepsilon t_i$ für alle $i \geq i_0$ gilt. Daher können wir eine Folge $(x_i)_{i\in\mathbb{N}} \subset C$ mit

$$\|x_i - (x+t_iv)\|_X \leq 2\varepsilon t_i$$

wählen und erhalten mit $f(x_i) \leq f(x)$ für alle $\varepsilon > 0$

$$\begin{aligned} f^0(x;v) \overset{f \text{ reg.}}{=} f'(x;v) &= \lim_{i\to\infty} \frac{f(x+t_iv) - f(x)}{t_i} \\ &= \lim_{i\to\infty} \frac{f(x+t_iv) - f(x_i) + f(x_i) - f(x)}{t_i} \\ &\leq \limsup_{i\to\infty} \frac{f(x+t_iv) - f(x_i)}{t_i} \leq 2K\varepsilon, \end{aligned}$$

so dass $f^0(x;v) \leq 0$, da $\epsilon > 0$ beliebig klein gewählt werden kann. K ist hier die Lipschitz-Konstante von f. ☐

Corollar 2.17
Sei $f : X \to \mathbb{R}$ Lipschitz-stetig nahe x, $0_{X^*} \notin \partial f(x)$ und $C := \{y \in X \mid f(y) \leq f(x)\}$. DANN gilt

$$N_C(x) \subset \bigcup_{\lambda \geq 0} \lambda \partial f(x),$$

und Gleichheit tritt ein, falls f regulär.

Beweis. Wir setzen

$$\tilde{K}(x) := \left\{ \zeta \in X^* \mid \langle \zeta, v\rangle \leq 0 \; \forall \, v \in X \text{ mit } f^0(x;v) \leq 0 \right\} = \text{Polar}\left(\left\{ w \in X \mid f^0(x;w) \leq 0 \right\} \right).$$

Es gilt $T_C(x) \supset \left\{ w \in X \mid f^0(x;w) \leq 0 \right\}$ nach Satz 2.16, also wie in Punkt (ii) der Bemerkung nach Definition 2.3

$$N_C(x) = \text{Polar}(T_C(x)) \subset \text{Polar}\left(\left\{ w \in X \mid f^0(x;v) \leq 0 \right\} \right) = \tilde{K}(x).$$

Die Behauptung folgt also, wenn wir $\tilde{K}(x) \subset \bigcup_{\lambda\geq 0} \lambda \partial f(x)$ zeigen. $\boxed{\text{BAUSTELLE}}$
Die Gleichheitsdiskussion verläuft wie in Satz 2.16. ☐

Corollar 2.18
Seien $f_1, \ldots, f_n : X \to \mathbb{R}$ strikt differenzierbar in x, $f_1(x) = \cdots = f_n(x) = 0$ und $C := \{y \in X \mid f_1(y) \leq 0, \ldots, f_n(y) \leq 0\}$. Ferner seien die $D_s f_i(x)$, $i = 1, \ldots, n$ positiv linear unabhängig, *d.h.*

$$\textstyle\sum_{i=1}^n \lambda_i D_s f_i(x) = 0, \quad \lambda_i \geq 0 \quad \Longrightarrow \quad \lambda_1 = \cdots = \lambda_n = 0.$$

DANN ist C regulär bei x und

$$N_C(x) = \left\{ \textstyle\sum_{i=1}^n \lambda_i D_s f_i(x) \mid \lambda_i \geq 0, i = 1, \ldots, n \right\}.$$

Bemerkung. Die positive lineare Unabhängigkeit modelliert die Bedingung „$0_{X^*} \notin \partial f(x)$“ in Corollar 2.17.

Beweis. Wir setzen $f(y) := \max_{i\in\{1,\dots,n\}} f_i(y)$, dann ist $f(x) = 0$, f regulär bei x und $C = \{ y \in X \mid f(y) \leq 0\}$. Mit Corollar 2.17 haben wir $N_C(x) = \bigcup_{\lambda\geq 0} \lambda\partial f(x)$. Nach Proposition 1.40 ist

$$\partial f(x) = \Big\{ \sum_{i=1}^n \mu_i D_s f_i(x) \Big| \sum_{i=1}^n \mu_i = 1, \mu_i \geq 0, i \in \underbrace{\{1,\dots,n\}}_{=I(x)} \Big\},$$

also $N_C(x) = \{ \sum_{i=1}^n \lambda_i D_s f_i(x) \mid \lambda_i \geq 0, i \in \{1,\dots,n\}\}$. □

Definition 2.19
$v \in X$ heißt *hypertangential an* $C \subset X$ *in* x, wenn es ein $\varepsilon > 0$ gibt, so dass

$$\forall y \in B_\varepsilon(x) \cap C, \quad t \in (0,\varepsilon), \quad w \in B_\varepsilon(v) \quad : \quad y + tw \in C.$$

Die Menge der an C in x hypertangentialen Vektoren $v \in X$ wird mit $HT_C(x)$ bezeichnet, d. h.

$$HT_C(x) = \{v \in X \mid \exists\, \varepsilon = \varepsilon(v) > 0\ \forall\, t \in (0,\varepsilon) : (B_\varepsilon(x) \cap C) + tB_\varepsilon(v) \subset C\}.$$

Bemerkung.

(i) v *hypertangential an* C *in* $x \Longrightarrow v \in T_C(x)$. In der Tat, zu vorgegebenen Folgen $(x_i)_{i\in\mathbb{N}} \subset C$, $x_i \to x$, $t_i \searrow 0$ haben wir für $i \gg 1$ nach Voraussetzung $x_i \in B_\varepsilon(x) \cap C, t_i \in (0,\varepsilon), w_i \in B_\varepsilon(v)$, so dass $x_i + t_i w_i \in C$. Mit Satz 2.10 folgt die Behauptung.

(ii) Es gibt Fälle, in denen es keine Hypertangenten gibt.

Satz 2.20 [Rockafellar[14]]

$$HT_C(x) \neq \emptyset \quad \Longrightarrow \quad HT_C(x) = \operatorname{int} T_C(x)$$

Beweis. $HT_C(x)$ ist nach Aufgabe 2.4 ein offener positiver Kegel; Punkt (i) der voranstehenden Bemerkung liefert also $HT_C(x) \subset \operatorname{int} T_C(x)$, zu zeigen bleibt $\operatorname{int} T_C(x) \subset HT_C(x)$. Dies folgt aber sofort aus

$$HT_C(x) + T_C(x) \subset HT_C(x); \tag{2.4}$$

denn für ein $v \in \operatorname{int} T_C(x)$ gibt es zu jedem $w \in HT_C(x) \subset T_C(x)$ ein $\lambda = \lambda(w) \in (0,1)$ mit $v - \lambda w \in \operatorname{int} T_C(x)$; wir haben also

$$v = \underbrace{\lambda w}_{\in HT_C(x)} + \underbrace{v - \lambda w}_{\in T_C(x)} \in HT_C(x),$$

[14] Ralph Tyrrell Rockafellar (*1935) Amerikanischer Mathematiker. Studium / Promotion in Harvard, seit 1973 Professor in Seattle. Rockafellar arbeitet auf den Gebieten Optimierung, Konvexe Analysis und Funktionalanalysis.

wobei wir ausgenutzt haben, dass $w \in HT_C(x)$ sofort $\lambda w \in HT_C(x)$ impliziert, setze hierzu $\varepsilon(\lambda w) := \lambda\varepsilon(w)$.
Seien nun $v_1 \in HT_C(x)$ und $v_2 \in T_C(x)$. Zum Beweis von (2.4) zeigen wir

$$\exists\ \varepsilon = \varepsilon(v_1, v_2) > 0\ \forall\ t \in (0, \varepsilon) : (B_\varepsilon(x) \cap C) + tB_\varepsilon(v_1 + v_2) \subset C. \tag{2.5}$$

Wir bemerken hierzu

$$\begin{aligned} &\exists\ \varepsilon_1 = \varepsilon_1(v_1)\ \forall\ t \in (0, \varepsilon_1) \ :\ (B_{\varepsilon_1}(x) \cap C) + tB_{\varepsilon_1}(v_1) \subset C & (2.6)\\ \text{und}\quad &\exists\ \varepsilon_2 = \varepsilon_2(v_2)\ \forall\ t \in (0, \varepsilon_2) \ :\ (B_{\varepsilon_2}(x) \cap C) + tv_2 \subset C + B_{t\frac{\varepsilon_1}{2}}(0). & (2.7)\end{aligned}$$

Hierbei ist (2.6) direkte Konsequenz der Definition des Hypertangentialraums; wäre (2.7) falsch, so hätten wir für alle $\varepsilon_2 > 0$ ein $y \in B_{\varepsilon_2}(x) \cap C$ und ein $t \in (0, \varepsilon_2)$, so dass für jedes $c \in C$ die Abschätzung $\|y + tv_2 - c\|_X \geq t\frac{\varepsilon_1}{2}$ gälte. Zu einem Folgenpaar $(y_i, t_i) \to (x, 0)$ lieferte Satz 2.10 eine Folge $v_i \to v_2$ mit $y_i + t_i v_i \in C$, und wir hätten

$$\|y_i + t_i v_2 - (y_i + t_i v_i)\|_X \leq t_i \|v_2 - v_i\|_X < t_i \frac{\varepsilon_1}{2} \text{ für } i \gg 1. \quad ↯$$

Wir setzen jetzt $\varepsilon := \min\left\{\varepsilon_2, \frac{\varepsilon_1}{2}, \frac{\varepsilon_1}{1+\varepsilon_1+\|v_2\|_X}\right\}$ und werden sehen, dass dieses (2.5) erfüllt. Seien nämlich $y \in B_\varepsilon(x) \cap C$, $t \in (0, \varepsilon)$ und $u \in B_1(0)$ vorgegeben. Dann ist $z := y + t(v_1 + v_2 + \varepsilon u) \overset{!}{\in} C$, denn wegen $\varepsilon \leq \varepsilon_2$ gibt es nach (2.7) ein $w \in B_1(0)$ mit $z' := y + tv_2 - t\frac{\varepsilon_1}{2}w \in C$. Die Ungleichung

$$\left\|x - z'\right\|_X = \left\|x - y - t(v_2 - \frac{\varepsilon_1}{2}w)\right\|_X \leq \|x - y\|_X + \varepsilon(\|v_2\|_X + \varepsilon_1) < \varepsilon + \varepsilon(\|v_2\|_X + \varepsilon_1) \leq \varepsilon_1$$

liefert sogar $z' \in C \cap B_{\varepsilon_1}(x)$. Eigenschaft (2.6) impliziert $z' + tB_{\varepsilon_1}(v_1) \subset C$, und dies können wir wegen $\left\|\varepsilon u + \frac{\varepsilon_1}{2}w\right\|_X < \varepsilon + \frac{\varepsilon_1}{2} \leq \varepsilon_1$ auf

$$z = \underbrace{y + tv_2 - t\frac{\varepsilon_1}{2}w}_{=z'} + t\left(v_1 + \varepsilon u + \frac{\varepsilon_1}{2}w\right)$$

übertragen, was (2.5) beweist. $\square$

Corollar 2.21
Sei $x \in C$, $HT_C(x) \neq \emptyset$. DANN ist $N_C(x)$ „abgeschlossen in x", d. h. für Folgen $x_i \to x$, $\zeta_i \overset{*}{\rightharpoonup} \zeta$, $\zeta_i \in N_C(x_i)$ gilt $\zeta \in N_C(x)$.

Beweis. Wir wählen $v \in HT_C(x)$ und erhalten $v \in T_C(y)$ für alle $y \in C$ nahe x, denn Folgen $(y_i)_{i\in\mathbb{N}} \subset C$, $y_i \to y \in B_{\varepsilon(v)}(x)$, $t_i \searrow 0$ erfüllen

$$\left.\begin{aligned} y_i &\in B_\varepsilon(x)\\ t_i &\in (0, \varepsilon)\end{aligned}\right\}\ \forall\ i \gg 1.$$

Nach Definition des Hypertangentialkegels folgt $y_i + t_i w \in C$ für alle $w \in B_\varepsilon(v)$. Also haben wir für die konstante Folge $v_i \equiv v$ die Eigenschaft $y_i + t_i v_i \in C$, mithin $v \in T_C(y)$ nach Satz 2.10.
Für Folgen $x_i \to x$, $\zeta_i \overset{*}{\rightharpoonup} \zeta$, $\zeta_i \in N_C(x_i)$ ergibt sich deshalb $v \in T_C(x_i)$ mit $i \gg 1$ und

$$0 \overset{\text{Def. 2.3 (ii)}}{\geq} \langle\zeta_i, v\rangle \quad \overset{\zeta_i \overset{*}{\rightharpoonup} \zeta}{\Longrightarrow} \quad 0 \geq \langle\zeta, v\rangle\ \forall\ v \in HT_C(x) \overset{\text{Satz 2.20}}{=} \operatorname{int} T_C(x).$$

Weil wir für $T_C(x)$ nach Aufgaben 2.2 (a) und 2.5 die Identität $T_C(x) = \overline{\operatorname{int} T_C(x)}$ haben, gilt $\langle\zeta, v\rangle \leq 0$ sogar für alle $v \in T_C(x)$, also $\zeta \in N_C(X)$ nach Definition 2.3 (ii). $\square$

Definition 2.22 [Epigraph]
Der *Epigraph* von $f : C \to \mathbb{R}$ ist die Menge

$$\text{epi}\, f := \{\, (x,r) \in C \times \mathbb{R} \,|\, r \geq f(x)\} \,.$$

Satz 2.23
Sei f Lipschitz-stetig nahe $x \in X$ mit Lipschitz-Konstante K. DANN gilt

(i) $\text{epi}\, f^0(x;\cdot) = T_{\text{epi}\, f}(x, f(x))$ *und*

(ii) f *regulär bei* $x \iff \text{epi}\, f$ *regulär bei* $(x, f(x))$.

Beweis von Satz 2.23 (i). Wir zeigen $(v,r) \in T_{\text{epi}\, f}(x, f(x)) \iff f^0(x;v) \leq r$.
„$\Rightarrow$“: Wir wählen Folgen $y_i \to x$ und $t_i \searrow 0$, die

$$f^0(x;v) = \lim_{i\to\infty} \frac{f(y_i + t_i v) - f(y_i)}{t_i}$$

realisieren. Wegen $(y_i, f(y_i)) \in \text{epi}\, f$ und $(v.r) \in T_{\text{epi}\, f}(x, f(x))$ erhalten wir nach Satz 2.10 eine Folge $(v_i, r_i) \to (v,r)$ mit $(y_i, f(y_i)) + t_i(v_i, r_i) \in \text{epi}\, f$, was äquivalent zu $f(y_i) + t_i r_i \geq f(y_i + t_i v_i)$ ist. Nun gilt

$$f^0(x;v) \xleftarrow{i\to\infty} \frac{f(y_i+t_iv)-f(y_i)}{t_i} + \underbrace{\frac{f(y_i+t_iv_i)-f(y_i+t_iv)}{t_i}}_{|\cdot| \leq K\frac{t_i\|v_i-v\|_X}{t_i}} = \frac{f(y_i+t_iv_i)-f(y_i)}{t_i} \quad \leq \quad r_i \xrightarrow{i\to\infty} r,$$

also $f^0(x;v) \leq r$.
„$\Leftarrow$“: Wir behaupten $(v, f^0(x;v) + \delta) \in T_{\text{epi}\, f}(x, f(x))$ für jedes $\delta \geq 0$. Um Satz 2.10 anwenden zu können, betrachten wir Folgen $(x_i, r_i) \to (x, f(x))$ mit $r_i \geq f(x_i)$, so dass $(x_i, r_i) \in \text{epi}\, f$, und $t_i \searrow 0$. Wir müssen eine Folge $(v_i, s_i) \to (v, f^0(x;v) + \delta)$ mit der Eigenschaft $(x_i, r_i) + t_i(v_i, s_i) \in \text{epi}\, f$ konstruieren. Hierzu wählen wir Folgen

$$\begin{aligned} v_i &\equiv v \\ \text{und} \qquad s_i &:= \max\left\{ f^0(x;v) + \delta, \frac{f(x_i + t_i v) - f(x_i)}{t_i} \right\}. \end{aligned}$$

Wir haben $s_i \xrightarrow{i\to\infty} f^0(x;v) + \delta$ und

$$r_i + t_i s_i \geq r_i + [f(x_i + t_i v) - f(x_i)] \geq f(x_i) + [f(x_i + t_i v) - f(x_i)] = f(x_i + t_i v),$$

folglich $(v, f^0(x;v) + \delta) \in T_{\text{epi}\, f}(x, f(x))$. □

Für den Beweis der zweiten Aussage von Satz 2.23 führen wir jetzt den Begriff der Dini-Ableitung ein.

Definition 2.24 [Dini-Ableitung]
Sei f eine auf einer Umgebung von $x \in X$ definierte reellwertige Funktion. Der Ausdruck

$$f'_+(x;v) := \liminf_{t\searrow 0} \frac{f(x + tv) - f(x)}{t}$$

wird DINI[15]-*Ableitung von f in Richtung $v \in X$ bei x* genannt.

Lemma 2.25
Sei f LIPSCHITZ-stetig nahe $x \in X$ und epi f *regulär in $(x, f(x))$.*
DANN ist $K_{\text{epi } f}(x, f(x)) = \text{epi } f'_+(x; \cdot)$.

Beweis von Lemma 2.25. „$\subset$“: Wir haben nach Definition 2.12

$$\begin{aligned} & K_{\text{epi } f}(x, f(x)) \\ = \; & \{(v, r) \in X \times \mathbb{R} \mid \forall\, \varepsilon > 0\ \exists\, t \in (0, \varepsilon), (w, s) \in B_\varepsilon(v, r) : (x, f(x)) + t(w, s) \in \text{epi } f\}\,. \end{aligned}$$

Sei $(v, r) \in K_{\text{epi } f}(x, f(x))$ vorgegeben. Wir wählen nun eine Nullfolge $t_i \searrow 0$, die

$$f'_+(x; v) = \lim_{i\to\infty} \frac{f(x + t_i v) - f(x)}{t_i} \tag{2.8}$$

realisiert. Nach Voraussetzung ist epi f regulär bei $(x, f(x))$. Definition 2.13 liefert also $T_{\text{epi } f}(x, f(x)) = K_{\text{epi } f}(x, f(x))$, und zu Folgen $x_i \equiv x$, $t_i \searrow 0$ liefert uns Satz 2.10 eine Folge $(v_i, r_i) \to (v, r)$ mit $(x, f(x)) + t_i(v_i, r_i) \in \text{epi } f$. Daraus erhalten wir $f(x) + t_i r_i \geq f(x + t_i v_i)$, mithin

$$f'_+(x; v) \xleftarrow{i\to\infty} \frac{f(x + t_i v_i) - f(x)}{t_i} \leq r_i \xrightarrow{i\to\infty} r,$$

d.h. $(v, r) \in \text{epi } f'_+(x; \cdot)$. Der Grenzübergang auf der linken Seite folgt aus (2.8) und der LIPSCHITZ-stetigkeit von f nahe x.
Der Beweis von „$\supset$“ wird als Aufgabe 2.7 gestellt. ☐

Beweis von Satz 2.23 (ii). „$\Rightarrow$“: Sei f regulär bei x. Wir definieren die Funktion $g : X \times \mathbb{R} \to \mathbb{R}$ durch $(\bar{x}, r) \mapsto f(\bar{x}) - r$. Folglich ist g regulär bei $(x, f(x))$. Nach Proposition 1.43 haben wir

$$\partial g(x, f(x)) \subset \partial_1 g(x, f(x)) \times \partial_2 g(x, f(x)) = \partial f(x) \times \{-1\},$$

und $0_{(X\times\mathbb{R})^*}$ liegt nicht in dieser Menge, so dass wir die Regularitätsaussage aus Satz 2.16 auf

$$\text{epi } f = \{(z, r) \in X \times \mathbb{R} \mid f(z) \leq r\} = \{(z, r) \in X \times \mathbb{R} \mid g(z, r) \leq 0 = g(x, f(x))\}$$

anwenden dürfen.
„$\Leftarrow$“: Sei epi f regulär bei $(x, f(x))$. Wir haben

$$\text{epi } f'_+(x; \cdot) \overset{\text{Lemma 2.25}}{=} K_{\text{epi } f}(x, f(x)) \overset{\text{Def. 2.13}}{=} T_{\text{epi } f}(x, f(x)) \overset{\text{Satz 2.23 (i)}}{=} \text{epi } f^0(x; \cdot).$$

Aufgabe 2.6 erbringt $f'_+(x; \cdot) = f^0(x; \cdot)$, insbesondere also

$$f'_+(x; v) = \liminf_{t\searrow 0} \frac{f(x + tv) - f(x)}{t} \overset{!}{=} \limsup_{t\searrow 0} \frac{f(x + tv) - f(x)}{t},$$

d. h. $\lim_{t\searrow 0} \frac{f(x+tv)-f(x)}{t}$ existiert für alle $v \in X$ und stimmt mit $f^0(x; v)$ überein. Das ist gerade die Regularität von f, vgl. Definition 1.30, Seite 20. ☐

[15] ULISSE DINI (1845 – 1918) Italienischer Mathematiker. DINI war ab 1871 Professor in Pisa, wo er auch studiert hatte. Er gilt als Begründer der italienischen Schule der reellen Funktionentheorie und beschäftigte sich unter anderem mit Bedingungen für die Differenzierbarkeit stetiger Funktionen und Verallgemeinerungen des Begriffs der gleichmäßigen Konvergenz. Bedeutende Schüler von ihm sind ARZELÀ, VITALI und VOLTERRA.

Für eine glatte Funktion $f : \mathbb{R}^n \to \mathbb{R}$ ist $(\nabla f, -1)$ normal zu $\mathrm{Graph}(f)$. Dies verallgemeinert nun das folgende

Corollar 2.26
Sei f LIPSCHITZ-stetig nahe $x \in X$. DANN gilt

$$\zeta \in \partial f(x) \iff (\zeta, -1) \in N_{\mathrm{epi}\, f}(x, f(x)).$$

Beweis. „$\Rightarrow$": Sei $(v, r) \in T_{\mathrm{epi}\, f}(x) \overset{\text{Satz 2.23 (i)}}{=} \mathrm{epi}\, f^0(x; \cdot)$ gegeben, d. h. $f^0(x; v) \leq r$. Für ein $\zeta \in \partial f(x)$ haben wir dann

$$\langle (\zeta, -1), (v, r) \rangle_{(X \times \mathbb{R})^* \times (X \times \mathbb{R})} = \langle \zeta, v \rangle_{X^* \times X} - r \leq f^0(x; v) - r \leq 0.$$

„$\Leftarrow$": Mit $(v, f^0(x; v)) \in T_{\mathrm{epi}\, f}(x)$ erhalten wir

$$\langle \zeta, v \rangle = \underbrace{\big\langle (\zeta, -1), (v, f^0(x; v)) \big\rangle}_{\leq 0} + f^0(x; v) \leq f^0(x; v).$$

$\square$

Damit können wir zwei wichtige Begriffe auch auf nicht LIPSCHITZ-stetige Funktionen übertragen.

Definition 2.27 [Verallgemeinerter Gradient, Regularität]
Sei $f : X \to \mathbb{R} \cup \{\pm\infty\}$ endlich an der Stelle $x \in X$. Dann setzen wir

$$\partial f(x) := \{\, \zeta \in X^* \mid (\zeta, -1) \in N_{\mathrm{epi}\, f}(x, f(x)) \}$$

und nennen f *regulär* bei x, sofern epi f regulär an der Stelle $(x, f(x))$ ist.

Bemerkungen.

(i) Diese Definition ist aufgrund von Corollar 2.26 und Satz 2.23 (ii) konsistent mit Definitionen 1.4 und 1.30.

(ii) Der verallgemeinerte Gradient kann für nicht LIPSCHITZ-stetige Funktionen leer sein.

(iii) $\partial f(x)$ ist schwach$*$-abgeschlossen, aber nicht mehr notwendig schwach$*$-kompakt.

Proposition 2.28
Sei $f : X \to \mathbb{R} \cup \{\pm\infty\}$ eine Funktion und $x \in X$ ein Punkt mit $f(x) < \infty$. Ist $f(x) = \inf_U f$ für eine Umgebung U von x, DANN gilt $0 \in \partial f(x)$ (insbesondere also $\partial f(x) \neq \emptyset$).

Beweis. Wir müssen $(0, -1) \in N_{\mathrm{epi}\, f}(x, f(x))$ zeigen, also $\langle (0, -1), (v, r) \rangle_{(X^* \times \mathbb{R}) \times (X \times \mathbb{R})} \leq 0$ oder $r \geq 0$ für alle $(v, r) \in T_{\mathrm{epi}\, f}(x, f(x))$. Dazu sei $(v, r) \in T_{\mathrm{epi}\, f}(x, f(x))$ vorgegeben. Aus Satz 2.10 folgt die Existenz einer gegen (v, r) konvergierenden Folge $(v_i, r_i) \subset X \times \mathbb{R}$ und einer Nullfolge $t_i \searrow 0$ mit $(x, f(x)) + t_i(v_i, r_i) \in \mathrm{epi}\, f$, d. h. $f(x) + t_i r_i \geq f(x + t_i v_i) \geq f(x)$, da $x + t_i v_i \in U$ für $i \gg 1$. Wegen $t_i r_i \geq 0$ haben wir $0 \leq r_i \to r$. $\square$

Proposition 2.29 [Indikatorfunktion]
Sei $x \in C \subset X$. DANN gilt

$$\zeta \in \partial i_C(x) \implies \zeta \in N_C(x)$$

und zusätzlich

$$i_C \text{ regulär bei } x \iff C \text{ regulär bei } x.$$

Beweis. $\zeta \in \partial i_C(x)$ ist nach Definition 2.27 äquivalent zu $(\zeta,-1) \in N_{\mathrm{epi}\, i_C}(x, \underbrace{i_C(x)}_{=0})$. Nun ist $\mathrm{epi}\, i_C = C \times [0,\infty)$, also $N_{\mathrm{epi}\, i_C}(x,0) = N_{C\times[0,\infty)}(x,0) \overset{\text{Cor. 2.11}}{=} N_C(x) \times N_{[0,\infty)}(0) = N_C(x) \times (-\infty, 0]$, mithin $\zeta \in N_C(x)$.
i_C ist nach Definition 2.27 genau dann regulär bei x, wenn $\mathrm{epi}\, i_C = C \times [0,\infty)$ regulär bei $(x,0)$ ist. Dies ist aber nach Definition 2.13 gleichbedeutend mit

$$\begin{array}{ccc} T_{C\times[0,\infty)}(x,0) & = & K_{C\times[0,\infty)}(x,0) \\ \overset{\text{Cor. 2.11}}{=} T_C(x) \times \underbrace{T_{[0,\infty)}(0)}_{=[0,\infty)} & & \overset{\text{Aufg. 2.8}}{=} K_C(x) \times \underbrace{K_{[0,\infty)}(0)}_{=[0,\infty)}\ . \end{array}$$

Hieraus folgt $T_C(x) = K_C(x)$, also C regulär nach Definition 2.13. □

Aufgaben

2.1 Beweisen Sie Proposition 2.2!

2.2 Sei X ein BANACHraum.

(a) Sei $x \in C \subset X$, $C \neq \emptyset$. Zeigen Sie, dass $T_C(x)$ (vgl. Definition 2.3) ein abgeschlossener, konvexer, (positiver) Kegel in X mit $0 \in T_C(x)$ ist. Zeigen Sie auch, dass $N_C(x)$ ein schwach$*$-abgeschlossener, konvexer, (positiver) Kegel in X^* mit $0 \in N_C(x)$ ist.
(Eine Menge M ist positiver Kegel*, falls mit m auch λm in M ist für alle $\lambda \geq 0$.)*

(b) Sei $x \in C \subset X$, $C \neq \emptyset$, C konvex. Dann ist der Normalkegel $\tilde{N}_C(x)$ im Sinne der konvexen Analysis definiert durch (vgl. Definition 2.7)

$$\tilde{N}_C(x) := \{\zeta \in X^* : \langle \zeta, x - c\rangle \geq 0 \ \text{ für alle } \ c \in C\}.$$

Zeigen Sie, dass $\tilde{N}_C(x)$ ein schwach$*$-abgeschlossener, konvexer, (positiver) Kegel in X^* ist.

2.3 Sei $x \in C \subset X$, $C \neq \emptyset$, X BANACHraum und $N_C(x)$ der Normalkegel von C in x. Zeigen Sie:

$$N_C(x) \subset \overline{\bigcup_{\lambda \geq 0} \lambda \partial d_C(x)}^{\,*}.$$

2.4 Sei $x \in C \subset X$, X ein BANACHraum. Zeigen Sie, dass die Menge aller Hypertangenten $HT_C(x)$ zu C bei x offen ist und dass $HT_C(x)$ ein positiver Kegel ist.
Bemerkung: Diese beiden Tatsachen wurden zu Beginn des Beweises von Satz 2.20 benutzt.

2.5 Zeigen Sie: Für eine konvexe abgeschlossene Teilmenge C eines BANACHraums X mit $\mathrm{int}\, C \neq \emptyset$ gilt

$$C = \overline{\mathrm{int}\, C}$$

Bemerkung: Diese Tatsache wurde am Schluss des Beweises von Korollar 2.21 benutzt.

2.6 Seien $f_1, f_2 : X \to \mathbb{R}$. Dann gilt $f_1 = f_2 \iff \mathrm{epi}\, f_1 = \mathrm{epi}\, f_2$.

2.7 Sei $f : X \to \mathbb{R}$ LIPSCHITZ-stetig nahe $x \in X$ mit der DINI-Ableitung (vgl. Definition 2.24 in der Vorlesung)

$$f'_+(x;v) := \liminf_{t \searrow 0} \frac{f(x+tv)-f(x)}{t}.$$

Zeigen Sie (vgl. Lemma 2.25 aus der Vorlesung):

$$\operatorname{epi} f'_+(x;.) \subset K_{\operatorname{epi} f}((x, f(x))).$$

2.8 Seien X_1, X_2 BANACHräume und $C_1 \subset X_1$, $C_2 \subset X_2$, sowie $C := C_1 \times C_2$ mit $x = (x_1, x_2) \in C$. Zeigen Sie:

$$T_C(x) = T_{C_1}(x_1) \times T_{C_2}(x_2) \tag{2.9}$$

$$N_C(x) = N_{C_1}(x_1) \times N_{C_2}(x_2) \tag{2.10}$$

$$K_C(x) \subset K_{C_1}(x_1) \times K_{C_2}(x_2). \tag{2.11}$$

Geben Sie eine allgemeine Bedingung an, die Gleichheit in (2.11) impliziert.
Hinweis: (2.9) und (2.10) sind Inhalt von Corollar 2.11 in der Vorlesung, verwenden Sie dafür Satz 2.10 und Polarisierung. (2.11) mit Gleichheit wurde für den Fall $C_2 = [0, \infty) \subset X_2 := \mathbb{R}$ in Proposition 2.29 für den Beweis der Identität $\partial i_{C_1}(x) = N_{C_1}(x_1)$ benutzt, wobei i_{C_1} die Indikatorfunktion von C_1 bezeichnet, vgl. Definition 2.1 (ii).

3 Endlich-dimensionaler Kalkül

Im Folgenden betrachten wir die Situation $X = \mathbb{R}^n \cong \mathbb{R}^{n\times 1}$, also $X^* \cong \mathbb{R}^{1\times n}$. Ω sei eine offene Teilmenge des $\mathbb{R}^n$. Wir beginnen dieses Kapitel mit dem Zitat des

Satz 3.1 [Satz von RADEMACHER[16]]
Sei $f : \mathbb{R}^n \to \mathbb{R}$ LIPSCHITZ-stetig auf $\Omega \subset \mathbb{R}^n$. DANN ist f fast überall differenzierbar, d. h.

$$\operatorname{Sing} f := \{x \in \Omega \,|\, \nabla f(x) \text{ existiert nicht}\}$$

ist eine Nullmenge bezüglich des n-dimensionalen LEBESGUE[17]-Maßes, kurz $|\operatorname{Sing} f| = 0$.

Einen Beweis findet man z. B. in [Eva98], p. 281.

Satz 3.2
Sei $f : \mathbb{R}^n \to \mathbb{R}$ LIPSCHITZ-stetig nahe $x \in \mathbb{R}^n$ und $S \subset \mathbb{R}^n$ eine Nullmenge. DANN ist

$$\partial f(x) = \operatorname{conv} \left\{ \lim_{i\to\infty} \nabla f(x_i) \,\middle|\, (x_i)_{i\in\mathbb{N}} \subset \mathbb{R}^n \setminus (\operatorname{Sing} f \cup S), x_i \to x \right\}.$$

Beweis. „$\supset$“: Sei $(x_i)_{i\in\mathbb{N}}$ eine gegen x konvergierende Folge mit $x_i \notin \operatorname{Sing} f \cup S$ für alle $i \in \mathbb{N}$, K die LIPSCHITZ-Konstante von f nahe x. Nach Proposition 1.14 ist $\nabla f(x_i) \in \partial f(x_i)$ für alle $i \gg 1$. Wegen $|\nabla f(x_i)| \le K$ für alle $i \gg 1$ gibt es ein $\zeta \in X^* \cong \mathbb{R}^{1\times n}$ und eine Teilfolge $(x_{i_k})_{k\in\mathbb{N}}$ mit $\nabla f(x_i) \to \zeta$. Nach Proposition 1.10 (ii) liegt ζ in $\partial f(x)$. Die Behauptung folgt

[16] HANS RADEMACHER (1892 – 1969) Deutsch-amerikanischer Mathematiker. Nach Studium in Göttingen und Promotion bei CARATHÉODORY war RADEMACHER in Berlin, Hamburg und Breslau tätig. Da er sich als Pazifist ab 1933 bedroht sah, emigrierte er in die U. S. A., wo er ab 1939 Professor in Philadelphia war; später lehrte er in New York. Er arbeitete auf vielen Gebieten der Mathematik und lieferte Beiträge zur Zahlentheorie, Geometrie, Topologie und Funktionentheorie.

[17] HENRY LÉON LEBESGUE (1875 – 1941) Französischer Mathematiker. Stipendien seines Geburtsortes ermöglichten LEBESGUE den Besuch höherer Schulen. Ab 1894 studierte er an der *Ecole Normale*. Trotz umfangreicher Lehraktivitäten am Lyceum von Nancy revolutionierte er, an Arbeiten von BAIRE, BOREL und JORDAN anknüpfend, mit seiner bedeutenden Dissertation *Intégrale, longueur, aire* („Integral, Länge, Fläche“, Nancy, 1902) den Integralbegriff. Sein Begriff der *messbaren Menge* ist eine Synthese von JORDANS Inhaltsdefinition mit BORELS Maßbegriff im $\mathbb{R}^n$. Im Gegensatz zu BORELS Ansatz ist die Menge der LEBESGUE-messbaren Funktionen abgeschlossen unter analytischen Operationen und Grenzprozessen. Die neue Theorie wurde anfangs nur zögernd aufgenommen. Bald erkannte man allerdings deren Tragweite und den Nutzen der in der klassischen Analysis nicht erreichbaren Resultate, wie z. B. die Sätze über monotone und dominierte Konvergenz, die die Vertauschbarkeit von Integrations- und Grenzprozess ermöglichen und der FOURIER-Analysis wichtige Impulse gaben. 1910 erweiterte LEBESGUE seine zunächst eindimensionale Theorie auf den $\mathbb{R}^n$; im selben Jahr wurde er als Lektor an die *Sorbonne* berufen. Im Ersten Weltkrieg befasste er sich als Leiter einer Kommission im Forschungsbeirat des Kriegsministeriums mit ballistischen Fragen. Erst 1919 wurde er zum Professor ernannt und wechselte 1921 an das *Collège de France*, wo er bis zu seinem Lebensende lehrte. Weiterhin beschäftigte er sich mit direkten Methoden der Variationsrechnung und lieferte Beiträge zum Dimensionsbegriff in der Topologie.

jetzt aus der Konvexität von $\partial f(x)$ (Proposition 1.5 (ii)).
„$\subset$": Nach Konstruktion ist $\Lambda := \{\lim_{i\to\infty} \nabla f(x_i) \,|\, (x_i)_{i\in\mathbb{N}} \subset \mathbb{R}^n \setminus (\mathrm{Sing}\, f \cup S), x_i \to x\}$ abgeschlossen und beschränkt, also kompakt. $\partial f(x)$ ist konvex und kompakt nach Proposition 1.5 (ii), (iii). Nach Proposition 1.8 (ii) folgt die Behauptung aus

$$\sigma_{\partial f(x)}(v) = f^0(x;v) \overset{!}{\leq} \sigma_\Lambda(v) \overset{\text{Aufg. 1.11}}{=} \sigma_{\mathrm{conv}\,\Lambda}(v) \qquad \text{für alle } v \in \mathbb{R}^n.$$

Diese Aussage ist direkte Konsequenz nachstehenden Lemmas 3.3. □

Lemma 3.3
Sei $f : \mathbb{R}^n \to \mathbb{R}$ LIPSCHITZ-stetig in der Nähe von $x \in \mathbb{R}^n$, $|S| = 0$. DANN gilt für alle $v \neq 0, \varepsilon > 0$

$$f^0(x;v) - \varepsilon \leq \limsup_{\substack{\bar{y}\to x \\ \bar{y}\notin \mathrm{Sing}\, f \cup S}} \nabla f(\bar{y}) \cdot v.$$

Beweis. Seien $v \neq 0$ und $\varepsilon > 0$ vorgegeben. Wähle $\delta > 0$ so klein, dass f LIPSCHITZ-stetig auf $B_\delta(x)$ und $\nabla f(y) \cdot v \leq \limsup_{\substack{\bar{y}\to x \\ \bar{y}\notin \mathrm{Sing}\, f \cup S}} \nabla f(\bar{y}) \cdot v + \varepsilon$ für alle $y \in B_\delta(x)$ ist. Setze $L_y := \left\{y + tv \,\middle|\, 0 < t < \frac{\delta}{2|v|}\right\}$, dann haben wir $\mathscr{H}^1(L_y \cap (\mathrm{Sing}\, f \cup S)) = 0$ für fast alle $y \in B_{\frac{\delta}{2}}(x)$, vgl. Aufgabe 3.1, wobei $\mathscr{H}^1$ das 1-dimensionale HAUSDORFF-Maß bezeichnet, siehe z.B. [EG92, S. 61].

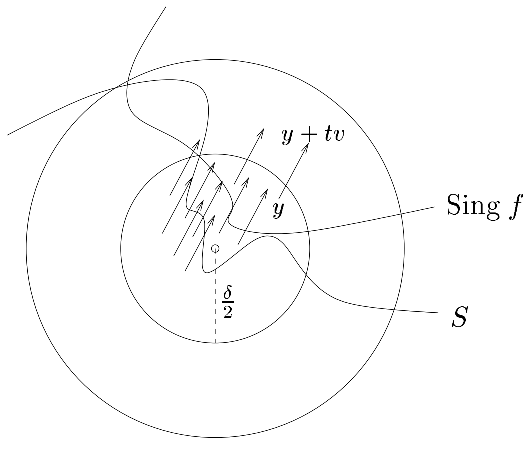

Abbildung 3.1: Zum Beweis von Lemma 3.3

Fast alle $y \in B_{\frac{\delta}{2}}(x)$ erfüllen also für jedes $t \in (0, \frac{\delta}{2|v|})$

$$f(y+tv) - f(y) = \int_0^t \nabla f(y+sv) \cdot v \,\mathrm{d}s \leq t \left(\limsup_{\substack{\bar{y}\to x \\ \bar{y}\notin \mathrm{Sing}\, f \cup S}} \nabla f(\bar{y}) \cdot v + \varepsilon \right).$$

Wegen der LIPSCHITZ-Stetigkeit gilt diese Aussage sogar für *alle* $y \in B_{\frac{\delta}{2}}(x)$. Hieraus schließen wir $f^0(x;v) = \limsup_{\substack{y\to x \\ t\searrow 0}} \frac{f(y+tv)-f(y)}{t} \leq \limsup_{\substack{\bar{y}\to x \\ \bar{y}\notin \mathrm{Sing}\, f \cup S}} \nabla f(\bar{y}) \cdot v + \varepsilon.$ □

Als direkte Konsequenz aus Satz 3.2 ergibt sich die Gleicheit der Stützfunktionen:

Corollar 3.4
Sei $f : \mathbb{R}^n \to \mathbb{R}$ LIPSCHITZ-stetig in der Nähe von $x \in \mathbb{R}^n$, $|S| = 0$. DANN ist

$$f^0(x; v) = \limsup_{\substack{\bar{y} \to x \\ \bar{y} \notin \operatorname{Sing} f \cup S}} \nabla f(\bar{y}) \cdot v \qquad \text{für alle } v \in \mathbb{R}^n.$$

Beispiel 3.5 Wir untersuchen die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$,

$$\begin{aligned} (x, y) &\mapsto \max\{\min\{x, -y\}, y - x\} \\ &= \left\{ \begin{array}{ccccl} x & \text{auf} & C_1 & := & \left\{(x,y) \in \mathbb{R}^2 \,\middle|\, y \leq 2x \text{ und } y \leq -x\right\} \\ -y & \text{auf} & C_2 & := & \left\{(x,y) \in \mathbb{R}^2 \,\middle|\, y \leq \frac{x}{2} \text{ und } y \geq -x\right\} \\ y - x & \text{auf} & C_3 & := & \left\{(x,y) \in \mathbb{R}^2 \,\middle|\, y \geq 2x \text{ oder } y \geq \frac{x}{2}\right\} \end{array} \right\}. \end{aligned}$$

Wegen $C_1 \cup C_2 \cup C_3 = \mathbb{R}^2$ und

$$Df(x,y) = \left\{ \begin{array}{ll} \binom{1}{0} & \text{in } \mathring{C}_1 \\ \binom{0}{-1} & \text{in } \mathring{C}_2 \\ \binom{-1}{1} & \text{in } \mathring{C}_3 \end{array} \right. \quad \text{liefert Satz 3.2, dass} \quad \partial f(0,0) = \operatorname{conv} \left\{ \tbinom{1}{0}, \tbinom{0}{-1}, \tbinom{-1}{1} \right\}.$$

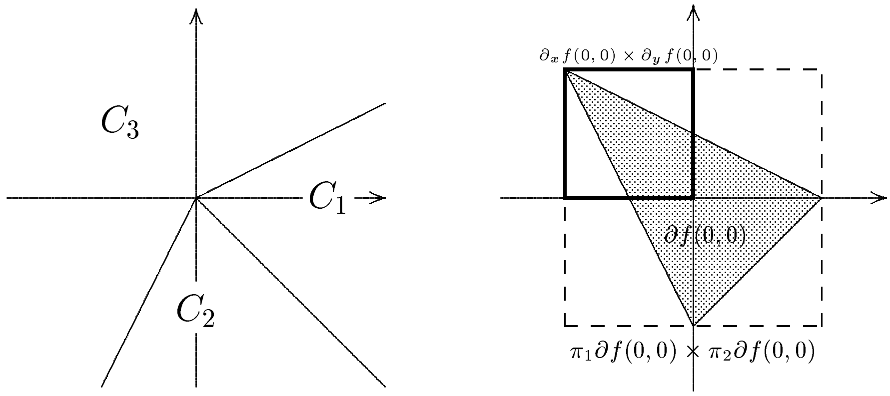

Abbildung 3.2: Zu Beispiel 3.5

Im Hinblick auf die in Propositionen 1.43 und 1.44 betrachtete Situation liegt es nun nahe, nach der Beziehung von $\partial f(0,0)$ zu $\partial_x f(0,0) \times \partial_y f(0,0)$ zu fragen. Wir haben $f(x,0) = \max\{0, -x\}$, also $\partial_x f(0,0) = [-1, 0]$, und $f(0,y) = \max\{0, y\}$, also $\partial_y f(0,0) = [0,1]$. Damit ist

$$\partial_x f(0,0) \times \partial_y f(0,0) \not\subset \partial f(0,0) \not\subset \partial_x f(0,0) \times \partial_y f(0,0).$$

Proposition 1.43 kann nicht angewendet werden, weil f nicht regulär in $(0,0)$ ist. Proposition 1.44 liefert immerhin $\partial_x f(0,0) \times \partial_y f(0,0) \subset \pi_1 \partial f(0,0) \times \pi_2 \partial f(0,0) = [-1,1]^2$. $\diamond$

Wir erweitern unseren Kalkül nunmehr auf vektorwertige Funktionen. Es läge nahe, hier die Menge $\left\{M \in \mathbb{R}^{m \times n} \,\middle|\, \text{für alle } i = 1, \ldots, m \text{ liegt die } i\text{-te Zeile von M in } \partial F^i(x)\right\}$ zu betrachten. Wir wählen jedoch eine restriktivere Definition.

Definition 3.6 [Verallgemeinerte JACOBIsche]
Sei $F : \mathbb{R}^n \to \mathbb{R}^m$ eine nahe $x \in \mathbb{R}^n$ LIPSCHITZ-*stetige* Funktion, d. h. es gibt ein $K > 0$ und eine Umgebung Ω von x, so dass

$$|F(\xi) - F(\eta)|_{\mathbb{R}^m} \leq K \, |\xi - \eta|_{\mathbb{R}^n} \qquad \forall\, \xi, \eta \in \Omega.$$

Dann existiert $DF(\xi)$ auf Ω bis auf eine Ausnahmemenge $\operatorname{Sing} F$ vom Maße Null, und wir bezeichnen die Menge

$$\partial F(x) := \operatorname{conv} \left\{ \lim_{i\to\infty} DF(x_i) \,\middle|\, (x_i)_{i\in\mathbb{N}} \subset \Omega \setminus \operatorname{Sing} F, x_i \to x \right\}$$

als *verallgemeinerte* JACOBI*sche* [18] von F bei x.

In [War81] wird gezeigt, dass die Konstruktion

$$\partial_S F(x) := \operatorname{conv} \left\{ \lim_{i\to\infty} DF(x_i) \,\middle|\, (x_i)_{i\in\mathbb{N}} \subset \Omega \setminus (\operatorname{Sing} F \cup S), x_i \to x \right\}$$

für eine beliebige Nullmenge S auf dieselbe verallgemeinerte JACOBIsche führt: $\partial F(x) = \partial_S F(x)$.
Wir führen nun einige elementare Eigenschaften der verallgemeinerten JACOBIschen an.

Proposition 3.7
Sei $F : \mathbb{R}^n \to \mathbb{R}^m$ eine nahe $x \in \mathbb{R}^n$ LIPSCHITZ*-stetige Funktion. DANN gilt:*

(i) $\partial F(x)$ *ist eine nicht-leere, konvexe und kompakte Teilmenge des* $\mathbb{R}^{m\times n}$.

(ii) $\partial F(.)$ *ist* „abgeschlossen bei x“, *d.h. für alle Folgen $x_i \to x$ und $(z_i)_{i\in\mathbb{N}} \subset \mathbb{R}^{m\times n}$, $z_i \in \partial F(x_i)$ mit $z_i \to z$, gilt $z \in \partial F(x)$.*

(iii) $\partial F(.)$ *ist oberhalbstetig im Sinne von Definition 1.9 (iii), d.h.*

$$\forall\, \varepsilon > 0\ \exists\, \delta > 0 : \partial F(B_\delta(x)) \subset B_\varepsilon(\partial F(x)) \subset \mathbb{R}^{m\times n}.$$

(iv) Seien $f^1, \ldots, f^m$ reellwertige und nahe $x \in \mathbb{R}^n$ LIPSCHITZ*-stetige Funktionen mit* LIPSCHITZ*-Konstanten $K_1, \ldots K_m$ und $F := \begin{pmatrix} f^1 \\ \vdots \\ f^m \end{pmatrix}$. DANN ist F* LIPSCHITZ*-stetig nahe x mit* LIPSCHITZ*-Konstante $K := \sqrt{K_1^2 + \cdots + K_m^2}$, und es gilt $\partial F(x) \subset B_K(0) \subset \mathbb{R}^{m\times n}$.*

(v) Für F wie in (iv) gilt:

$$\partial F(x) \subset \left\{ M \in \mathbb{R}^{m\times n} \,\middle|\, \text{für alle } i = 1, \ldots, m \text{ liegt die } i\text{-te Zeile von M in } \partial f^i(x) \right\}$$

Im Fall $m = 1$ gilt Gleichheit der Mengen.

[18] CARL GUSTAV JACOB JACOBI (1804 – 1851) Deutscher Mathematiker. Studium der Philosophie, Philologie und Mathematik (zunächst autodidaktisch) in Berlin, 1825 Promotion und Habilitation. Ab 1826 in Königsberg, 1834 Gründung eines mathematisch-physikalischen Seminars. Ab 1844 arbeitete er wieder in Berlin. JACOBI war ein außerordentlich produktiver Forscher; er beschäftigte sich mit elliptischen Funktionen, Zahlentheorie, analytischer Mechanik, partiellen Differentialgleichungen, Geometrie und Algebra. Mit seiner Königsberger Schule übte er großen Einfluss auf andere Universitäten aus. JACOBI trat auch als Musiker in Erscheinung und betätigte sich als politischer Redner während der Revolution 1848/49, was ihm finanzielle Repressalien einbrachte.

Beweisskizze. (i) folgt wie zu Beginn des Beweises von Satz 3.2: Für eine Folge $x_i \to x$, $x_i \notin \operatorname{Sing} F$, ist $DF(x_i) \in \partial F(x_i)$ beschränkt. Somit gibt es eine konvergente Teilfolge.
(ii) folgt aus (iii), denn wenn wir für alle $\varepsilon > 0$ ein $i_0 \in \mathbb{N}$ haben, so dass $z_i \in B_\varepsilon(\partial F(x))$ für alle $i \geq i_0$, muss z in der (abgeschlossenen) Menge $\partial F(x)$ liegen.
(iii) Angenommen, es gäbe ein $\varepsilon > 0$, so dass wir für alle $i \in \mathbb{N}$ ein $y_i \in B_{\frac{1}{i}}(x) \setminus \operatorname{Sing} F$ mit $DF(y_i) \notin B_\varepsilon(\partial F(x))$ finden könnten. Dann hätten wir $y_i \to x$, also $|DF(y_i)| \leq K$. Hieraus folgte sofort die Konvergenz einer Teilfolge gegen ein $z \in DF(x)$ nach Definition 3.6. ↯
(iv) prüft man sofort nach mit Hilfe von Definition 3.6.
(v) folgt aus Definition 3.6 und Satz 3.2. □

Proposition 3.8
Sei $F : \mathbb{R}^n \to \mathbb{R}^m$ eine nahe $x \in \mathbb{R}^n$ LIPSCHITZ-stetige Funktion und Σ eine Nullmenge im $\mathbb{R}^n$. DANN gilt für beliebige $v \in \mathbb{R}^n \cong \mathbb{R}^{n\times 1}$, $w \in \mathbb{R}^m \cong \mathbb{R}^{1\times m}$

$$\partial F(x) \cdot v = \partial_\Sigma F(x) \cdot v \qquad \textit{und} \qquad w \cdot \partial F(x) = w \cdot \partial_\Sigma F(x).$$

Bemerkung. Mit dem oben zitierten Resultat von WARGA [War81] folgt dieses schwächere Resultat sofort, aber unabhängig davon gilt der folgende

Beweis. Die Voraussetzungen von Proposition 1.8 (ii) sind für die betreffenden Mengen in $\mathbb{R}^{m\times 1}$ bzw. $\mathbb{R}^{1\times n}$ erfüllt. Zu zeigen ist $\sigma_{\partial F(x)\cdot v}(u) = \sigma_{\partial_S F(x)\cdot v}(u)$ für alle $u \in \mathbb{R}^m \cong \mathbb{R}^{1\times m}$. Mit $g_u(y) := F(y) \cdot u$ gilt für jedes $u \in \mathbb{R}^m$ $\operatorname{Sing} g_u \subset \operatorname{Sing} F$ und damit

$$\begin{aligned}
\sigma_{\partial F(x)\cdot v}(u) &= \sup_{\zeta \in \partial F(x)\cdot v} \langle \zeta, u\rangle &&= \limsup_{\substack{y\to x\\ y\notin \operatorname{Sing} F}} \langle DF(y)\cdot v, u\rangle \\
&= \limsup_{\substack{y\to x\\ y\notin \operatorname{Sing} F}} \left\langle v \cdot DF(y)^\top, u\right\rangle &&= \limsup_{\substack{y\to x\\ y\notin \operatorname{Sing} F}} \langle v, DF(y)\cdot u\rangle \\
&= \limsup_{\substack{y\to x\\ y\notin \operatorname{Sing} F}} \langle v, \nabla g_u(y)\rangle &&\overset{\substack{\text{Cor. 3.4}\\ \text{für } S := \operatorname{Sing} F \setminus \operatorname{Sing} g_u}}{=} g_u^0(x, v) \\
&\overset{\substack{\text{Cor. 3.4}\\ \text{für } S := \Sigma}}{=} \limsup_{\substack{y\to x\\ y\notin \operatorname{Sing} F\cup\Sigma}} \langle v, \nabla g_u(y)\rangle &&= \limsup_{\substack{y\to x\\ y\notin \operatorname{Sing} F\cup\Sigma}} \langle v, DF(y)\cdot u\rangle \\
&= \limsup_{\substack{y\to x\\ y\notin \operatorname{Sing} F\cup\Sigma}} \left\langle v \cdot DF(y)^\top, u\right\rangle &&= \limsup_{\substack{y\to x\\ y\notin \operatorname{Sing} F\cup\Sigma}} \langle DF(y)\cdot v, u\rangle \\
&= \sup_{\zeta \in \partial_\Sigma F(x)\cdot v} \langle \zeta, u\rangle &&= \sigma_{\partial_\Sigma F(x)\cdot v}(u).
\end{aligned}$$

Der Beweis für die zweite Aussage verläuft analog. □

Proposition 3.9 [Vektorieller Mittelwertsatz]
Sei U eine offene Menge im $\mathbb{R}^n$ und $F : U \to \mathbb{R}^m$ LIPSCHITZ-stetig auf U. DANN gilt für $x, y \in U$ mit $[x,y] \subset U$

$$F(y) - F(x) \in \operatorname{conv} \partial F([x,y]) \cdot (y - x),$$

d. h. $F(y) - F(x)$ ist Konvexkombination von Elementen der Form $Z \cdot (y - x)$, wobei $Z \in \partial F(u)$ für ein $u \in [x,y]$.

Beweisskizze. Sei x fest vorgegeben. Dann gilt $\mathscr{H}^1([x,y] \cap \operatorname{Sing} F) = 0$ für fast alle $y \in U$, vgl. Aufgabe 3.1. Wir können also

$$F(y) - F(x) = \underbrace{\int_0^1 \underbrace{DF(x + t(y-x))}_{\in \partial F([x,y])}(y-x)\,\mathrm{d}t}_{\overset{\text{Aufg. 3.3}}{\in} \operatorname{conv} \partial F([x,y])\cdot(y-x)}$$

schreiben, und für allgemeine y folgt die Behauptung aus der Stetigkeit von F und der Oberhalbstetigkeit von ∂F. □

Satz 3.10 [Kettenregel III]
Sei $F : \mathbb{R}^n \to \mathbb{R}^m$ LIPSCHITZ-*stetig nahe* $x \in \mathbb{R}^n$, $g : \mathbb{R}^m \to \mathbb{R}$ LIPSCHITZ-*stetig nahe* $F(x) \in \mathbb{R}^m$ *und* $f := g \circ F : \mathbb{R}^n \to \mathbb{R}$. *DANN gilt*

$$\partial f(x) \subset \operatorname{conv} \partial g(F(x)) \circ \partial F(x).$$

Ist g strikt differenzierbar, so haben wir

$$\partial f(x) = D_s g(F(x)) \circ \partial F(x).$$

Beweis. Für die Inklusion haben wir

$$\forall\, v \in \mathbb{R}^n\ \exists\, \zeta_0 \in \partial g(F(x)), \eta_0 \in \partial F(x) : f^0(x;v) \le \zeta_0 \eta_0 v \quad \Big(\le \sup_{\xi \in \partial g(F(x)) \circ \partial F(x)} \xi v \Big)$$

zu zeigen, denn dann gilt nach Aufgabe 1.11

$$f^0(x;v) \le \sigma_{\operatorname{conv} \partial g(F(x)) \circ \partial F(x)}(v) = \sigma_{\partial g(F(x)) \circ \partial F(x)}(v).$$

Tatsächlich haben wir

$$\begin{aligned}
\frac{f(y+tv) - f(y)}{t} &= \frac{g(F(y+tv)) - g(F(y))}{t} \\
&\overset{\text{Satz 1.33}}{=} \frac{\zeta_t(F(y+tv) - F(y))}{t} \qquad \begin{array}{r} \text{für ein} \\ \text{mit} \end{array} \begin{array}{l} \zeta_t \in \partial g(u_t) \\ u_t \in [F(y), F(y+tv)] \end{array} \\
&\overset{\text{Prop. 3.9}}{=} \frac{\zeta_t}{t} w_t \cdot tv \qquad \text{für ein } w_t \in \operatorname{conv} \partial F([y, y+tv]) \\
&\le \zeta_t \eta_t^{k_{\max}} v,
\end{aligned}$$

wobei $w_t = \sum_{k=1}^P \lambda_k \eta_t^k$ mit $\eta_t^k \in \partial F([y, y+tv])$, $\lambda_k \ge 0$, $\sum_{k=1}^P \lambda_k = 1$ und $k_{\max} \in \{1, \ldots, P\}$ so gewählt sei, dass $\zeta_t w_t \le \sum_{k=1}^P \lambda_k \zeta_t \eta_t^{k_{\max}} = \zeta_t \eta_t^{k_{\max}}$. Wir betrachten nun Folgen $y_i \to x$, $t_i \searrow 0$, die $\frac{f(y_i + t_i v) - f(y_i)}{t_i} = f^0(x;v)$ realisieren. Obige Rechnung liefert uns zu jedem $i \gg 1$ ein $u_{t_i} \in [F(y_i), F(y_i + t_i v)] \to \{F(x)\}$, ein $\zeta_{t_i} \in \partial g(u_{t_i})$ mit $|\zeta_{t_i}| \le K_g$ (LIPSCHITZ-Konstante von g) und ein $\eta_{t_i}^{k_{\max}} \in \partial F([y_i, y_i + t_i v])$ mit $\left|\eta_{t_i}^{k_{\max}}\right| \le K_F$ (LIPSCHITZ-Konstante von F). Also gibt es eine gegen ein ζ_0 konvergierende Teilfolge, und ζ_0 liegt nach Proposition 1.10 in $\partial g(F(x))$; ebenso gibt es eine gegen ein η_0 konvergierende Teilfolge, wobei η_0 nach Proposition 3.7 (ii) in $\partial F(x)$ liegt. Damit ist $f^0(x;v) \le \zeta_0 \eta_0 v$, und die Inklusion ist bewiesen.

Zur Gleichheitsdiskussion merken wir an, dass nach Proposition 1.15 die strikte Differenzierbarkeit von g die Identität $\partial g(F(x)) = \{D_s g(F(x))\}$ garantiert. Damit ergibt sich für jedes $v \in \mathbb{R}^n$

$$\begin{array}{rlcl}
 & \sigma_{\mathrm{conv}\, D_s g(F(x))\partial F(x)}(v) & \overset{\text{Aufg. 1.11}}{=} & \sigma_{D_s g(F(x))\partial F(x)}(v) \\
= & \sup\limits_{\eta\in\partial F(x)} \langle D_s g(F(x))\eta, v\rangle & \overset{\text{Def. 3.6}}{=} & \limsup\limits_{\substack{y\to x\\ y\notin \mathrm{Sing}\, F}} D_s g(F(x))DF(y)\cdot v \\
\overset{\text{Prop. 3.8}}{=} & \limsup\limits_{\substack{y\to x\\ y\notin \mathrm{Sing}\, F\cup\mathrm{Sing}\, f}} D_s g(F(x))DF(y)\cdot v & \overset{!}{\leq} & \limsup\limits_{\substack{y\to x\\ y\notin \mathrm{Sing}\, F\cup\mathrm{Sing}\, f}} \nabla f(y)\cdot v \\
\overset{\text{Cor. 3.4}}{=} & f^0(x;v) & = & \sigma_{\partial f(x)},
\end{array}$$

wobei die markierte Ungleichung folgendermaßen zustande kommt: Angenommen, diese Ungleichung wäre falsch. Dann gäbe es ein $v \in \mathbb{S}^{n-1}$ mit

$$\varepsilon := \frac{1}{2}\left(\limsup_{\substack{y\to x\\ y\notin \mathrm{Sing}\, F\cup\mathrm{Sing}\, f}} D_s g(F(x))DF(y)\cdot v - \limsup_{\substack{y\to x\\ y\notin \mathrm{Sing}\, F\cup\mathrm{Sing}\, f}} \nabla f(y)\cdot v\right) > 0.$$

Lemma 3.11 liefert uns ein $\delta > 0$ mit

$$|D_s g(F(x))DF(\xi) - \nabla f(\xi)| < \frac{\varepsilon}{2} \tag{3.1}$$

für jedes $\xi \in B_\delta(x) \setminus (\mathrm{Sing}\, F \cup \mathrm{Sing}\, f)$. Wegen der Oberhalbstetigkeit von ∂F gibt es ein $\tilde{\delta} > 0$ mit

$$|DF(\xi) - DF(x)| < \frac{\varepsilon}{4\,\|D_s g(F(x))\|} \tag{3.2}$$

für alle $\xi \in B_{\tilde{\delta}}(x)$. Nach Definition von ε könnten wir $y_1, y_2 \in B_\delta(x) \cap B_{\tilde{\delta}}(x) \setminus (\mathrm{Sing} F \cup \mathrm{Sing} f)$ mit

$$D_s g(F(x))DF(y_1)\cdot v - \nabla f(y_2)\cdot v > \varepsilon$$

finden. Dies stünde im Widerspruch zu

$$\begin{array}{rl}
\multicolumn{2}{l}{|D_s g(F(x))DF(y_1)\cdot v - \nabla f(y_2)\cdot v|} \\
\leq & \underbrace{|D_s g(F(x))DF(y_1)\cdot v - D_s g(F(x))DF(y_2)\cdot v|}_{\leq \|D_s g(F(x))\|(|DF(y_1)-DF(x)|+|DF(x)-DF(y_2)|)} + |D_s g(F(x))DF(y_2)\cdot v - \nabla f(y_2)\cdot v| \\
< & \varepsilon \quad \text{nach (3.1) und (3.2).}
\end{array}$$

□

Lemma 3.11
Unter den Voraussetzungen von Satz 3.10 gibt es bei strikter Differenzierbarkeit von g zu jedem $0 < \varepsilon \ll 1$ ein $\delta > 0$ mit

$$\nabla f(y) \subset B_\varepsilon(D_s g(F(x))DF(y)) \qquad \text{für alle } y \in B_\delta(x) \setminus (\mathrm{Sing}\, F \cup \mathrm{Sing}\, f).$$

Beweis. Wir können ohne Einschränkung $K_F \geq 1$ annehmen. Wähle ein $\delta > 0$ so klein, dass F Lipschitz-stetig auf $B_\delta(x)$, g Lipschitz-stetig nahe $F(y)$ für alle $y \in B_\delta(x)$ und $\partial g(F(B_\delta(x))) \subset B_{\frac{\varepsilon}{K_F}}(D_s g(F(x)))$ ist; letzteres wird durch die Oberhalbstetigkeit von $\partial g(.)$

und die LIPSCHITZ-Stetigkeit von F ermöglicht, vgl. Proposition 1.10 (iii). Wir behaupten nun

$$\nabla f(y) \subset \partial g(F(y))DF(y) \qquad \forall\, y \in B_\delta(x) \setminus (\operatorname{Sing} F \cup \operatorname{Sing} f).$$

In der Tat gilt für jedes $v \in \mathbb{R}^n$ nach dem Mittelwertsatz, Satz 1.33,

$$\nabla f(y) \cdot v = \lim_{t \searrow 0} \frac{g(F(y+tv)) - g(F(y))}{t} = \lim_{t \searrow 0} \frac{\zeta_t(F(y+tv) - F(y))}{t} = \zeta_0 DF(y) \cdot v,$$

wobei $\zeta_t \in \partial g(u_t)$, $u_t \in [F(y), F(y+tv)]$ und $\zeta_0 \in \partial g(F(y))$. Für die letzte Gleichheit haben wir die Kompaktheit und die Propositionen 1.10 und 3.7 benutzt. Damit gilt

$$\nabla f(y) \cdot v \in \partial g(F(y))DF(y) \cdot v \quad \forall\, v \in \mathbb{R}^n,$$

woraus folgt $\nabla f(y) \in \partial g(F(y))DF(y)$. Wäre nämlich $\nabla f(y)$ nicht in der konvexen Menge $\partial g(F(y))DF(y)$ enthalten, so gäbe es nach dem Trennungssatz für konvexe Mengen ([Alt02, S. 223]) ein $w \in \mathbb{R}^n$ mit $\nabla f(y) \cdot w > \langle \eta, w \rangle$ für alle $\eta \in \partial g(F(y))DF(y)$. $\lightning$
Also

$$\nabla f(y) \in B_{\frac{\varepsilon}{K_F}}(D_s g(F(x)))DF(y) \subset B_\varepsilon(D_s g(F(x)))DF(y).$$

$\square$

Corollar 3.12
Sei $F : \mathbb{R}^n \to \mathbb{R}^m$ LIPSCHITZ-stetig nahe $x \in \mathbb{R}^n$, $G : \mathbb{R}^m \to \mathbb{R}^k$ LIPSCHITZ-stetig nahe $F(x) \in \mathbb{R}^m$. DANN gilt

$$\partial(G \circ F)(x) \cdot v \subset \operatorname{conv} \partial G(F(x)) \circ \partial F(x) \cdot v \qquad \text{für alle } v \in \mathbb{R}^n.$$

Ist G stetig differenzierbar in der Nähe von $F(x)$, so gilt

$$\partial(G \circ F)(x) \cdot v = DG(F(x)) \circ \partial F(x) \cdot v \qquad \text{für alle } v \in \mathbb{R}^n.$$

Beweis. Sei $\alpha \in \mathbb{R}^k$ vorgegeben. Dann ist die Abbildung $z \mapsto \langle \alpha, z \rangle$ glatt und besitzt die konstante Ableitung $z \mapsto \alpha^\top$. Wir haben

$$\begin{aligned}
\langle \alpha, \partial(G \circ F)(x) \cdot v \rangle &\overset{\substack{\text{Satz 3.10} \\ \text{Gleichheit}}}{=} \partial(\langle \alpha, G \circ F \rangle)(x) \cdot v \\
&\overset{\text{Satz 3.10}}{\subset} \operatorname{conv} \partial(\langle \alpha, G \rangle)(F(x))\partial F(x) \cdot v \\
&\overset{\text{Satz 3.10}}{\subset} \operatorname{conv} \langle \alpha, \partial G(F(x))\partial F(x) \cdot v \rangle \\
&\subset \langle \alpha, \operatorname{conv} \partial G(F(x))\partial F(x) \cdot v \rangle.
\end{aligned}$$

Damit ergibt sich die behauptete Inklusion mit Aufgabe 3.4.

Ist G nun sogar stetig differenzierbar, so gilt

$$
\begin{aligned}
&DG(F(x))\partial F(x)\\
\overset{\text{Def. 3.6}}{=}\quad & DG(F(x))\mathrm{conv}\left\{\lim_{i\to\infty} DF(x_i)\,\middle|\, (x_i)_{i\in\mathbb{N}}\subset\mathbb{R}^n\setminus \operatorname{Sing} F \text{ mit } x_i\to x\right\}\\
=\quad & \mathrm{conv}\left\{DG(F(x))\lim_{i\to\infty} DF(x_i)\,\middle|\, (x_i)_{i\in\mathbb{N}}\subset\mathbb{R}^n\setminus \operatorname{Sing} F \text{ mit } x_i\to x\right\}\\
\overset{G\in C^1}{=}\quad & \mathrm{conv}\left\{\lim_{i\to\infty} DG(F(x_i))DF(x_i)\,\middle|\, (x_i)_{i\in\mathbb{N}}\subset\mathbb{R}^n\setminus \operatorname{Sing} F \text{ mit } x_i\to x\right\}\\
\overset{\substack{\operatorname{Sing} F\supset\\ \operatorname{Sing} G\circ F}}{\subset}\quad & \mathrm{conv}\left\{\lim_{i\to\infty} DG(F(x_i))DF(x_i)\,\middle|\, (x_i)_{i\in\mathbb{N}}\subset\mathbb{R}^n\setminus \operatorname{Sing} G\circ F \text{ mit } x_i\to x\right\}\\
\overset{\text{Def. 3.6}}{=}\quad & \partial(G\circ F)(x) \overset{\text{s. o.}}{\subset} \mathrm{conv}\ \underbrace{DG(F(x))\,\underbrace{\partial F(x)}_{\text{konvex}}}_{\text{konvex}}\\
=\quad & DG(F(x))\partial F(x).
\end{aligned}
$$

Dies beweist die Gleichheit. □

Aufgaben

3.1 Sei $S\subset\mathbb{R}^n$ mit $|S|=0$, $v\in S^{n-1}$. Zeigen Sie:

$$\mathscr{H}^1(L_y\cap S)=0$$

für fast alle $y\in B_1(0)\subset\mathbb{R}^n$, wobei

$$L_y := \{y+tv : 0<t<1\}.$$

Bemerkung: Eine Aussage dieses Typs wurde im Beweis von Lemma 3.3 und Proposition 3.9 in der Vorlesung benutzt.

3.2 Die Funktion $f:\mathbb{R}^2\to\mathbb{R}$ sei definiert durch

$$f(x,y) := 5x\chi_{\{y>x\}}\chi_{\{x>0\}} - 4x\chi_{\{x\le 0\}} + 5|y|\chi_{\{x>y\}}\chi_{\{-x<y\}} + 5x\chi_{\{-x>y\}}\chi_{\{x>0\}}.$$

Berechnen Sie $\partial f(0,0)$.

3.3 Sei $C\subset\mathbb{R}^n$ abgeschlossen, $-\infty<a<b<+\infty$. Sei $f\in L^1([a,b],\mathbb{R}^n)$ mit $f(t)\in C$ für fast alle $t\in[a,b]$. Zeigen Sie:

$$\frac{1}{b-a}\int_a^b f(t)\,dt\in\mathrm{conv}\,C.$$

Ist die Abgeschlossenheit von C erforderlich?
Bemerkung: Eine solche Aussage wurde bei dem Beweis der vektoriellen Mittelwertinklusion Proposition 3.9 benutzt.

3.4 Seien $F:\mathbb{R}^n\to\mathbb{R}^m$ Lipschitz nahe $x\in\mathbb{R}^n$ und $G:\mathbb{R}^m\to\mathbb{R}^k$ Lipschitz nahe $F(x)$ und $v\in\mathbb{R}^n$ gegeben. Zeigen Sie, dass aus

$$\alpha^\top\partial(G\circ F)(x)v\subset\alpha^\top\mathrm{conv}\,\{\partial G(F(x))\partial F(x)v\}\ \text{ für alle }\ \alpha\in\mathbb{R}^k$$

die Aussage

$$\partial(G\circ F)(x)v\subset\mathrm{conv}\,\{\partial G(F(x))\partial F(x)v\}$$

folgt. *Diese Tatsache wurde für den Beweis von Corollar 3.12 benutzt.*

4 Integral- und Maximalfunktionen

Integralfunktionen

Zentrales Hilfsmittel für einige Beweise aus diesem Abschnitt sind der Satz 1 aus Paragraph 2 und Satz 1 aus Paragraph 3 von [IL74], die wir hier in einer schwächeren Version zitieren.

Proposition 4.1 [Ioffe, Levin 1972]
Sei $(T, \mathscr{T}, \mu)$ *ein positiver Maßraum,* Ω *eine offene, konvexe Teilmenge eines* $\mathbb{R}$-Banach*raumes* X, $0 \in \Omega$. *Weiterhin gebe es für jedes* $t \in T$ *eine Abbildung* $\hat{u}_t : \Omega \to \mathbb{R}$ *mit*

- $\hat{u}_t$ *konvex für alle* $t \in T$,
- $\hat{u}_t$ *stetig in* 0 *für alle* $t \in T$ *und*
- $\Big(t \mapsto \hat{u}_t(z)\Big) \in L^1(\mu)$ *für alle* $z \in \Omega$.

Ist außerdem eine der Bedingungen

a) T *abzählbar,*

b) X *separabel,*

c)* T *separabel, metrisch,* μ *regulär , (vgl. Definition A.15) und* $t \mapsto \hat{u}_t(z)$ *oberhalbstetig für alle* $z \in \Omega$ *sowie* $\hat{u}$ *stetig in* 0

erfüllt, DANN ist $\hat{u} : \Omega \to \mathbb{R}$, $z \mapsto \int_T \hat{u}_t(z)\,\mathrm{d}\mu(t)$ *stetig in* 0, *und es gilt*

$$\partial \hat{u}(0) = \int_T \partial \hat{u}_t(0)\,\mathrm{d}\mu(t). \tag{4.1}$$

Bemerkung. Mit der Notation (4.1) beschreiben wir abkürzend die Situation, dass es zu jedem $\zeta \in \partial \hat{u}(0)$ eine Abbildung $T \to X^*$, $t \mapsto \zeta_t$ mit $\zeta_t \in \partial \hat{u}_t(0)$ für μ-fast alle $t \in T$ gibt, die

- $\forall\, v \in X : \Big(t \mapsto \langle \zeta_t, v\rangle_{X^* \times X}\Big) \in L^1(\mu)$ und
- $\langle \zeta, v\rangle_{X^* \times X} = \int_T \langle \zeta_t, v\rangle_{X^* \times X}\,\mathrm{d}\mu(t)$

erfüllt.

In Verallgemeinerung dieses Resultats beweisen wir für den Fall nicht notwendig konvexer, aber Lipschitz-stetiger Funktionen den

Satz 4.2
Sei $(T, \mathscr{T}, \mu)$ *ein positiver Maßraum,* U *eine offene Teilmenge eines* $\mathbb{R}$-BANACH-*Raumes* X. *Weiterhin sei zu jedem* $t \in T$ *eine Abbildung* $f_t : U \to \mathbb{R}$ *mit den folgenden beiden Eigenschaften gegeben.*

- *Für jedes* $x \in U$ *sei die Abbildung* $t \mapsto f_t(x)$ *messbar.*
- *Es gibt ein* $k \in L^1(\mu)$, *so dass für* μ*-fast alle* $t \in T$ *und alle* $x, y \in U$ *die Abschätzung* $|f_t(x) - f_t(y)| \leq k(t) \, \|x - y\|_X$ *erfüllen.*

(i) Die reellwertige Abbildung $f : x \mapsto \int_T f_t(x) \, \mathrm{d}\mu(t)$ *sei an der Stelle* $\bar{x} \in U$ *definiert. DANN ist* f *überall auf* U *definiert und* LIPSCHITZ-*stetig auf* U.

(ii) Falls außerdem

a) T *abzählbar, ODER*

b) X *separabel, ODER*

c) T *separabel, metrisch,* μ *regulär (vgl. Definition A.15) und* $T \to \mathfrak{P}(X^*)$, $t \mapsto \partial f_t(y)$ *oberhalbstetig (bzgl. der schwach*$*$*-Topologie in* X^**) für alle* $y \in U$

gilt, DANN haben wir

$$\partial f(x) \subset \int_T \partial f_t(x) \, \mathrm{d}\mu(t) \tag{4.2}$$

für $x \in U$.

(iii) Wenn zusätzlich sogar $f_t(.)$ *für alle* $t \in T$ *regulär bei* x *sind, so ist* f *regulär bei* x, *und es gilt Gleichheit in* (4.2).

Insbesondere wird diese Situation im Hinblick auf die Variationsrechnung interessant, wenn die LAGRANGE[19]-Funktionen nur noch LIPSCHITZ-stetig sind, vgl. (4.5).
Die Definitionen hier verwandter maßtheoretischer Begriffe findet man in [Alt02, p. 42], für die Definition der Regularität von Maßen verweisen wir auf Definition A.15 im Anhang.
Nachstehender Beweis ließe sich alternativ — ohne konvexe Analysis — mit einem Minimax-Prinzip führen, vgl. [Aub79, p. 396].

[19] JOSEPH LOUIS LAGRANGE (1736 – 1813) Französischer Mathematiker, Physiker, Astronom. Beginn der mathematischen Laufbahn mit 16 Jahren, Professor mit 19 Jahren. Ausgehend von Arbeiten EULERS zur Variationsrechnung führte er den Begriff der *ersten Variation* ein und löste das *isoperimetrische Problem.* Später schuf er eine allgemeine Methode zur Lösung von Extremalproblemen mit Nebenbedingungen. Ab 1766 arbeitete er in Berlin auf dem Gebiet der Differentialgleichungen. Weiterhin beschäftigte er sich mit der Potenzreihendarstellung von Funktionen und der Lösbarkeit algebraischer Gleichungen. 1786 kehrte er nach Frankreich zurück und wurde Mitglied der Pariser Akademie. Dort beschäftigte er sich auch mit Philosophie und Chemie. Seine Ideen zur Mechanik bilden die Grundlagen für die Arbeiten von HAMILTON und JACOBI. Weitere Beiträge betreffen die Theorie gewöhnlicher und partieller Differentialgleichungen, Differentialgeometrie und die Himmelsmechanik (Drei-Körper-Problem). Nach Auflösung der Akademie arbeitete LAGRANGE in einer Kommission zur Reform von Maßen und Gewichten und der Einführung des metrischen Systems.

Beweis von Satz 4.2. Sei $y \in U$. Wegen der Messbarkeit von $t \mapsto f_t(y)$ ist

$$|f(y) - f(x)| \leq \int_T |f_t(y) - f_t(x)| \,\mathrm{d}\mu(t) \overset{\text{(ii)}}{\leq} \underbrace{\|k\|_{L^1(\mu)}}_{=:K} \|y - x\|_X$$

für $x, y \in U$. Damit haben wir $|f(y)| \leq |f(y) - f(\bar{x})| + |f(\bar{x})| < \infty$. Dies zeigt (i).
Für $\zeta \in \partial f(x) \neq \emptyset$ (Proposition 1.5 (i)) und $v \in X$ gilt

$$\langle \zeta, v\rangle \leq f^0(x; v) = \limsup_{\substack{y \to x \\ \lambda \searrow 0}} \int_T \frac{\overbrace{f_t(y + \lambda v) - f_t(y)}^{\leq k(t)\|v\|_X}}{\lambda} \,\mathrm{d}\mu(t) \leq \int_T f_t^0(x; v) \,\mathrm{d}\mu(t). \tag{4.2a}$$

Der letzte Umformungsschritt beruht auf folgender stärkeren Version des Lemmas von FATOU[20]: Ist eine Folge integrierbarer Funktionen $(u_i)_{i \in \mathbb{N}}$ punktweise durch eine integrierbare Majorante beschränkt, so folgt $\int \liminf u_i \,\mathrm{d}\mu \leq \liminf \int u_i \,\mathrm{d}\mu$ und $\int \limsup u_i \,\mathrm{d}\mu \geq \limsup \int u_i \,\mathrm{d}\mu$, vgl. [BF89], p. 293, Aufgabe 8.
Wir setzen nun $\hat{f}_t(v) := f_t^0(x; v)$ und $\hat{f}(v) := \int_T \hat{f}_t(v) \,\mathrm{d}\mu(t)$ und bemerken, dass $\hat{f}_t(.)$ nach Proposition 1.3 (ii), (iii), (iv) stetig in $0 \in X$ und konvex sind, und damit auch $\hat{f}(.)$. Wegen $\langle \zeta, v - 0\rangle = \langle \zeta, v\rangle \leq \hat{f}(v) = \hat{f}(v) - \hat{f}(0)$ für alle $v \in X$ liegt ζ nach Definition 1.21 in Sub $\hat{f}(0) \overset{\text{Prop. 1.22}}{=} \partial \hat{f}(0)$. Proposition 4.1 liefert nun $\partial \hat{f}(0) = \int_T \partial \hat{f}_t(0) \,\mathrm{d}\mu(t)$, wobei $\Omega := B_1(0) \subset X$ in Proposition 4.1 gewählt wird. Proposition 4.1 ist tatsächlich anwendbar; denn Lemma 4.3 weiter unten sichert, dass Bedingung c) aus Satz 4.2 die Bedingung c*) aus Proposition 4.1 impliziert. Darüberhinaus garantiert Lemma 4.3 die Messbarkeit von $t \mapsto \hat{f}_t(v)$ für alle $v \in X$. Hieraus folgt mit $\left|\hat{f}_t(v)\right| \leq k(t) \|v\|_X$ schließlich $\big(t \mapsto \hat{f}_t(v)\big) \in L^1(\mu)$. Zum Beweis der Inklusion bleibt noch $\partial \hat{f}_t(0) \subset \partial f_t(x)$ herzuleiten. Hierzu betrachten wir Folgen $y_i \to 0$, $\lambda_i \searrow 0$, die $\lim_{i\to\infty} \frac{\hat{f}_t(y_i + \lambda_i v) - \hat{f}_t(y_i)}{\lambda_i} = \hat{f}_t^0(0; v)$ realisieren. Nun ist die linke Seite definitionsgemäß gleich

$$\lim_{i\to\infty} \frac{f_t^0(x; y_i + \lambda_i v) - f_t^0(x; y_i)}{\lambda_i} \overset{\text{Prop. 1.3 (iii)}}{\leq} \lim_{i\to\infty} \frac{f_t^0(x; \lambda_i v)}{\lambda_i} = \lim_{i\to\infty} f_t^0(x; v) = f_t^0(x; v),$$

was den Beweis von (ii) abschließt.
Nun ist noch die Gleichheitsdiskussion (iii) zu führen. Wir haben

$$\begin{aligned} f^0(x; v) &\overset{(4.2a)}{\leq} \int_T f_t^0(x; v) \,\mathrm{d}\mu(t) \overset{f_t \text{ reg.}}{=} \int_T f_t'(x, v) \,\mathrm{d}\mu(t) \\ &= f'(x; v) \leq f^0(x; v), \end{aligned}$$

wobei das Integral $\int_T f_t'(x, v) \,\mathrm{d}\mu(t)$ existiert, weil der Integrand durch $k(.) \|v\|_X$ dominiert wird. Daraus ergibt sich nun $f'(x; v) = f^0(x; v)$ und damit die Regularität von f. Sei jetzt $\zeta = \int_T \langle \zeta_t, \cdot\rangle \,\mathrm{d}\mu(t) \in \int_T \partial f_t(x) \,\mathrm{d}\mu(t)$, dann ist $f^0(x; v) = \int_T f_t^0(x; v) \,\mathrm{d}\mu(t) \geq \int_T \langle \zeta_t, v\rangle \,\mathrm{d}\mu(t) = \langle \zeta, v\rangle$, also $\zeta \in \partial f(x)$. Zusammen mit der zuvor bereits gezeigten Inklusion (4.2) beweist dies die Gleichheit der Mengen. □

[20] PIERRE JOSEPH LOUIS FATOU (1878 – 1929) Französischer Mathematiker, ab 1901 am Observatorium in Paris. Er beschäftigte sich neben der Astronomie mit Differentialgleichungssystemen, Numerik, Funkionalgleichungen und Funktionentheorie.

Lemma 4.3
Seien die Voraussetzungen in (i) und (ii) aus Satz 4.2 erfüllt.
DANN ist $t \mapsto \hat{f}_t(v) = f_t^0(x; v)$ für alle $v \in X$ messbar. Darüber hinaus impliziert die Bedingung c) aus Satz 4.2 die Bedingung c) aus Proposition 4.1.*

Beweis. Im *Fall a)* ist T abzählbar, und damit jedes Urbild einer messbaren Menge in $\mathbb{R}$ abzählbar, also messbar. Für den *Fall b)* nutzen wir die Existenz einer abzählbaren dichten Teilmenge $M \subset X$ und die Stetigkeit der Abbildung f_t für μ-fast alle $t \in T$; dann ist nämlich

$$\hat{f}_t(v) = f_t^0(x; v) = \limsup_{\substack{\lambda \searrow 0, \lambda \in \mathbb{Q} \\ y \to x, y \in M}} \underbrace{\frac{f_t(y + \lambda v) - f_t(y)}{\lambda}}_{\text{messbar}} \tag{4.3}$$

eine messbare Funktion.
Zum *Fall c)* zeigen wir zunächst, dass die Abbildung $t \mapsto f_t^0(x; v)$ für vorgegebene $x, v \in X$ oberhalbstetig ist: Für $t_i \to t$ soll also $f_t^0(x; v) \geq \limsup_{i \to \infty} f_{t_i}^0(x; v)$ gelten. In Proposition 1.5 (v) haben wir $f_t^0(y; v) = \max_{\zeta \in \partial f_t(y)} \langle \zeta, v \rangle$ hergeleitet. Wir wählen also eine Folge $(\zeta_i)_{i \in \mathbb{N}}$, $\zeta_i \in \partial f_{t_i}(x)$, mit $\langle \zeta_i, v \rangle \geq f_{t_i}^0(x, v) - \frac{1}{i}$. (Geht auch anders; das Maximum wird angenommen.)
Zu jedem $i \in \mathbb{N}$ garantiert die Oberhalbstetigkeit von $t \mapsto \partial f_t(y)$ die Existenz eines $\eta_i \in \partial f_t(x)$ mit $\langle \eta_i, v \rangle \geq \langle \zeta_i, v \rangle - \frac{1}{i}$, denn für alle $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für alle $\tau \in B_\delta(t)$

$$\begin{aligned} \partial f_\tau(x) &\subset \{\zeta \in X^* \mid \exists\, \eta \in \partial f_t(x) : |\langle \zeta - \eta, v \rangle| < \varepsilon\} \\ &= \left\{\zeta \in X^* \;\middle|\; \exists\, \eta \in \partial f_t(x) : \zeta - \eta \in v^{-1}(-\varepsilon, \varepsilon)\right\}. \end{aligned}$$

Damit erhalten wir $f_t^0(x; v) \geq \langle \eta_i, v \rangle \geq f_{t_i}^0(x; v) - \frac{2}{i}$, was nach Bilden des $\limsup$ die Oberhalbstetigkeit liefert.
Hieraus leiten wir nun die Messbarkeit her:

$$\begin{aligned} (f_t^0)^{-1}(a, b) &= (f_t^0)^{-1}(-\infty, b) \setminus (f_t^0)^{-1}(-\infty, a] \\ &= (f_t^0)^{-1}(-\infty, b) \setminus (f_t^0)^{-1}\left(\bigcap_{i=1}^{\infty}(-\infty, a + \tfrac{1}{i})\right) \\ &= \underbrace{(f_t^0)^{-1}(-\infty, b)}_{\text{offen } \Rightarrow \in \mathscr{T}} \setminus \underbrace{\bigcap_{i=1}^{\infty} \underbrace{(f_t^0)^{-1}(-\infty, a + \tfrac{1}{i})}_{\text{offen } \Rightarrow \in \mathscr{T}}}_{\in \mathscr{T}} \in \mathscr{T}, \end{aligned}$$

wobei die fixierten Argumente $x, v \in X$ in $f_t^0(x; v)$ unterdrückt haben. □

Bemerkung. Zum Fall c) ist anzumerken, dass wir *unter unseren Rahmenbedingungen* weder die Regularität , von μ noch die Separabilität von T explizit ausgenutzt haben; die Oberhalbstetigkeit von $t \mapsto f_t^0(x; v)$ impliziert bereits die Messbarkeit. Proposition 4.1 von IOFFE und LEVIN hingegen benötigt diese Voraussetzung, siehe insbesondere [IL74, p. 16]

Wir möchten in den nächsten beiden Sätzen den Spezialfall abgeschlossener Teilräume von L^p-Räumen betrachten. Man beachte, dass wir dabei Voraussetzungen und Notation ändern; insbesondere wird Y nicht die Rolle von U als offene Teilmenge von X spielen.

Satz 4.4 [L^∞-Situation]
Sei $(T, \mathscr{T}, \mu)$ ein σ-finiter positiver Maßraum, Y ein separabler $\mathbb{R}$-BANACHraum und X

ein abgeschlossener Teilraum von $L^\infty(\mu, Y)$. *Weiterhin sei zu jedem* $t \in T$ *eine Abbildung* $g_t : Y \to \mathbb{R}$ *mit den folgenden drei Eigenschaften gegeben.*

- $t \mapsto g_t(v)$ *messbar für alle* $v \in Y$,
- $g(x) := \int_T g_t(x(t))\,\mathrm{d}\mu(t) < \infty$ *für ein* festes $x \in X$,
- $\exists\ \varepsilon > 0, k \in L^1(\mu)$, so dass für μ-fast alle $t \in T$ $\forall$ $x_1, x_2 \in B_\varepsilon(x(t)) \subset Y$ gilt: $|g_t(x_1) - g_t(x_2)| \le k(t)\,\|x_1 - x_2\|_Y$.

DANN ist g LIPSCHITZ-*stetig nahe* x, *und es gilt*

$$\partial g(x) \subset \int_T \partial g_t(x(t))\,\mathrm{d}\mu(t). \tag{4.4}$$

Falls g_t *regulär bei* $x(t)$ *für alle* $t \in T$ *sind, so ist* g *regulär bei* x, *und wir haben Gleichheit in* (4.4).

Beweis. Der Nachweis der LIPSCHITZ-Stetigkeit verläuft wie im Beweis von Satz 4.2. Sei nun $\eta \in \partial g(x)$, $v \in X$. Wir setzen $\hat{g}_t(v) := g_t^0(x; v)$ und $\hat{g}(v) := \int_T \hat{g}_t(v)\,\mathrm{d}\mu(t)$.
Konvexität und Stetigkeit von $\hat{g}_t$ sind Konsequenz aus Proposition 1.3. Die Messbarkeit der Abbildung $t \mapsto \hat{g}_t(v)$ folgt wie in (4.3), und natürlich ist $\bar{v} \mapsto \hat{g}_t(\bar{v})$ für $\bar{v} \in Y$ stetig. Daher ist $t \mapsto \hat{g}_t(v(t))$ für $v \in X$ messbar. Analog zum Beweis von Satz 4.2 folgt $\eta \in \operatorname{Sub} \hat{g}(0) = \partial \hat{g}(0)$. Wegen $\partial \hat{g}_t(0) = \partial g_t(x(t))$ und $\partial \hat{g}(0) \subset \partial g(x)$ (vgl. Beweis von Satz 4.2 (ii)) liefert Proposition 4.1 b) die Inklusion (4.4).
Zur Gleichheitsdiskussion: Seien die g_t regulär bei x für alle $t \in T$. Dann haben wir für $v \in X$

$$\begin{aligned} g^0(x; v) &\ge \limsup_{\lambda \searrow 0} \frac{g(x + \lambda v) - g(x)}{\lambda} \ge \liminf_{\lambda \searrow 0} \frac{g(x + \lambda v) - g(x)}{\lambda} \\ &\overset{\substack{[\text{BF89}] \\ \text{p. } 293_8}}{\ge} \int_T g_t'(x(t); v(t))\,\mathrm{d}\mu(t) \overset{g_t \text{ reg.}}{=} \int_T g_t^0(x(t); v(t))\,\mathrm{d}\mu(t) \overset{(4.2a)}{\ge} g^0(x; v), \end{aligned}$$

wobei wir für die letzte Ungleichung wieder ein FATOU-Argument benutzt haben wie in (4.2a). Hieraus ergibt sich die Regularität von g in x. Ist dann ein $\eta \in X^*$ und eine Familie von Abbildungen $(\eta_t)_{t \in T}$ mit $\eta_t \in \partial g_t(x(t))$ und $\langle \eta, v \rangle_{X^* \times X} = \int_T \langle \eta_t, v(t) \rangle_{Y^* \times Y}\,\mathrm{d}\mu(t)$ gegeben, so folgt

$$\begin{aligned} \langle \eta, v \rangle = \int_T \langle \eta_t, v(t) \rangle\,\mathrm{d}\mu(t) &\le \int_T g_t^0(x(t); v(t))\,\mathrm{d}\mu(t) \\ &= \int_T g_t'(x(t); v(t))\,\mathrm{d}\mu(t) = g'(x; v) = g^0(x; v), \end{aligned}$$

also $\eta \in \partial g(x)$. □

Beispiel 4.5 Man betrachte den Fall $(T, \mathscr{T}, \mu) = ([a, b], \mathscr{B}, L^1)$, $-\infty < a < b < +\infty$, wobei $\mathscr{B}$ die Klasse der BOREL[21]-Mengen auf $[a, b]$ bezeichne, $Y := \mathbb{R}^n \times \mathbb{R}^n$ und

$$X := \left\{ (z, p) \in L^\infty([a, b], \mathbb{R}^n \times \mathbb{R}^n) \,\middle|\, \exists\, c \in \mathbb{R}^n : z(s) = c + \int_a^s p(\bar{s})\,\mathrm{d}\bar{s} \text{ f.f.a. } s \in [a.b] \right\}.$$

[21] EMILE BOREL (1871 – 1956) Französischer Mathematiker, 1897 Sekretär der *Société Mathématique de France*, 1905 deren Präsident, Professor an der Sorbonne für Funktionentheorie (1909) und Wahrscheinlichkeitstheorie und mathematische Physik (1920), 1921 Mitglied der *Académie des Sciences*, 1925 Marine-

In Aufgabe 4.1 soll bewiesen werden, dass X ein abgeschlossener Unterraum von $L^\infty([a,b],\mathbb{R}^n\times\mathbb{R}^n)$ ist. Für eine LAGRANGE-Funktion $F:\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ betrachten wir das Funktional

$$\begin{array}{rccl} \mathscr{F}: & X & \to & \mathbb{R} \\ & u & \mapsto & \int_a^b F(t,u(t),u'(t))\,\mathrm{d}t \end{array} \tag{4.5}$$

und nehmen an, dass $\mathscr{F}(X)\neq\{+\infty\}$. Wir fordern, dass F nachstehende Eigenschaften erfülle.

- $t\mapsto F(t,z,p)$ sei für alle $(z,p)\in Y$ messbar.
- Für jedes $(\hat z,\hat p)\in X$ gebe es $\varepsilon>0$, $k\in L^1([a,b])$, so dass für L^1-fast alle $t\in[a,b]$

$$|F(t,z_1,p_1)-F(t,z_2,p_2)|\le k(t)\,(|z_1-z_2|+|p_1-p_2|)$$

  für alle $(z_1,p_1),(z_2,p_2)\in B_\varepsilon(\hat{z(t)},\hat{p(t)})$ gelte.

Mit den Bezeichnungen $g_t(z,p):=F(t,z,p)$ und $g(z,p):=\int_a^b g_t(z(t),p(t))\,\mathrm{d}t$ liefert Satz 4.4 für jedes $\eta\in\partial g(\hat z,\hat p)$ und $t\in[a,b]$ ein $(Q(t),P(t))\in\partial g_t(\hat z(t),\hat p(t))$ mit

$$\langle\eta,(z,p)\rangle_{X^*\times X}=\int_a^b\big(Q(t)\cdot z(t)+P(t)\cdot p(t)\big)\,\mathrm{d}t\qquad\forall\,(z,p)\in X.$$

Ist $(\hat z,\hat p)$ zudem ein lokales Maximum von $\mathscr{F}=g$, so ist $0\in\partial g(\hat z,\hat p)$ nach Proposition 1.26, also $\int_a^b(Q(t)\cdot z(t)+P(t)\cdot p(t))\,\mathrm{d}t=0$ für alle $(z,p)\in X$. Hieraus folgt nach Aufgabe 4.2

$$P\in H^{1,1}\quad\text{und}\quad\dot P=Q\text{in Integralform.}$$

Ist nun $F\in C^1$, dann gilt $\partial g_t(\hat z(t),\hat p(t))=\begin{pmatrix}F_z(t,\hat z(t),\hat p(t))\\ F_p(t,\hat z(t),\hat p(t))\end{pmatrix}$.
Hieraus ergeben sich die klassischen EULER[22]-LAGRANGE-Gleichungen

$$\tfrac{\mathrm{d}}{\mathrm{d}t}F_p(t,\hat z(t),\hat p(t))=F_z(t,\hat z(t),\hat p(t)),$$

vgl. auch [VdM03], Kapitel 1 und 3. ◇

minister, 1928 Gründungsdirektor des *Institut Poincaré*, 1934 Präsident der *Académie des Sciences*, 1941 von Besatzungstruppen inhaftiert. Aus der Vielzahl der Arbeiten sind insbesondere die zu quasianalytischen Funktionen, Maßtheorie, Funktionentheorie, und Wahrscheinlichkeitstheorie zu nennen. Auf letzterem Gebiet nahm er bereits Gedankengänge KOLMOGOROWS vorweg und verschärfte J. BERNOULLIS *Starkes Gesetz der großen Zahlen*. Weiterhin beschäftigte er sich mit Spieltheorie. BOREL gilt als einer der bedeutendsten französischen Mathematiker des 20. Jahrhunderts.

[22] LEONHARD EULER (1707 – 1783) Schweizer Mathematiker; Professor in St. Petersburg und Berlin. EULER ist einer der bedeutendsten Mathematiker aller Zeiten; im vorliegenden Rahmen kann seine Biographie nur kurz beleuchtet werden. Er erzielte bahnbrechende Resultate auf allen mathematischen Gebieten und prägte die Entwicklung der Mathematik über das 18. Jahrhundert hinaus. — EULER wurde schon als Dreizehnjähriger an der Philosophischen Fakultät der Universität Basel immatrikuliert; unter dem Einfluss von JOHANN I BERNOULLI wandte er sich der Mathematik zu und folgte 1727 dessen Söhnen an die neu gegründete Akademie nach St. Petersburg. Dort wurde er Professor für Physik (1731) und Mathematik (1733). Wegen politischer Unruhen nach dem Tode der Zarin ANNA I folgte er 1741 einem Ruf an die Berliner Akademie und wurde 1746 Leiter deren mathematischer Abteilung. Nach Differenzen mit dem preußischen König kehrte er 1766 nach St. Petersburg zurück. Trotz seiner Erblindung kurze Zeit darauf setzte EULER seine Forschungen bis zu seinem Tod unvermindert fort. — EULER ist Autor zahlreicher Bücher über die verschiedensten Bereiche der Mathematik und führte viele noch heute gebräuchliche Bezeichnungen ein, wie etwa die Funktionsschreibweise $f(x)$, das Summenzeichen $\sum$, i für die imaginäre Einheit und e für die Basis der Exponentialfunktion. Ein besonderer Schwerpunkt in seinem Schaffen stellt die Analysis dar, in

Satz 4.6 [L^p-Situation]
*Sei $(T, \mathscr{T}, \mu)$ ein positiver vollständiger Maßraum mit $\mu(T) < \infty$, Y ein separabler $\mathbb{R}$-*BANACH*raum und X ein abgeschlossener Teilraum von $L^p(\mu, Y)$, $p \in [1, \infty)$, $\frac{1}{p} + \frac{1}{q} = 1$. Weiterhin sei zu jedem $t \in T$ eine Abbildung $h_t : Y \to \mathbb{R}$ mit den folgenden drei Eigenschaften gegeben.*

- *$t \mapsto h_t(v)$ messbar für alle $v \in Y$,*
- *$h(x) := \int_T h_t(x)\,\mathrm{d}\mu(t) < \infty$ für ein* festes *$x \in X$,*
- *es gilt* eine *der beiden nachstehenden Optionen.*

  (i) $\exists\ k \in L^q(\mu)$, so dass für μ-fast alle $t \in T, x_1, x_2 \in Y : |h_t(x_1) - h_t(x_2)| \leq k(t)\,\|x_1 - x_2\|_Y$

  (ii) Jede Funktion h_t ist LIPSCHITZ*-stetig nahe jeden Punktes von Y, und es gibt eine Konstante $C > 0$ mit*

$$\Big(\zeta \in \partial h_t(y) \Longrightarrow \|\zeta\|_{Y^*} \leq C(1 + \|y\|_Y^{p-1})\Big)\ \forall\, t \in T, y \in Y.$$

DANN ist h LIPSCHITZ*-stetig auf allen beschränkten Teilmengen von X, und es gilt für jedes $x \in X$*

$$\partial h(x) \subset \int_T \partial h_t(x(t))\,\mathrm{d}\mu(t). \tag{4.6}$$

Falls für alle $t \in T$ die Funktionen h_t regulär bei $x(t)$ sind, so ist h regulär bei x, und wir haben Gleichheit in (4.6).

Bemerkung. Für $p \in [1, \infty)$, $\frac{1}{p} + \frac{1}{q} = 1$, ist $\big(L^p(\mu, Y)\big)^* = L^q(\mu, Y^*)$ nach [Alt02, p. 159].

Beweisskizze zu Satz 4.6. Im Fall (i) zeigt man wie im Beweis von Satz 4.2 die globale LIPSCHITZ-Stetigkeit von h mit Hilfe der HÖLDER[23]-Ungleichung. Es gelte nun (ii). Für ein vorgegebenes $m > 0$ wähle $u, v \in L^p(\mu, Y)$ mit $\|u\|_{L^p(\mu,Y)}, \|v\|_{L^p(\mu,Y)} \leq m$. Nach dem Mittelwertsatz 1.33 gibt es zu μ-fast jedem $t \in T$ ein $\theta_t \in \partial h_t(w_t)$ mit $w_t \in [u(t), v(t)] \subset Y$ und

$$h_t(u(t)) - h_t(v(t)) = \langle \theta_t, u(t) - v(t) \rangle_{Y^* \times Y}.$$

deren Mittelpunkt er statt der Geometrie den Funktionenbegriff stellte. Sein Hauptwerkzeug bilden hierbei die Potenzreihen; weiter ermittelte er grundlegende Zusammenhänge zwischen Funktionen, wie etwa $e^{ix} = \cos x + i \sin x$, und legte damit die Grundlagen der Funktionentheorie. Auch der Ansatz $y(x) = e^{ax}$ als Lösung der homogenen linearen Differentialgleichung und die Polygonzugmethode zur approximativen Lösung von Anfangswertproblemen gehen auf EULER zurück. Partielle Differentialgleichungen waren ebenfalls ein Forschungsgegenstand EULERS, z. B. im Hinblick auf die *Schwingende Saite*. Nicht zuletzt verhalf er der Variationsrechnung durch Betrachten allgemeiner Funktionale vom Typ $\int F(x, u(x), u'(x))\,\mathrm{d}x$ zum Status eines eigenen Gebiets und leitete die EULER-LAGRANGEschen Differentialgleichungen $F_z = \frac{\mathrm{d}}{\mathrm{d}x} F_p$ als notwendige Bedingung für einen Extremierer her. Ideen zu divergenten Reihen bildeten die Grundlage zu Forschungen von BOREL u. a. Anfang des 20. Jahrhunderts. Ähnliches gilt für seine Polyederformel. In der Algebra bewies EULER, dass alle Nullstellen eines Polynoms von der Form $a + ib$ sind. Auf dem Gebiet der Zahlentheorie konnte er die von FERMAT aufgestellten Vermutungen teilweise beweisen oder widerlegen. Ferner führte er die *Zeta-Funktion* $\zeta(s) = \sum_1^\infty n^{-s} = \prod \frac{1}{1-p^{-s}}$ ein. Weiterhin beschäftigte er sich mit Physik und Astronomie (*Drei-Körper-Problem*).

[23] LUDWIG OTTO HÖLDER (1859 – 1937) Deutscher Mathematiker, Professor in Göttingen, Tübingen, Königsberg, Leipzig. Beiträge zu fast allen Gebieten der Mathematik, vor allem Algebra, Funktionentheorie, Mechanik, Grundlagenforschung.

Demnach haben wir

$$\begin{aligned} |h_t(u(t)) - h_t(v(t))| &\leq C(1 + \|w_t\|_Y^{p-1})\, \|u(t) - v(t)\|_Y \\ &\leq \underbrace{C(p)\Big(1 + \|u(t)\|_Y{}^{p-1} + \|v(t)\|_Y{}^{p-1}\Big)}_{=:\tau(t),\tau\in L^q(\mu)} \|u(t) - v(t)\|_Y\,. \end{aligned}$$

Integration ergibt

$$\begin{aligned} |h(u) - h(v)| &\leq \int_T |h_t(u(t)) - h_t(v(t))| \,\mathrm{d}\mu(t) \leq \int_T \tau(t)\, \|u(t) - v(t)\| \,\mathrm{d}\mu(t) \\ &\overset{\text{Hölder}}{\leq} \|\tau\|_{L^q(\mu)}\, \|u - v\|_{L^p(\mu,Y)}\,, \end{aligned}$$

wobei $\|\tau\|_{L^q(\mu)} \leq C(p,m)$; also ist h LIPSCHITZ-stetig auf $\overline{B_m(0)} \subset X$. Bei diesem Schritt wurde $\mu(T) < \infty$ für die HÖLDER-Ungleichung benutzt.
Im Fall (i) erhalten wir parallel zum Beweis von Satz 4.4 über das FATOU-Argument für $\theta \in \partial h(x)$

$$\langle \theta, v\rangle \leq h^0(x;v) \leq \int_T h_t^0(x(t);v(t))\,\mathrm{d}\mu(t), \tag{4.7}$$

in (ii) haben wir wegen

$$\left|\frac{h_t(y(t) + \lambda v(t)) - \ h_t(y(t))}{\lambda}\right| \leq \tau(t)\,\|v(t)\|_Y \overset{\text{Hölder}}{\in} L^1(\mu)$$

eine integrierbare Majorante, und wir können damit auch hier das FATOU-Argument anwenden, um (4.7) zu erhalten.
Mit den Bezeichnungen $\hat{h}_t(v) := h_t^0(x;v)$ und $\hat{h}(v) := \int_T \hat{h}_t(v(t))\,\mathrm{d}\mu(t)$ erhalten wir konvexe Funktionen $\hat{h}_t$ und $\hat{h}$ und $\theta \in \partial h(x)$ liegt wegen (4.7) in Sub $\hat{h}(0)$. Dies bleibt auch nach Einschränkung von v auf $X \cap L^\infty(\mu,Y) \subset L^p(\mu,Y)$ richtig. Außerdem ist $X \cap L^\infty(\mu,Y)$ ein abgeschlossener Unterraum von $L^\infty(\mu,Y)$, folglich dürfen wir Satz 4.4 anwenden; dessen weitere Voraussetzungen werden in Lemma 4.7 überprüft.
Die Gleichheitsdiskussion läuft wie im Beweis zu Satz 4.4. □

Lemma 4.7
Es gelten die Voraussetzungen von Satz 4.6.
Für $g_t := \hat{h}_t$ sind die Voraussetzungen von Satz 4.4 erfüllt.

Beweis.

- Die Messbarkeit von $t \mapsto g_t(v)$ zeigt man wie im Fall b) von Satz 4.2.
- Offenbar ist $g(0) = \int_T g_t(0)\,\mathrm{d}\mu(t) = \int_T h^0(x;0) = 0$.
- LIPSCHITZ-Voraussetzung:

  (i) $\partial h_t(x(t))$ ist wegen $\mu(T) < \infty$ durch $k \in L^q(\mu)$ beschränkt.

(ii) Wir haben für alle $\eta_t \in \partial h_t(x(t))$

$$\|\eta_t\|_{Y^*} \le \underbrace{C(1 + \|x(t)\|_Y^{p-1})}_{\in L^q(\mu)},$$

also

$$\begin{aligned} g_t(v_1) - g_t(v_2) &= \max_{\eta' \in \partial h_t(x(t))} \langle \eta', v_1 \rangle - \max_{\eta'' \in \partial h_t(x(t))} \langle \eta'', v_2 \rangle \\ &= \langle \eta_{\max}, v_1 \rangle - \max_{\eta'' \in \partial h_t(x(t))} \langle \eta'', v_2 \rangle \\ &\le \langle \eta_{\max}, v_1 - v_2 \rangle \le \|\eta_{\max}\|_{Y^*} \|v_1 - v_2\|_Y \\ &\le \alpha(t) \|v_1 - v_2\|_Y \quad \text{für ein } \alpha \in L^1(\mu). \end{aligned}$$

Vertauschen von v_1 und v_2 liefert die gesuchte Abschätzung.

□

Maximalfunktionen

Generalvoraussetzung für diesen Abschnitt

Im Folgenden sei T ein topologischer Raum, X ein $\mathbb{R}$-Banachraum, $x \in X$.
Zu jedem $t \in T$ sei eine nahe x Lipschitz-stetige Abbildung $f_t : X \to \mathbb{R}$ gegeben.
Setze $f : X \to \mathbb{R} \cup \{\infty\}$, $x \mapsto \sup_{t \in T} f_t(x)$.

Ferner sollen folgende Bedingungen gelten:

(i) T folgenkompakt,

(ii) $t \mapsto f_t(y)$ ist für alle y aus einer geeigneten Umgebung U von x oberhalbstetig,

(iii) für alle $t \in T$ und alle $y, z \in U$ gelte $|f_t(y) - f_t(z)| \le K \|y - z\|_X$,
und $(f_t(y))_{t \in T}$ sei für alle $y \in U$ beschränkt,

(iv) X separabel ODER (iv)' T metrisch.

Definition 4.8

- $$\partial_{[T]} f_t(x) := \overline{\operatorname{conv}} \left\{ \zeta \in X^* \;\middle|\; \begin{array}{l} \exists\, \zeta_i \in \partial f_{t_i}(x_i), x_i \to x, t_i \to t \\ \zeta \text{ Häufungspunkt von } (\zeta_i)_{i \in \mathbb{N}} \\ \text{bezüglich der schwach}*\text{-Konvergenz} \end{array} \right\}$$
- $(\tau, y) \mapsto \partial_{[T]} f_\tau$ heißt (schwach∗-)*abgeschlossen bei* (t, x), wenn $\partial_{[T]} f_t(x) = \partial f_t(x)$.
- Für $S \subset T$ setzen wir
$$P[S] := \left\{ \mu \text{ Radon}^{24}\text{-Maß auf } T \text{ mit } \mu(T) = 1,\ \operatorname{supp} \mu \subset S \right\}.$$

[24] Johann Karl August Radon (1887 – 1956) Österreichischer Mathematiker, Professor in Hamburg, Greifswald, Erlangen, Breslau und Wien. Arbeiten auf dem Gebiet der Variationsrechnung, der Differentialgeometrie und der absolut additiven Mengenfunktionen.

Bemerkungen.

- „$\supset$“ in der Definition „*abgeschlossen bei*“ gilt natürlich immer.
- Ist t isoliert in T, so ist jede gegen t konvergierende Folge bis auf endlich viele Folgenglieder konstant t. Daher folgt $\partial_{[T]} f_t(x) = \partial f_t(x)$ in dem Fall mit Proposition 1.10 (ii).
- f ist endlich auf U, und das Supremum wird angenommen. Wir wählen hierzu ein festes $y \in U$ und erhalten $|f_t(y)| \le |f_t(y) - f_t(x)| + |f_t(x)| \overset{\text{(iii)}}{\le} K \|y - x\|_X + |f_t(x)| < \infty$, wobei wegen (iii) eine Schranke unabhängig von $t \in T$ existiert. Folglich ist für eine Maximalfolge $t_i \to t$ in T mit $f_{t_i}(y) \to \sup_{t \in T} f_t(y)$ die Folge $(f_{t_i}(y))_{i \in \mathbb{N}}$ notwendig beschränkt, und (i) liefert die Existenz einer gegen ein $t_* \in T$ konvergierenden Teilfolge $(t_i)_{i \in \mathbb{N}}$. Damit haben wir $\sup_{t \in T} f_t(y) = \lim_{i \to \infty} f_{t_i}(y) \overset{\text{(ii)}}{\le} f_{t_*}(y) \le \sup_{t \in T} f_t(y)$. Also
$$f(y) = \max_{t \in T} f_t(y) \ < \infty \quad \text{auf } U.$$
- Es gilt $|f(x) - f(y)| \le K \|x - y\|_X$, siehe Aufgabe 1.15.
- Die *Inzidenzmenge* $M(y) := \{t \in T \mid f_t(y) = f(y)\}$ ist nicht-leer ($t_*(y) \in M(y)$) und folgenabgeschlossen: Betrachte eine Folge $t_i \to t$ in T mit $t_i \in M(y)$ für fast alle $i \in \mathbb{N}$. Dann ist $f(y) \ge f_t(y) \overset{\text{(iii)}}{\ge} \limsup_{i \to \infty} f_{t_i}(y) = f(y)$, also $f_t(y) = f(y)$ und damit $t \in M(y)$.

Satz 4.9
Wir haben
$$\partial f(x) \subset \left\{ \int_T \partial_{[T]} f_t(x) \, \mathrm{d}\mu(t) \,\middle|\, \mu \in P[M(x)] \right\}. \tag{4.8}$$
Falls für alle $t \in M(x)$ die Funktionen f_t regulär bei x sind und $(\tau, y) \mapsto \partial_{[T]} f_\tau(y)$ abgeschlossen bei (t, x) für alle $t \in M(x)$, so ist f regulär, und es gilt
$$\partial f(x) = \left\{ \int_T \partial f_t(x) \, \mathrm{d}\mu(t) \,\middle|\, \mu \in P[M(x)] \right\}. \tag{4.9}$$

Corollar 4.10 [Konvexer Fall]
U und alle f_t seien konvex, die Abbildungen $t \mapsto f_t(y)$, $y \in U$, stetig.
DANN ist f konvex auf U, und es gilt (4.9).

Beweis. Nach Aufgabe 4.4 ist f konvex. Aufgrund von Proposition 1.32 (ii) sind die f_t regulär. Satz 4.9 garantiert uns (4.9), sofern wir die Abgeschlossenheit von $(\tau, y) \mapsto \partial_{[T]} f_\tau(y)$ bei (t, x) für alle $t \in M(x)$ zeigen.

Sei also $(\zeta_i)_{i \in \mathbb{N}}$ für $x_i \to x$, $t_i \to t \in M(x)$ eine Folge mit $\zeta_i \in \partial f_{t_i}(x_i) \overset{\text{Prop. 1.22}}{=} \operatorname{Sub} f_{t_i}(x_i)$ und Häufungspunkt ζ. Um $\zeta \in \partial f_t(x)$ herzuleiten, wählen wir $v \in X$ so klein, dass $x_i, x_i + v \in U$ für $i \gg 1$, also
$$f_{t_i}(x + v) - f_{t_i}(x) \overset{\text{(iii)}}{\ge} \underbrace{f_{t_i}(x_i + v) - f_{t_i}(x_i)}_{\overset{\text{Def. 1.21}}{\ge} \langle \zeta_i, v \rangle} - 2K \underbrace{\|x_i - x\|_X}_{\to 0}.$$

Wir haben die Konvergenz $\langle \zeta_{i_k}, v\rangle \to \langle \zeta, v\rangle$ einer Teilfolge $\zeta_{i_k} \stackrel{*}{\rightharpoonup} \zeta$; aufgrund der Stetigkeit der Abbildungen $t \mapsto f_t(x)$ gilt $f_t(x+v) - f_t(x) \geq \langle \zeta, v\rangle$, also $\zeta \in \mathrm{Sub}\, f_t(x) = \partial f_t(x)$. $\square$

Corollar 4.11 [Differenzierbarer Fall]
Seien alle f_t strikt differenzierbar auf U, d.h. $\partial f_t(y) = \{D_s f_t(y)\}$ für alle $t \in T$, $y \in U$. Ist ferner $(t,y) \mapsto D_s f_t(y)$ stetig auf $T \times U$, DANN ist f regulär bei x, und es gilt

$$\partial f(x) = \left\{ \int_T D_s f_t(x)\, \mathrm{d}\mu(t) \,\middle|\, \mu \in P[M(x)] \right\}.$$

Der Beweis wird als Aufgabe 4.5 gestellt.

Beispiel 4.12 [(Banach) Supremum-Norm auf $C^0([a,b],\mathbb{R}^n)$] Für $T := [a,b]$ versehen wir den Raum $X := C^0(T,\mathbb{R}^n)$ mit der üblichen Norm $\|y\|_X := \max_{t\in T} |y(t)|$. Die Abbildungen $f_t(y) := |y(t)|$ sind konvex in y und stetig in t.
Sei ein $x \in X$ vorgegeben, dessen Betrag $|x(\cdot)|$ an genau einer Stelle $t \in [a,b]$ sein Maximum annimmt. Dann ist mit $M(x)$ auch $P[M(x)]$ einelementig, und aus Corollar 4.10 folgt mittels Proposition 1.15 (ii) die strikte Differenzierbarkeit von $f(y) = \|y\|_X$ in $x \in X$. $\diamond$

Beweis von Satz 4.9. Wir betrachten den Fall (iv) und stellen die Behandlung von (iv)' als Aufgabe 4.6. Den Beweis führen wir in vier Schritten.
I. Wir betrachten die durch $g : U \times X \to \mathbb{R}$, $(x;v) \mapsto \max\left\{\langle \zeta, v\rangle \,\middle|\, \zeta \in \partial_{[T]} f_t(x), t \in M(x)\right\}$ gegebene Funktion. Sie ist offenbar konvex in v. Das Maximum wird angenommen, da $M(x)$ abgeschlossen und $\partial_{[T]} f_t(x)$ nach Satz A.12 schwach$*$-kompakt: $\partial_{[T]} f_t(x)$ ist schwach$*$-abgeschlossen (und konvex nach Definition 4.8); zusätzlich ist $\partial_{[T]} f_t(x) \subset B_K(0) \subset X^*$. Wir wollen nun

$$f^0(x;v) \leq g(x;v) \qquad \text{für alle } v \in X \tag{4.10}$$

zeigen. Hierzu wählen wir Folgen $y_i \to x$, $\lambda_i \searrow 0$, die $f^0(x;v) = \lim_{i\to 0} \frac{f(y_i+\lambda_i v) - f(y_i)}{\lambda_i}$ realisieren. Wir finden $t_i \in M(y_i + \lambda_i v)$, so dass $\frac{f(y_i+\lambda_i v)-f(y_i)}{\lambda_i} \leq \frac{f_{t_i}(y_i+\lambda_i v)-f_{t_i}(y_i)}{\lambda_i}$. Auf letzteren Term können wir den Mittelwertsatz 1.33 anwenden und erhalten $\zeta_i \in \partial f_{t_i}(y_i^*)$ mit $y_i^* \in [y_i, y_i+\lambda_i v]$ und $\frac{f_{t_i}(y_i+\lambda_i v)-f_{t_i}(y_i)}{\lambda_i} = \langle \zeta_i, v\rangle$. Da T folgenkompakt ist und $\|\zeta_i\|_{X^*} \leq K$, können wir (ohne Umbenennen) zu Teilfolgen $t_i \to t \in T$ und $\zeta_i \stackrel{*}{\rightharpoonup} \zeta$ übergehen. Hieraus folgt

$$f^0(x;v) \leq \langle \zeta, v\rangle .$$

Wegen $y_i^* \to x$ liefert Definition 4.8 dann $\zeta \in \partial_{[T]} f_t(x)$.
Wir zeigen nun $t \in M(x)$, denn hieraus folgt $\langle \zeta, v\rangle \leq g(x;v)$ und damit die Behauptung (4.10). Für alle $\tau \in T$ haben wir $f_{t_i}(y_i+\lambda_i v) \geq f_\tau(y_i + \lambda_i v)$ wegen $t_i \in M(y_i+\lambda_i v)$. Damit folgt

$$\begin{aligned} f_{t_i}(x) &= f_{t_i}(x) - f_{t_i}(y_i+\lambda_i v) + f_{t_i}(y_i+\lambda_i v) \\ &\stackrel{\text{(iii)}}{\geq} f_{t_i}(y_i+\lambda_i v) - K\,\|y_i+\lambda_i v - x\|_X \\ &\geq f_\tau(y_i+\lambda_i v) - K\,\|y_i+\lambda_i v - x\|_X \\ &\geq f_\tau(x) - 2K\,\underbrace{\|y_i+\lambda_i v - x\|_X}_{\to 0}. \end{aligned}$$

Der Grenzübergang liefert (mit (ii)) $f_t(x) \geq f_\tau(x)$ für alle $\tau \in T$. Also $f_t(x) = \max_{\tau\in T} f_\tau(x) = f(x)$, d. h. $t \in M(x)$.

II. $g(x;0) = 0$, für alle $\zeta \in \partial f(x)$ ist mithin $\langle \zeta, v - 0\rangle = \langle \zeta, v\rangle \overset{\text{I.}}{\leq} g(x;v) - g(x;0)$, also $\zeta \in \operatorname{Sub}\big(g(x;\cdot)\big)(0)$. [IL74], p. 33f., liefert uns ein $\mu \in P[M(x)]$ und eine Familie $(\zeta_t)_{t\in T}$, $\zeta_t \in \partial_{[T]} f_t(x)$, mit $\langle \zeta, v\rangle = \int_T \langle \zeta_t, v\rangle \,\mathrm{d}\mu(t)$. Dies beweist (4.8). (Alternativ verwende man das bereits zitierte Minimax-Prinzip aus [Aub79].)
III. Es mögen nun die zusätzlichen Annahmen für (4.9) gelten. Wir setzen $\theta := \liminf_{\lambda\searrow 0} \frac{f(x+\lambda v)-f(x)}{\lambda} \leq f^0(x;v)$. Wir werden nun $\theta \geq f^0(x;v)$ zeigen. Dazu wählen wir ein $t \in M(x)$ und erhalten $\frac{f(x+\lambda v)-f(x)}{\lambda} \geq \frac{f_t(x+\lambda v)-f_t(x)}{\lambda}$, also

$$\theta \geq f_t'(x,v) \overset{f_t \text{ reg.}}{=} f_t^0(x;v) \qquad \text{für alle } t \in M(x).$$

Hieraus ergibt sich $\theta \geq \max_{t\in M(x)} f_t^0(x;v) \overset{!}{=} g(x;v) \overset{\textbf{I.}}{\geq} f^0(x;v)$, wobei wir die markierte Identität über

$$g(x;v) = \max\{\langle \zeta, v\rangle \mid \zeta \in \partial f_t(x), t \in M(x)\} \overset{\text{Prop. 1.5}}{=} \max\big\{ f_t^0(x;v) \,\big|\, t \in M(x)\big\}$$

erhalten. Dies zeigt die Regularität von f.
IV. Es bleibt noch die Gleichheitsdiskussion zu führen.
Wir haben für $\zeta \in \big\{ \int_T \partial f_t(x)\,\mathrm{d}\mu(t) \,\big|\, \mu \in P[M(x)]\big\}$

$$\begin{aligned}
\langle \zeta, v\rangle = \int_T \langle \zeta_t, v\rangle \,\mathrm{d}\mu(t) \quad &\leq \quad \int_T f_t^0(x;v)\,\mathrm{d}\mu(t) \\
&\overset{f_t \text{ reg.}}{=} \quad \int_T f_t'(x;v)\,\mathrm{d}\mu(t) \\
&= \quad \int_T \lim_{\lambda\searrow 0} \frac{f_t(x+\lambda v) - f_t(x)}{\lambda}\,\mathrm{d}\mu(t) \\
&\overset{\substack{\text{(iii)}\\ \text{dom. Konv.}}}{=} \quad \limsup_{\lambda\searrow 0} \int_T \frac{f_t(x+\lambda v) - f_t(x)}{\lambda}\,\mathrm{d}\mu(t) \\
&\overset{\mu\in P[M(x)]}{\leq} \quad \limsup_{\lambda\searrow 0} \int_T \frac{f(x+\lambda v) - f(x)}{\lambda}\,\mathrm{d}\mu(t) \\
&= \quad \limsup_{\lambda\searrow 0} \frac{f(x+\lambda v) - f(x)}{\lambda} \underbrace{\mu(T)}_{=1} \quad \leq \quad f^0(x;v),
\end{aligned}$$

also $\zeta \in \partial f(x)$. □

Aufgaben

4.1 Seien $-\infty < a < b < +\infty$, $F : \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n \to \mathbb{R}$, $k \in L^1([a,b])$, so dass

$$F(.,z,p) \text{ messbar für alle } z,p \in \mathbb{R}^n$$
$$|F(t,z_1,p_1) - F(t,z_2,p_2)| \leq k(t)|(z_1 - z_2, p_1 - p_2)| \text{ für fast alle} t \in [a,b].$$

Betrachten Sie das Variationsintegral

$$\mathscr{F}(u) := \int_a^b F(t,u(t),u'(t))\,dt$$

für $u \in H^{1,\infty}([a,b],\mathbb{R}^n)$, und zeigen Sie, dass die Voraussetzungen von Satz 4.4 der Vorlesung erfüllt sind, sofern das Integral existiert.
Bemerkung: Vergleichen Sie mit Beispiel 4.5.

4.2 F erfülle die Voraussetzungen aus Aufgabe 4.1. Zeigen Sie: Falls für $(Q(t), P(t)) \in \partial F(t, \hat{z}(t), \hat{p}(t))$, wobei $\hat{z}, \hat{p} \in L^\infty([a,b], \mathbb{R}^n)$ fixiert sind, die Identität

$$0 = \int_a^b \Big[Q(t)v(t) + P(t)v'(t)\Big]\, dt \text{ für alle } v \in H^{1,\infty}([a,b], \mathbb{R}^n), \tag{4.11}$$

gilt, dann ist $P \in H^{1,1}([a,b], \mathbb{R}^n)$ und erfüllt $Q = \dot{P}$. *Hinweis: Dies wurde im Beispiel 4.5 benutzt.*

4.3 Sei $(T, \mathscr{T}, \mu)$ ein σ-endlicher, positiver Maßraum, Y ein separabler BANACHraum, X ein abgeschlossener Teilraum von $L^\infty(T, Y)$, $f_t : Y \to \mathbb{R}$ für $t \in T$, und

$$f(z) := \int_T f_t(z(t)) d\mu(t) \text{ für } z \in X$$

erfülle $f(x) < \infty$ für ein $x \in X$. Zusätzlich existiere $\epsilon > 0, k \in L^1(T, \mathbb{R})$, so dass

$$|f_t(v_1) - f_t(v_2)| \leq k(t)\|v_1 - v_2\|_Y$$

für μ-fast alle $t \in T$ und für alle $v_1, v_2 \in B_\epsilon(x(t))$. Zusätzlich seien die Abbildungen $t \mapsto f_t(v)$ messbar für alle $v \in Y$. Zeigen Sie: Falls f_t regulär bei $x(t)$ für alle $t \in T$, dann ist f regulär bei x und erfüllt

$$\partial f(x) = \int_T \partial f_t(x(t))\, d\mu(t).$$

Hinweis: Sie dürfen die in der Vorlesung gezeigte Inklusion $\partial f(x) \subset \int_T \partial f_t(x(t))\, d\mu(t)$ *benutzen.*

4.4 Sei T ein kompakter Raum, X ein BANACHraum, $f : T \times X \to \mathbb{R}$ eine Funktion mit $t \mapsto f(t,x)$ oberhalbstetig und $x \mapsto f(t,x)$ konvex. Zeigen Sie: Die Funktion $\tilde{f}(x) := \sup_T f(t,x)$ ist konvex.
Bemerkung: Diese Tatsache wurde z. B. im Beweis von Corollar 4.10 benutzt.

4.5 Sei X separabler BANACHraum, T ein folgenkompakter topologischer Raum, und für ein $x \in X$ seien die Funktionen $f_t : X \to \mathbb{R}$ LIPSCHITZ-stetig nahe x für alle $t \in T$ und die Abbildungen $t \mapsto f_t(y)$ seien oberhalbstetig für alle $y \in X$. Zudem gebe es eine Konstante K, so dass für alle $t \in T$

$$|f_t(\bar{x}) - f_t(y)| \leq K\|\bar{x} - y\|_X \text{ für alle } \bar{x}, y \in X, \text{ und } \sup_T f_t(x) \leq K,$$

und darüberhinaus gelte $\partial f_t(y) = \{D_s f_t(y)\}$ für alle $y \in X$, und die Abbildung

$$(t, y) \mapsto D_s f_t(y)$$

sei stetig.
Zeigen Sie: Dann ist die Funktion $f(y) := \max_T f_t(y)$ regulär bei x, und es gilt

$$\partial f(x) = \left\{ \int_T D_s f_t(x)\, d\mu(t) : \mu \in P[M(x)] \right\}.$$

Hinweis: Das ist die Aussage von Corollar 4.11 aus der Vorlesung.

4.6 * Beweisen Sie Satz 4.9 unter der Voraussetzung (iv)', dass T metrisierbar ist. *Hinweis: Verwenden Sie dazu [IL74, p. 34].*

5 LAGRANGE-Multiplikatorregel & Variationsprinzip von EKELAND

Satz 5.1 [LAGRANGE-Multiplikatorregel]
Sei X ein $\mathbb{R}$-BANACHraum, C abgeschlossen in X. Wir betrachten in einer Umgebung von C LIPSCHITZ-stetige Funktionen $f : X \to \mathbb{R}$, $g : X \to \mathbb{R}^n$ und $h : X \to \mathbb{R}^m$ mit

$$|(f,g,h)(c_1) - (f,g,h)(c_2)| \leq K \|c_1 - c_2\|_X \qquad \text{für alle } c_1, c_2 \in C.$$

Das Infimum von f auf der Menge

$$B := \{y \in C \mid g_i(y) \leq 0, i = 1, \ldots, n, h_j(y) = 0, j = 1, \ldots, m\}$$

werde in $x \in B$ angenommen. DANN gibt es für alle $\hat{K} > K$ nicht simultan verschwindende

$$\lambda \geq 0, \qquad r \in \mathbb{R}^n, r_i \geq 0, i = 1, \ldots, n, \qquad s \in \mathbb{R}^m,$$

so dass die Funktion

$$\begin{array}{rccl} L: & X \times (\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m) \times \mathbb{R} & \to & \mathbb{R} \\ & (y, (\lambda, r, s), \hat{K}) & \mapsto & \lambda f(y) + g(y) \cdot r + h(y) \cdot s + \hat{K} \, |(\lambda, r, s)| \, d_C(y) \end{array}$$

die Bedingungen

$$g(x) \cdot r = 0 \qquad \text{und} \qquad 0 \in \partial_y L(x, (\lambda, r, s), \hat{K})$$

erfüllt.

Bemerkungen.

- Wegen $r_i \geq 0$ haben wir auf B

$$r \cdot g(x) = 0 \iff \Big(g_i(x) < 0 \Rightarrow r_i = 0\Big).$$

- Falls Ungleichungs- bzw. Gleichungsnebenbedingungen nicht auftreten, so verschwinden die entsprechenden Terme in $0 \in \partial_y L$.
- Es würde genügen, B durch $B \cap B_\delta(x)$ zu ersetzen.
- Aus Propositionen 1.25, 1.27 und 2.4 erhalten wir

$$0 \in \partial_y L(x, (\lambda, r, s), \hat{K}) \subset \lambda \partial f(x) + \textstyle\sum_{i=1}^n r_i \partial g_i(x) + \sum_{j=1}^m s_j \partial h_j(x) + \hat{K} \, |(\lambda, r, s)| \underbrace{\partial d_C(x)}_{\subset N_C(x)}.$$

- Ist (λ, r, s) ein LAGRANGE-Multiplikator, so auch $(\kappa\lambda, \kappa r, \kappa s)$ für $\kappa > 0$, so dass man $|(\lambda, r, s)| = 1$ oder $\lambda = 1$ normieren kann.

Beweis von Satz 5.1. Sei

$$T := \left\{ t = (\lambda, r, s) \,\middle|\, |t| = 1, \lambda \geq 0, r \in \mathbb{R}^n \text{ mit } r_i \geq 0, i = 1, \dots, n, s \in \mathbb{R}^m \right\}.$$

Wir definieren eine Abbildung $F : X \to \mathbb{R}$ durch

$$F(y) := \max_{t=(\lambda,r,s)\in T} (\lambda, r, s) \cdot (f(y) - f(x) + \varepsilon, g(y), h(y)),$$

es ist also insbesondere $F(x) = \max_T \left\{ \lambda\varepsilon + r \cdot \underbrace{g(x)}_{\leq 0} + s \cdot \underbrace{h(x)}_{=0} \right\} = \varepsilon$. Für $\hat{K} > K$ können wir $\varepsilon > 0$ so klein wählen, dass die Abbildung

$$G^{(z)} := F + \sqrt{\varepsilon}\, \|z - \cdot\|_X$$

für alle $z \in \overline{B_1(x)}$ LIPSCHITZ-stetig mit LIPSCHITZ-Konstante $\hat{K}$ ist. Es gilt $F > 0$ auf C. Anderenfalls folgte aus der Existenz eines $c \in C$ mit $F(c) \leq 0$

- $h(c) = 0$, denn sonst könnten wir $s \in \mathbb{R}^m$, $|s| = 1$, so wählen, dass $s \cdot h(c) > 0$. Dann wäre $(0, 0, s) \in T$ und $F(c) \geq (0, 0, s) \cdot (f(c) - f(x) + \varepsilon, g(c), h(c)) > 0$. ↯

- $g_i(c) \leq 0$ für $i = 1, \dots, n$, denn sonst könnten wir für ein i_* mit $g_{i_*}(c) > 0$ den Vektor r mit $r_{i_*} = 1$ und $r_{\neq i_*} = 0$ wählen, so dass $g(c) \cdot r > 0$. Dann wäre $(0, r, 0) \in T$ und $F(c) \geq (0, r, 0) \cdot (f(c) - f(x) + \varepsilon, g(c), h(c)) > 0$. ↯

- $f(c) - f(x) + \varepsilon \leq 0$, denn sonst könnten wir $\lambda = 1$ wählen. Dann wäre $(\lambda, 0, 0) \in T$ und $F(c) \geq (\lambda, 0, 0) \cdot (f(c) - f(x) + \varepsilon, g(c), h(c)) > 0$. ↯

Die ersten beiden Punkte implizieren $c \in B$, der dritte $f(c) \leq f(x) - \varepsilon$, was der Minimalität von $f(x)$ auf B widerspricht. Mithin $F > 0$ auf C.
Also gilt $F(x) \leq \inf_C F + \varepsilon$. Das EKELANDsche Variationsprinzip, Satz 5.2, liefert uns nun ein $u \in \overline{B_{\sqrt{\varepsilon}}(x)} \cap C$ mit $G^{(u)}(c) \geq G^{(u)}(u) = F(u)$. Aus Proposition 2.5 folgt wegen der LIPSCHITZ-Stetigkeit der $G^{(\cdot)}$ für alle y in einer Umgebung von u

$$\begin{aligned} G^{(u)}(u) + \hat{K}\underbrace{d_C(u)}_{} = 0 \quad &\leq \quad G^{(u)}(y) + \hat{K} d_C(y) \\ &= \quad \max_T\, (\lambda, r, s) \cdot (f(y) - f(x) + \varepsilon, g(y), h(y)) + \sqrt{\varepsilon}\, \|y - u\|_X + \hat{K} d_C(y) \\ &= \quad \underbrace{\max_T\, L(y, (\lambda, r, s), \hat{K}) - \lambda f(x) + \lambda\varepsilon}_{=:H(y)} + \sqrt{\varepsilon}\, \|y - u\|_X\,. \end{aligned}$$

Vermöge Propositionen 1.26 und 1.27 erhalten wir

$$0 \in \partial H(u) + \sqrt{\varepsilon}\partial(\|\cdot - u\|_X)(u) \quad = \quad \partial H(u) + \sqrt{\varepsilon}\, \overline{B_1(0)}.$$

Unser Ziel ist nun die Anwendung von Satz 4.9 auf $H(.)$. Hierzu zeigen wir

$$\partial_{[T]} L(y, t, \hat{K}) = \partial L(y, t, \hat{K}). \tag{5.1}$$

Für alle $t_1, t_2 \in T$ und y_1, y_2 in einer Umgebung von u gilt

$$|L(y,t_1,\hat{K}) - L(y,t_2,\hat{K})| = |(t_1 - t_2) \cdot ((f,g,h)(y_1) - (f,g,h)(y_2))| \leq |t_1 - t_2|\, \hat{K}\, \|y_1 - y_2\|_X .$$

Mit Proposition 1.5 (iv) erhalten wir

$$\partial\left(L(\cdot,t_1,\hat{K}) - L(\cdot,t_2,\hat{K})\right) \subset B_{2|t_1-t_2|\hat{K}}(0),$$

also $\partial L(\cdot,t_1,\hat{K}) \subset \partial L(\cdot,t_2,\hat{K}) + B_{2|t_1-t_2|\hat{K}}(0)$ nach Proposition 1.27. Sei nun $\zeta \in \partial_{[T]}(L(\cdot,t,\hat{K}))(y)$. Dann gibt es Folgen $y_i \to y$, $t_i \to t$ und $\zeta_i \xrightarrow{*} \zeta$ mit $\zeta_i \in \partial L(\cdot,t_i,\hat{K})(y_i) \subset \partial L(\cdot,t,\hat{K}) + B_{2|t_i-t|\hat{K}}(0)$. Also finden wir eine Folge $(\eta_i)_{i\in\mathbb{N}}$ aus $\partial L(\cdot,t,\hat{K})(y)$ mit $\eta_i \xrightarrow{*} \zeta$. Mit Proposition 1.10 (i) folgt wie gewünscht $\partial_{[T]}L(y,t,\hat{K}) \subset \partial L(y,t,\hat{K})$ und damit (5.1).
Da $F > 0$ auf C, gibt es genau ein $t_u \in T$ mit $F(u) = t_u \cdot (f(u) - f(x) + \varepsilon, g(u), h(u))$, d.h. $\int_T \partial L(u,t,\hat{K})\,\mathrm{d}\mu = \partial L(u,t_u,\hat{K})$, und (4.8) liefert

$$\begin{aligned} 0 \quad &\in \quad \left\{ \int_T \partial_{[T]} L(u,t,\hat{K})\,\mathrm{d}\mu(t) \,\middle|\, \mu \in P[M(u)] \right\} + \sqrt{\varepsilon}\,\overline{B_1(0)} \\ &\overset{(5.1)}{=} \quad \partial L(u,t_u,\hat{K}) + \sqrt{\varepsilon}\,\overline{B_1(0)}. \end{aligned}$$

Da $u \in \overline{B_{\sqrt{\varepsilon}}(x)}$, erhalten wir mit $\varepsilon \to 0$ eine Folge (u_ε) mit $u_\varepsilon \xrightarrow{\varepsilon \searrow 0} x$, und wir dürfen annehmen, dass $t_{u_\varepsilon} \to t^* \in T$, da T kompakt. Wegen der schwach$*$-Abgeschlossenheit von $(t,y) \mapsto \partial_{[T]}L(y,t,\hat{K})$ folgt $0 \in \partial L(x,t^*,\hat{K})$. $\square$

Satz 5.2 [Ekeland[25] 1975]
Sei (V,d) ein vollständiger metrischer Raum und $F : V \to \mathbb{R} \cup \{\infty\}$ unterhalbstetig und durch $-C$ nach unten beschränkt. Ferner gebe es ein $u \in V$ und $\varepsilon > 0$ mit $F(u) \leq \inf_V F + \varepsilon$. DANN gibt es für alle $\lambda > 0$ ein $v \in V$ mit

(i) $F(v) \leq F(u)$,

(ii) $d(u,v) \leq \lambda$,

(iii) $F(v) + \frac{\varepsilon}{\lambda}\underbrace{d(v,v)}_{=0} \leq F(w) + \frac{\varepsilon}{\lambda} d(v,w)$ *für alle* $w \in V$.

Der Beweis folgt einer Idee aus [BP63].
Wir definieren für $\alpha > 0$ durch

$$(v_1,r_1) \succeq_\alpha (v_2,r_2) :\Longleftrightarrow r_1 - r_2 \geq \alpha d(v_1,v_2)$$

eine reflexive und transitive Teilordnung auf $V \times \mathbb{R}$. Die Menge $T^{(v_1,r_1)}_{\succeq_\alpha} := \{(v,r) \in V \times \mathbb{R} \,|\, (v_1,r_1) \succeq_\alpha (v,r)\}$ ist für alle (v,r) nicht-leer und (folgen)abgeschlossen. Wir formulieren und beweisen zunächst das folgende Hilfslemma.

Lemma 5.3
Sei S eine abgeschlossene Teilmenge von $V \times \mathbb{R}$, so dass ein $m \in \mathbb{R}$ mit $r \geq m$ für alle $(v,r) \in S$ existiert.
DANN gibt es zu jedem $(v_1,r_1) \in S$ genau ein $(\bar{v},\bar{r}) \in S \cap T^{(v_1,r_1)}_{\succeq_\alpha}$, das bezüglich $\succeq_\alpha$ minimal auf S ist.

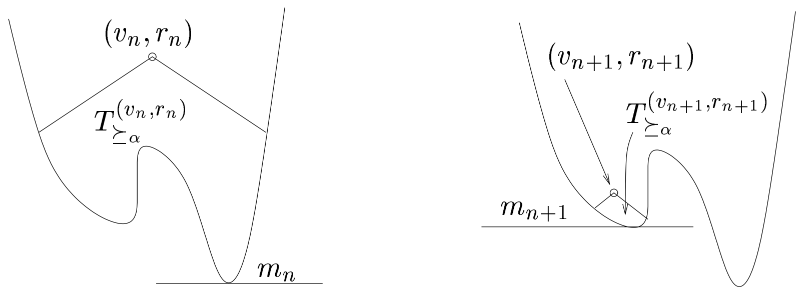

Abbildung 5.1: Zum Beweis von Lemma 5.3

Beweis. Wir konstruieren induktiv eine Folge $(v_i, r_i)_{i\in\mathbb{N}}$. (v_1, r_1) ist gegeben. Sei nun (v_n, r_n) bereits konstruiert. Wir setzen

$$m_n := \inf\left\{ r \in \mathbb{R} \,\middle|\, \exists v \in V \text{ mit } (v, r) \in T_{\succeq\alpha}^{(v_n,r_n)} \right\} \ \big(\geq m\big)$$

und wählen $(v_{n+1}, r_{n+1}) \in T_{\succeq\alpha}^{(v_n,r_n)}$ mit $r_n - r_{n+1} \geq \frac{1}{2}(r_n - m_n)$. Wegen der Transitivität haben wir $T_{\succeq\alpha}^{(v_{n+1},r_{n+1})} \subset T_{\succeq\alpha}^{(v_n,r_n)}$ und $m_{n+1} \geq m_n$. Dann gilt

$$0 \leq r_{n+1} - m_{n+1} \leq \frac{1}{2}(r_n - m_n) \leq 2^{-n}(r_1 - m),$$

also folgt für alle $(v, r) \in T_{\succeq\alpha}^{(v_{n+1},r_{n+1})}$

$$0 \leq r_{n+1} - r = r_{n+1} - m_{n+1} + \underbrace{m_{n+1} - r}_{\leq 0} \leq 2^{-n}(r_1 - m) \to 0$$

und $d(v_{n+1}, v) \leq \frac{r_{n+1}-r}{\alpha} \leq \frac{2^{-n}}{\alpha}(r_1 - m) \to 0$. Wegen der Vollständigkeit von V gibt es nun ein Grenzelement $(\bar{v},\bar{r}) \in V \times \mathbb{R}$ mit $\{(\bar{v}, \bar{r})\} = \bigcap_{n\in\mathbb{N}} T_{\succeq\alpha}^{(v_n,r_n)}$, das in der Tat minimal auf S ist: Sei (w, s) ein minimales Element auf S. Dann gilt $(v_n, r_n) \succeq_\alpha (w, s)$ für alle $n \in \mathbb{N}$, also $(w, s) \in \bigcap_{n\in\mathbb{N}} T_{\succeq\alpha}^{(v_n,r_n)}$. $\square$

Beweis von Satz 5.2. Wir wenden Lemma 5.3 auf $S :=$ epi F, $\alpha := \frac{\varepsilon}{\lambda} > 0$, $m := F(u) - \varepsilon$ und $(v_1, r_1) := (u, F(u))$ an. Dies ist möglich, denn $(w, s) \in$ epi F impliziert $s \overset{!}{\geq} F(w) \geq \inf F \overset{\text{Vorauss.}}{\geq} F(u) - \varepsilon = m$. Also erhalten wir ein in epi F minimales (v, r). Es ergibt sich sofort $r = F(v)$. Weiterhin erhalten wir wegen $(u, F(u)) \succeq_\alpha (v, F(v))$

$$F(u) - F(v) \geq \alpha d(u, v) \geq 0, \tag{5.2}$$

also (i). Die Minimalität liefert weiterhin für alle $w \neq v$

$$F(w) + \alpha \underbrace{d(v, w)}_{\geq 0} > F(v),$$

das ist (iii). Die Bedingung (ii) folgt aus (5.2) mit der Eigenschaft von u:

$$d(u, v) \leq \frac{F(u) - F(v)}{\alpha} \leq \frac{F(v) + \varepsilon - F(v)}{\alpha} = \frac{\varepsilon}{\alpha} = \lambda.$$

$\square$

[25] Ivar Ekeland (*1944) Französischer Mathematiker, Professor an der *Université Paris-Dauphine*.

6 Fixpunktsatz & Umkehrsatz

Fixpunktsatz

Sei T eine stetige Selbstabbildung auf einem vollständigen metrischen Raum (V, d).
Existiert ein $\sigma \in (0,1)$ mit

$$d(Tx, Ty) \leq \sigma d(x, y), \tag{6.1}$$

für alle $x, y \in V$, so gibt es nach dem Banach*schen Fixpunktsatz* genau ein $v \in V$ mit $Tv = v$.
Wir möchten nun eine Art „lokale Variante“ dieses Resultats herleiten. Dies ist jedoch problematisch: Fordert man z.B. (6.1) für eine Umgebung eines Punktes $z \in V$, so wird diese Bedingung für die diskrete Metrik (ein-Punkt-Umgebungen) trivial.
Zunächst erweitern wir die Definition von (x, y) durch

$$(x, y) := \{z \in V \setminus \{x, y\} \mid d(x, z) + d(z, y) = d(x, y)\}$$

auf metrische Räume.

Definition 6.1 (i) Ein metrischer Raum heißt *konvex*, falls $(x, y) \neq \emptyset$ für alle $x \neq y$, $x, y \in V$.

(ii) Wir nennen eine Abbildung $T : V \to V$ *gerichtete Kontraktion*, wenn sie stetig ist und es ein $\sigma \in (0,1)$ mit

$$\forall\, v \in V, Tv \neq v\ \exists\, w \in (v, Tv) : d(Tv, Tw) \leq \sigma d(v, w)$$

gibt.

Satz 6.2 [Fixpunktsatz]
Sei (V, d) ein vollständiger metrischer Raum und $T : V \to V$ eine gerichtete Kontraktion. DANN gibt es ein $v \in V$ mit $Tv = v$.

Bemerkung. Der Fixpunkt braucht nicht eindeutig sein, wie man an nachfolgendem Beispiel erkennt.

Beispiel 6.3 Sei $V := \mathbb{R}^2$, $d(u, v) := |u_1 - v_1| + |u_2 - v_2|$, wobei $u_i, v_i, i = 1, 2$ die Komponenten von u und v sind. Dann ist $(u, v) = [u_1, u_2] \times [v_1, v_2] \setminus \{u, v\}$. Wir betrachten die Abbildung

$$T : x \mapsto (\tfrac{3}{2}x_1 - \tfrac{1}{3}x_2, x_1 + \tfrac{1}{3}x_2).$$

Wegen $d(T(x_1, x_2), T(x_1', x_2)) = \frac{5}{2}|x_1 - x_1'|$ ist T *keine* Kontraktion, d. h. (6.1) gilt nicht global. Allerdings ist T eine *gerichtete* Kontraktion: Sei $y := Tx \neq x$. Dann ist $x_2 \neq y_2$

(andernfalls folgte sofort $x_1 = y_1$ und damit $x = y$). Wähle ein $w_2 \in (x_2, y_2)$. Dann ist $w := (x_1, w_2) \in (x, Tx)$, und wir haben $d(Tx, Tw) = \frac{2}{3}|x_2 - w_2| = \frac{2}{3}d(x,z)$. Mit $\sigma := \frac{2}{3}$ folgt die Behauptung. Nach Satz 6.2 gibt es einen nicht eindeutigen Fixpunkt; in der Tat ist *jeder* Punkt der Form $(t, \frac{3}{2}t)$, $t \in \mathbb{R}$, invariant unter T. ◇

Beweis von Satz 6.2. Die Abbildung $F : V \to \mathbb{R}$, $v \mapsto d(v, Tv)$ ist nicht-negativ und stetig, da $T : V \to V$ und $d : V \times V \to \mathbb{R}$ stetig sind. Unser Ziel ist die Minimierung von F. Dies kann jedoch wegen fehlender Kompaktheit nicht über eine Minimalfolge geschehen. Ein Ausweg besteht in der Anwendung des EKELANDschen Variationsprinzips, Satz 5.2, für $\varepsilon := \frac{1-\sigma}{2} > 0$ und $\lambda = 1$, wobei wir $F(v_i) \le \inf_V F + \epsilon$ für eine Minimalfolge $(v_i)_{i\in\mathbb{N}}$ für $i \gg 1$ verifizieren können. Wir erhalten also mit Satz 5.2 ein $v \in V$ mit

$$F(v) \le F(w) + \frac{1-\sigma}{2} d(v,w) \quad \text{für alle } w \in V. \tag{6.2}$$

Wenn $F(v) = 0$, sind wir fertig. Andernfalls, also falls $v \neq w$, leiten wir einen Widerspruch her. Wegen der gerichteten Kontraktionseigenschaft gibt es damit ein $w \in (v, Tv)$ mit

$$\begin{aligned}
0 &\le \sigma d(v,w) - d(Tv,Tw) \\
&\le \sigma d(v,w) + d(Tv,w) - d(w,Tw) \\
&\overset{w\in(v,Tv)}{=} \sigma d(v,w) + d(v,Tv) - d(v,w) - d(w,Tw) \\
&= (\sigma-1)d(v,w) + d(v,Tv) - d(w,Tw) \\
&\overset{(6.2)}{\le} (\sigma-1)d(v,w) + \frac{1-\sigma}{2}d(v,w) \\
&= \frac{\sigma-1}{2}d(v,w) \le 0.
\end{aligned}$$

also haben wir $v = w$ im Widerspruch zu $w \in (v, Tv)$. □

Umkehrsatz

Satz 6.4 [Umkehrsatz]
Sei $F : \mathbb{R}^n \to \mathbb{R}^n$ LIPSCHITZ-stetig nahe x_0 und Rang $(\partial F(x_0)) = n$, *d.h.* Rang $M = n$ *für alle $M \in \partial F(x_0) \subset \mathbb{R}^{n\times n}$.*
DANN gibt es Umgebungen U von x_0, V von $F(x_0)$ und eine LIPSCHITZ-stetige Funktion $G : V \to \mathbb{R}^n$ mit

$$G \circ F = \mathrm{id}_U \qquad \textit{und} \qquad F \circ G = \mathrm{id}_V.$$

Bemerkung.

(i) Im Fall $F \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ ist F strikt differenzierbar, also $\partial F(\cdot)$ einelementig. Nach dem klassischen Umkehrsatz ist dann auch $G \in C^1$.

(ii) Nach Satz 3.1 existiert $DF(\cdot)$ fast überall. Es genügt jedoch nicht, die Maximalität des Rangs an diesen Stellen zu fordern: Die reelle Betragsfunktion ist ein Gegenbeispiel.

Beispiel 6.5 Gegeben sei die Funktion $F:\mathbb{R}^2\to\mathbb{R}^2$, $(x,y)\mapsto(|x|+y, 2x+|y|)$. Nach Definition 3.6 ist

$$\begin{aligned}\partial F(0,0) &= \operatorname{conv}\left\{\begin{pmatrix}1&1\\2&1\end{pmatrix},\begin{pmatrix}1&1\\2&-1\end{pmatrix},\begin{pmatrix}-1&1\\2&1\end{pmatrix},\begin{pmatrix}-1&1\\2&-1\end{pmatrix}\right\}\\ &= \left\{\begin{pmatrix}s&1\\2&t\end{pmatrix}\,\middle|\, s,t\in[-1,1]\right\}.\end{aligned}$$

Wegen $\operatorname{Det}\left(\begin{smallmatrix}s&1\\2&t\end{smallmatrix}\right)=st-2\le -1$ können wir Satz 6.4 anwenden und erhalten die Existenz einer Umkehrabbildung $G=F^{-1}$ nahe $F(0,0)=(0,0)\in\mathbb{R}^2$. ◇

Zum Beweis von Satz 6.4 formulieren wir zunächst drei Lemmata.

Lemma 6.6
Unter den Voraussetzungen von Satz 6.4 gibt es $\delta, r>0$, so dass

$$\forall\, v\in\mathbb{S}^{n-1}\ \exists\, w\in\mathbb{S}^{n-1}\ \forall\, M\in\partial F(B_r(x_0)) : \langle w, Mv\rangle\ge\delta.$$

Beweis. Die Menge $\partial F(x_0)\mathbb{S}^{n-1}$ ist kompakt und enthält wegen der Rang-Bedingung nicht 0. Also gibt es ein $\delta>0$ mit $\operatorname{dist}(\partial F(x_0)\mathbb{S}^{n-1},0)\ge 2\delta$. Dann ist für $0<\varepsilon\ll 1$ immerhin

$$\operatorname{dist}(B_\varepsilon(\partial F(x_0)v),0)\ge\delta \tag{6.3}$$

für alle $v\in\mathbb{S}^{n-1}$. Nach Proposition 3.7 (iii) ist ∂F oberhalbstetig, d. h. es gibt ein $r>0$ mit $\partial F(B_r(x_0))\subset B_\varepsilon(\partial F(x_0))$. Nach eventueller Verkleinerung von r erhalten wir zusätzlich, dass F auf $\overline{B_r(x_0)}$ LIPSCHITZ-stetig mit einer LIPSCHITZ-Konstanten K ist.
Sei nun $v\in\mathbb{S}^{n-1}$ mit (6.3) vorgegeben. Wegen der Konvexität von $B_\varepsilon(\partial F(x_0))$ und (6.3) folgt nun aus dem Trennungssatz ([Alt02, Ü2.9]) die Existenz eines $w\in\mathbb{S}^{n-1}$ mit $\langle w,Mv\rangle\ge\delta$ für alle $M\in B_\varepsilon(\partial F(x_0))$, also speziell für alle $M\in\partial F(B_r(x_0))$. □

Lemma 6.7
Unter den Voraussetzungen von Satz 6.4 gilt für $\delta, r>0$ aus Lemma 6.6

$$|F(x)-F(y)|\ge\delta\,|x-y|\qquad\forall\, x,y\in\overline{B_r(x_0)}.$$

Beweis. Seien $x,y\in B_r(x_0)$, $x\ne y$, für Randelemente folgt die Behauptung mit der Stetigkeit von F. Wir setzen $v:=\frac{y-x}{|y-x|}\in\mathbb{S}^{n-1}$, $\lambda:=|y-x|$, also $y=x+\lambda v$. Für $v_x^\perp:=\{x+z\,|\,z\in\mathbb{R}^n,\langle z,v\rangle=0\}$ gilt nach einem FUBINI[26]-Argument (vgl. Aufgabe 3.1)

$$\mathscr{H}^1(\{x'+tv\,|\,t\ge 0\}\cap\operatorname{Sing}F)=0\quad\text{für }\mathscr{H}^{n-1}\text{-fast alle }x'\in v_x^\perp.$$

Damit gibt es ein $x'\in v_x^\perp$ beliebig nahe an x mit $x'+tv\in B_r(x_0)$ für alle $t\in[0,\lambda]$, und $DF(x'+tv)$ existiert für fast alle $t\in[0,\lambda]$.
Mithin gilt $F(x'+\lambda v)-F(x')=\int_0^\lambda DF(x'+tv)v\,\mathrm{d}t$. Nach Lemma 6.6 gibt es also ein $w=w(v)\in\partial B_1(0)=\mathbb{S}^{n-1}$ mit

$$\begin{aligned}\delta\,|x-y| &\le \int_0^\lambda w\cdot DF(x'+tv)v\,\mathrm{d}t = w\,\big(F(x'+\lambda v)-F(x')\big)\\ &\le \big|F(x'+\lambda v)-F(x')\big|.\end{aligned}$$

Die Behauptung folgt mit der Stetigkeit von F. □

[26] GUIDO FUBINI (1879 – 1943) Italienischer Mathematiker, Professor in Turin. Nach der Emigration 1939 arbeitete er in Princeton. FUBINI lieferte bedeutende Beiträge zur Differentialgeometrie, der Theorie diskontinuierlicher Gruppen und automorpher Funktionen.

Lemma 6.8
Unter den Voraussetzungen von Satz 6.4 gilt für $\delta, r > 0$ aus Lemma 6.6

$$B_{\frac{\delta r}{2}}(F(x_0)) \subset F(B_r(x_0)).$$

Beweis. Sei $y \in B_{\frac{\delta r}{2}}(F(x_0))$ und $x \in \overline{B_r(x_0)}$ mit $|y - F(x)|^2 = \min_{\overline{B_r(x_0)}} |y - F(\cdot)|^2$. Dann ist $x \in B_r(x_0)$, denn anderenfalls hätten wir

$$\begin{aligned} \frac{\delta r}{2} &> |y - F(x_0)| \geq |F(x) - F(x_0)| - |y - F(x)| \\ &\overset{\substack{\text{Lemma 6.7} \\ x \text{ minimal}}}{\geq} \delta\,|x - x_0| - |y - F(x_0)| \overset{x \in \partial B_r(x_0)}{>} \delta r - \frac{\delta r}{2} = \frac{\delta r}{2}. \qquad \lightning \end{aligned}$$

Proposition 1.26 liefert mit Satz 3.10

$$0 \in \partial(|y - F(\cdot)|^2)(x) \subset 2(y - F(x))^T \partial F(x).$$

Lemma 6.6 impliziert die Invertierbarkeit aller Matrizen aus $\partial F(B_r(x_0))$. Hieraus folgt $y - F(x) = 0$ und $y \in F(B_r(x_0))$. □

Beweis von Satz 6.4. Wir definieren die Funktion G auf $V := B_{\frac{\delta r}{2}}(F(x_0))$ durch

$$v \mapsto G(v) := x \in B_r(x_0) \quad \text{mit } F(x) = v,$$

was nach Lemma 6.8 möglich und nach Lemma 6.7 wohldefiniert ist. Wegen der Stetigkeit von F kann man eine Umgebung U von x_0 mit $F(U) \subset V$ finden. Lemma 6.7 impliziert außerdem die LIPSCHITZ-Stetigkeit von G mit Konstante $\frac{1}{\delta}$. □

Corollar 6.9 [Impliziter Funktionensatz]
Sei $H : \mathbb{R}^m \times \mathbb{R}^k \to \mathbb{R}^k$ LIPSCHITZ-*stetig nahe* $(\hat{y}, \hat{z})$, $H(\hat{y}, \hat{z}) = 0$ *und* Rang $\pi_z \partial H(\hat{y}, \hat{z}) = k$, *wobei*

$$\pi_z \partial H(\hat{y}, \hat{z}) = \left\{ M \in \mathbb{R}^{k \times k} \,\middle|\, \exists\, N \in \mathbb{R}^{k \times m} : (N, M) \in \partial H(\hat{y}, \hat{z}) \subset \mathbb{R}^{k \times (m+k)} \right\}.$$

DANN gibt es eine Umgebung $\Upsilon \subset \mathbb{R}^m$ *von* $\hat{y}$ *und eine nahe* $\hat{y}$ LIPSCHITZ-*stetige Funktion* $\zeta : \Upsilon \to \mathbb{R}^k$ *mit*

$$H(y, \zeta(y)) = 0 \quad \textit{für alle } y \in \Upsilon.$$

Beweis. Definiere $F : \mathbb{R}^{m \times k} \to \mathbb{R}^{m \times k}$, $(y, z) \mapsto (y, H(y, z))$. Nach Konstruktion ist F LIPSCHITZ-stetig nahe $(\hat{y}, \hat{z})$. An allen Stellen (y, z), an denen F differenzierbar ist, gilt

$$DF(y, z) = \begin{pmatrix} \mathbb{1}_m & 0 \\ D_y H(y, z) & D_z H(y, z) \end{pmatrix},$$

im Sinne von Definition 3.6 erhalten wir Rang $\partial F(\hat{y}, \hat{z}) = m + k$. Mithin dürfen wir Satz 6.4 anwenden und erhalten Umgebungen U von $(\hat{y}, \hat{z})$ und V von $F(\hat{y}, \hat{z}) = (\hat{y}, H(\hat{y}, \hat{z})) = (\hat{y}, 0)$ und eine nahe $(\hat{y}, 0)$ LIPSCHITZ-stetige Funktion $G : V \to \mathbb{R}^{m+k}$, $G(\cdot) = (G^1(\cdot), G^2(\cdot)) \in \mathbb{R}^m \times \mathbb{R}^k$ mit

$$\left(G^1(u, v), H\left(G^1(u, v), G^2(u, v)\right)\right) = F(G(u, v)) \overset{!}{=} (u, v) \quad \forall\, (u, v) \in V.$$

Der ersten Komponente entnehmen wir $G^1(u, v) = u$ für alle u nahe $\hat{y}$, hiermit vereinfacht sich die zweite zu $H(u, G^2(u, v)) = v$ für alle v nahe 0. Also liefert $\zeta(u) := G^2(u, 0)$ die Identität $H(u, \zeta(u)) = 0$ für alle u nahe $\hat{y}$. □

7 Elastische Stäbe mit Selbstkontakt

In der Natur treten viele „schlanke“ Objekte mit elastischen Eigenschaften und Selbstkontaktphänomenen auf. Ein einfaches Beispiel stellen Elektrokabel dar; das hier grob skizzierte Modell ist motiviert durch *Selbstkontaktphänomene* in der Mikrobiologie: DNA-Moleküle sind in der Regel verdrillt und oft auch verknotet. In einem Zellkern finden sich DNA-Moleküle von etwa 1.2 Meter Länge. Auf einen Basketball hochgerechnet, entspricht dies rund 200 Kilometern; ein hoher Verdrillungs- und Verknotungsgrad liefert also zumindest eine Möglichkeit, diese Länge äußerst kompakt in dem vorgegebenen Volumen des Zellkerns unterzubringen. Aber auch andere funktionelle Zusammenhänge zwischen Verdrillung (*supercoiling*) und Verknotung einerseits und den chemisch-biologischen Prozessen andererseits, denen DNA-Moleküle ausgesetzt sind, wurden inzwischen in der DNA-Forschung nachgewiesen. Auch auf anderen Skalierungsebenen stößt man auf Selbstkontaktphänomene: makrozelluläre bakterielle Verbände bilden verdrehte Strukturen ähnlich einem überbeanspruchten Telefonkabel, so etwa beim von Mendelsohn und Tilby untersuchten *bacillus subtilis*; elastische Filamente treffen auf elastische Membranen; elastische Drähte berühren sich selbst.
Im folgenden skizzieren wir die Modellierung elastischer geschlossener, schlanker Objekte mit Selbstkontakt im Rahmen der Cosserat[27]-*Theorie elastischer Stäbe.* Eine ausführliche Darstellung und Literaturhinweise findet man in [GMSvdM02], [SvdM03c, SvdM03b, SvdM03a].
Für vorgegebene Funktionen $u, v : I \to \mathbb{R}^3$, $I = (a, b)$ ein endliches reelles Intervall, und Anfangsdaten $\gamma(0) = \gamma_0 \in \mathbb{R}^3$, $D(0) = D_0 \in SO(3)$ beschreibt das System gewöhnlicher Differentialgleichungen

$$\begin{aligned} d_k'(s) &= \left(\textstyle\sum_{i=1}^3 u^i(s) d_i(s)\right) \wedge d_k(s), \\ \gamma'(s) &= \textstyle\sum_{i=1}^3 v^i(s) d_i(s) \end{aligned} \tag{7.1}$$

eine Kurve $\gamma : I \to \mathbb{R}^3$ und ein begleitendes Dreibein $D = (d_1, d_2, d_3) : I \to SO(3)$.

Man kann zeigen, dass für $p, q > 1$ genau dann eindeutige Lösungen $\gamma \in H^{1,q}(I, \mathbb{R}^3)$, $D \in H^{1,p}(I, SO(3))$ zu gegebenen Anfangswerten γ_0, D_0 existieren, die (7.1) für fast alle $s \in I$ erfüllen, wenn $u \in L^p(I, \mathbb{R}^3)$, $v \in L^q(I, \mathbb{R}^3)$.
Wir haben somit eine eineindeutige Beziehung von (γ, D) und

$$w := (u, v, \gamma_0, D_0) \in L^p(I, \mathbb{R}^3) \times L^q(I, \mathbb{R}^3) \times \mathbb{R}^3 \times SO(3) =: X$$

[27] Eugéne Maurice Pierre Cosserat (1866 – 1931) Französischer Mathematiker, Mechaniker und Astronom; Professor in Toulouse. Cosserat arbeitete neben seinen astronomischen Verpflichtungen als Direktor des Observatoriums (ab 1908) in Geometrie — seine Dissertation beschäftigt sich mit der Weiterentwicklung Plückers Theorie der Raumerzeugung durch gerade Linien — und Elastizitätstheorie. Hier wendete er beispielsweise die Methode des begleitenden Dreibeins auf Funktionalgleichungen von Ellipsoiden an und studierte Fragen der Flächenverbiegung. Cosserats mechanische Arbeiten wurden bald von der Relativitätstheorie überholt.

und schreiben $(\gamma[w], D[w])$. Für eine Funktion $W : \mathbb{R}^3 \times \mathbb{R}^3 \times I \to \mathbb{R}$ („*Energiedichte*") mit

(W1) $W(u, v, \cdot)$ messbar auf I für alle $(u, v) \in \mathbb{R}^3 \times \mathbb{R}^3$,

(W2) $W(\cdot, \cdot, s)$ konvex auf $\mathbb{R}^3 \times \mathbb{R}^3$ für alle $s \in I$,

(W3) $\exists\, c > 0 : c\,(|u|^p + |v|^q) \leq W(u, v, s)$ für alle $(u, v, s) \in \mathbb{R}^3 \times \mathbb{R}^3 \times I$
(Wachstums-/Koerzivitätsbedingung)

definieren wir auf X die *innere elastische Energie*

$$E(u, v, \gamma_0, D_0) := \int_I W(u(s), v(s), s)\, \mathrm{d}s.$$

Um in sich geschlossene Objekte zu modellieren, beschränken wir uns auf

$$\mathscr{L} := \left\{ w = (u, v, \gamma_0, D_0) \in X \,\middle|\, \gamma^{(w)}(a) = \gamma^{(w)}(b), D^{(w)}(a) = D^{(w)}(b) \right\}.$$

Weiterhin sollen Knotentyp κ und GAUSS[28]sche Verschlingungszahl $\ell_0 \in \mathbb{Z}$ (und damit der Verdrillungszustand des Objekts) vorgeschrieben werden, indem wir für $0 < \varepsilon \ll 1$

$$\mathscr{T} := \{ w = (u, v, \gamma_0, D_0) \in X \mid \gamma[w] \simeq \kappa, \ell(\gamma[w], \gamma[w] + \varepsilon d_1[w]) = \ell_0 \}$$

betrachten. Hierbei bezeichnen wir mit $\cong$ die Relation der *Isotopie*, d.h. grob gesprochen die Existenz einer doppelpunktfreien Homotopieabbildung zwischen $\gamma[w]$ und κ, während mit $l(\alpha, \beta)$ die GAUSSsche Verschlingungszahl (siehe [Spi70, S. 402]) zweier punktfremder Kurven $\alpha, \beta : I \to \mathbb{R}^3$ bezeichnet wird.
Wir untersuchen nun zu $\theta > 0$ Kurven γ mit

(C1) $\Delta[\gamma[w]] \geq \theta$ (globaler Krümmungsradius),

(C2) $\gamma[w] \in \mathscr{L} \cap \mathscr{T}$.

[28] CARL FRIEDRICH GAUSS (1777 – 1855) Deutscher Mathematiker, Astronom, Geodät und Physiker; Professor in Göttingen. GAUSS ist einer der bedeutendsten deutschen Mathematiker; seine Leistungen können hier nur ansatzweise skizziert werden. Aus armen Verhältnissen stammend, erhielt GAUSS, dessen außergewöhnliche Begabung schon früh aufgefallen war, durch Stipendien Zugang zu höherer Schulbildung. Ab 1795 studierte er an der Göttinger Universität. Schon zuvor hatte er nach Lektüre von Arbeiten NEWTONS, EULERS, LAGRANGEs u. a. eigene Forschungsergebnisse erzielt, z. B. zur Primzahlverteilung und Fehlerrechnung (Methode der kleinsten Quadrate). Die Konstruktion des regelmäßigen Siebzehnecks mit Zirkel und Lineal 1796, die Veröffentlichung seines Buches *Disquisitiones arithmeticae* (1801), in dem er die Theorie der Kongruenzen, die quadratischen Formen und die Kreisteilung behandelt, und der Beweis des *Fundamentalsatzes der Algebra* in seiner Dissertation 1799 ließen ihn zu den führenden Mathematikern seiner Zeit aufrücken. Weitere Arbeiten betreffen komplexen Zahlen, elliptische Funktionen, nicht-EUKLIDische Geometrie, Himmelsmechanik, Störungsrechnung (Berechnung der Bahn des Planetoiden *Ceres* 1801; Monographie *Theoria motus corporum coelestium ...* 1809), Reihenkonvergenz, transzendente Funktionen, Funktionentheorie; Geodäsie (Vermessung des Königreiches Hannover ab 1820); Optik, Kapillarität; Telegraphie und viele andere Gebiete. GAUSS war politisch konservativ. Den Bruch der liberalen Verfassung durch den Hannoveraner König bedauernd, hat er sich nicht zugunsten seiner sieben protestierenden Kollegen (darunter die Brüder GRIMM) eingesetzt. Auch stand er der 1848er Revolution wegen ihrer Unruhen ablehnend gegenüber, obgleich er für die Einigung Deutschlands eintrat.

Hierbei ist der *globale Krümmungsradius* eine rein geometrisch definierte Bedingung an die Kurve γ in Termen der Bogenlängenparametrisierung $\Gamma : [0, L] \to \mathbb{R}^3$:

$$\triangle[\gamma] := \inf_{\substack{s \neq s_1 \neq s_2 \neq s \\ s, s_1, s_2 \in [0,L]}} R(\Gamma(s), \Gamma(s_1), \Gamma(s_2)),$$

wobei $R(x, y, z)$ den Radius des (eindeutig bestimmten) kleinsten Kreises durch die Punkte $x, y, z \in \mathbb{R}^3$ ist. Man kann zeigen, dass die Bedingung (C1) die Doppelpunktfreiheit von γ impliziert; darüberhinaus gilt unter der Bedingung (C1), dass Γ von der Klasse $C^{1,1}([0, L], \mathbb{R}^3)$ ist und der Abschätzung $\|\Gamma''\|_{L^\infty([0,L],\mathbb{R}^3)} \leq \triangle[\gamma]^{-1}$ genügt, siehe [SvdM03c, Theorem 1]. Die Bedingungen (C1) und (C2) sind nichtlinear und nicht-konvex, zudem ist (C1) durch die Bildung eines Infimums eine *nichtglatte* Nebenbedingung.

Satz 7.1 [Existenzsatz]
Wenn überhaupt ein $\bar{w} \in X$ *mit* (C1), (C2) *existiert, und wenn* $|\tilde{\gamma}_0| \leq C$ *für alle* $\tilde{w} = (\tilde{u}, \tilde{v}, \tilde{\gamma}_0, \tilde{D}_0) \in X$*, dann gibt es auch ein* $w \in X$ *mit* (C1), (C2) *und*

$$E(w) = \inf_{X \text{ mit (C1), (C2)}} E(.).$$

Der Beweis verwendet die direkte Methode der Variationsrechnung (siehe [VdM03, Kapitel 5]), die Hauptarbeit liegt darin, zu zeigen, dass die Bedingungen (C1) und (C2) eine bezüglich der schwachen Konvergenz in X abgeschlossene Teilmenge $C \subset X$ definieren; für die Details verweisen wir auf [GMSvdM02].
Die in den Kapiteln 4 und 5 entwickelten Methoden der nichtglatten Analysis erlauben nicht direkt die Herleitung einer EULER-LAGRANGEschen Gleichung, da die Funktion

$$R(.,.,.) : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$$

nahe der Diagonalen $\{x = y = z\}$ singulär wird, was die Anwendung der nichtglatten Version der Kettenregel, Satz 1.36, und von Satz 4.9 zunächst verbietet. Wenn aber für den Minimierer $w \in X$ die physikalisch naheliegende *strikte* Ungleichung

$$\|\Gamma''[w]\|_{L^\infty([0,L],\mathbb{R}^3)} < \triangle[\gamma[w]]^{-1} \tag{7.2}$$

gilt, dann kann man beweisen, dass $\triangle[\gamma[w]] = \min_{s,\sigma \in Q} r(\Gamma[w](s), \Gamma[w](\sigma), \Gamma[w]'(\sigma))$ ist, wobei nun $r(x, y, t)$ der Radius des eindeutig bestimmten Kreises ist, der die Punkte $x, y \in \mathbb{R}^3$ enthält und in y tangential zu $t \in S^2$ ist. Die Parametermenge Q ist eine *kompakte* Teilmenge von $[0, L]^2 \subset \mathbb{R}^2$ mit *postivem* Abstand zur Diagonale $\{s = \sigma\}$, vgl. [SvdM03c]. Die nichtglatte LAGRANGE-Multiplikatorregel Satz 5.1, Satz 4.9 und eine geegnete Verfeinerung der Kettenregel, Satz 1.36, liefern unter geeigneten Glattheitsannahmen an $W = W(u, s)$ nach ausgedehnter Rechnung (siehe [SvdM03b]) die EULER-LAGRANGE-Gleichung

$$0 = \lambda_E W_u(u(s), s) + m_0 + \int_s^L d_3[w](t) \wedge (f_0 - f_c(t))\, dt \quad \text{für fast alle} \quad s \in [0, L],$$

wobei für $\tau \in [0, L]$,

$$f_c(\tau) := \int_{Q_\tau} \frac{\Gamma[w](s) - \Gamma[w](\sigma)}{|\Gamma[w](s) - \Gamma[w](\sigma)|}\, d\mu(s, \sigma), \tag{7.3}$$

$$Q_\tau := \{(s, \sigma) \in [0, L] \times [0, L] : \sigma \leq \tau \leq s\} \text{ für } \tau \in [0, L]. \tag{7.4}$$

Hierbei sind $\lambda_E \geq 0$, $f_0, m_0 \in \mathbb{R}^3$ LAGRANGE-Multiplikatoren und μ ist ein Radon Maß mit Träger in der *Kontaktmenge*, d.h. in $\{(s,\sigma) \in [0,L]^2 : r(\Gamma[w](s), \Gamma[w](\sigma), \Gamma[w]'(\sigma)) = \theta\}$.

Neben der expliziten Form der damit auch mechanisch sinnvoll interpretierbaren Kontaktmomente lässt sich aus dieser Gleichung auch höhere Regularität ablesen: $D[w] \in W^{2,\infty}([0,L], \mathbb{R}^{3\times 3})$ und $\gamma \in W^{3,\infty}([0,L], \mathbb{R}^3) \cong C^{2,1}([0,L], \mathbb{R}^3)$, was z.B. bedeutet, dass die Krümmung der minimierenden Kurve γ noch LIPSCHITZ-stetig ist.

Für den anderen Extremfall der Nichtscherbarkeit und unendlichen Dehnbarkeit (v beliebig, $u \equiv 0$) kann man mit ähnlichen Methoden auch eine EULER-LAGRANGEsche Gleichung für *ideale Knoten* herleiten, also für längenminimierende Kurven in einer vorgeschriebenen Knotenklasse mit einer gegebenen “Mindestdicke”, also unter der Bedingung (C1). Hier erhält man eine Darstellungsformel für den Krümmungsvektor γ'' des Minimierers, die Aussagen über die geometrische Gestalt der idealen Knoten zulässt, siehe [SvdM03a].

Anhang

Ausführliche Darstellungen der hier nur umrissenen Fakten findet man z. B. in [HS71] und [Rud91].

Topologie

Wir bezeichnen mit $\mathfrak{P}(X)$ die Potenzmenge von X.

Definition A.1 [Topologie]
Ein Paar $(M, \mathfrak{T})$ heißt *topologischer Raum*, wenn M eine Menge und $\mathfrak{T} \subset \mathfrak{P}(M)$ eine *Topologie* auf M ist, d. h. die folgenden Eigenschaften erfüllt:

(i) $\emptyset, M \in \mathfrak{T}$

(ii) $W_1, W_2 \in \mathfrak{T} \Longrightarrow W_1 \cap W_2 \in \mathfrak{T}$

(iii) $\mathscr{T} \subset \mathfrak{T} \Longrightarrow \bigcup_{T \in \mathscr{T}} T \in \mathfrak{T}$

Eine weitere Topologie $\mathfrak{S}$ auf M heißt *gröber* als $\mathfrak{T}$, wenn $\mathfrak{S} \subset \mathfrak{T}$.
$(M, \mathfrak{T})$ heißt HAUSDORFF-*Raum*, falls je zwei Punkte aus V disjunkte Umgebungen in $\mathfrak{T}$ haben.

Definition A.2 [Stetige reellwertige Funktionen]
Eine Funktion $f : M \to \mathbb{R}$ heißt *stetig* (bezüglich $\mathfrak{T}$), falls $f^{-1}(a, b) \in \mathfrak{T}$ für alle $a < b \in \mathbb{R}$.

Definition A.3 [Kompaktheit]
Ein topologischer Raum M heißt *kompakt*, wenn es zu jeder Überdeckung von M durch eine Familie offener Mengen $\{U_i\}_{i \in I}$ eine endliche Teilmenge $J \subset I$ mit $M \subset \bigcup_{i \in J} U_i$ gibt. Eine Teilmenge $A \subset M$ heißt *kompakt*, falls A mit der Relativtopologie $\mathfrak{T}|_A := \{U \cap A \,|\, U \in \mathfrak{T}\}$ ein kompakter topologischer Raum ist.

Eigenschaften A.4 kompakter topologischer Räume

(i) Eine abgeschlossene Teilmenge eines kompakten Raums ist kompakt.

(ii) Eine kompakte Teilmenge eines HAUSDORFF-*Raums ist abgeschlossen.*

Definition A.5 [Topologischer Vektorraum]
Sei V ein $\mathbb{R}$-Vektorraum und gleichzeitig $(V, \mathfrak{T})$ ein topologischer Raum. Dann heißt V *topologischer Vektorraum*, wenn

(i) $\{v\}$ für alle $v \in V$ abgeschlossen ist und

(ii) Addition und Skalarmultiplikation stetig sind.

Letzteres bedeutet, für alle $v_1, v_2 \in V$ und jede Umgebung W von v_1+v_2 gibt es Umgebungen W_1 von v_1 und W_2 von v_2 mit $W_1 + W_2 \subset W$, und für $v \in V, \alpha \in \mathbb{R}$ und jede Umgebung W von $\alpha \cdot v$ gibt es ein $r > 0$ und eine Umgebung U von v mit $B_r(\alpha) \cdot U \subset W$. ([Rud91], 1.6)

Eigenschaften A.6 topologischer Vektorräume
Sei V ein topologischer Vektorraum. DANN gilt:

(i) V ist ein HAUSDORFF*-Raum.*

(ii) $A, B \subset V \Longrightarrow \overline{A} + \overline{B} \subset \overline{A + B}$

(iii) Ist U Unterraum von V, so auch $\overline{U}$.

(iv) $C \subset V$, C konvex $\Longrightarrow \mathring{C}$, $\overline{C}$ konvex

(v) $D \subset V$, D beschränkt $\Longrightarrow \overline{D}$ beschränkt

Die Beweise findet man in [Rud91], 1.12 und 1.13 b,c,d,f.

Bemerkung. Ein normierter Vektorraum ist ein topologischer Vektorraum bezüglich der von der Norm induzierten Topologie.

Definition A.7 [Dualraum]
Die Menge V^* aller stetigen linearen Abbildungen $f : V \to \mathbb{R}$ wird *Dualraum* von V genannt.

Eigenschaften A.8 des Dualraums

(i) Ist $\langle \xi, v_1 \rangle_{V^* \times V} = \langle \xi, v_2 \rangle_{V^* \times V}$ für $v_1, v_2 \in V$ und alle *$\xi \in V^*$, so folgt $v_1 = v_2$.*

(ii) Ist X ein HILBERT[29]*-Raum, so ist*

$$x \mapsto \big(y \mapsto \langle y, x \rangle_{X \times X}\big)$$

ein isometrisch konjugiert linearer Isomorphismus $X \to X^*$. Dies ist die Aussage des RIESZ*schen*[30]*Darstellungssatzes, dessen Beweis in [Alt02], p. 147 f. zu finden ist.*

(iii) Jedes $x \in V$ induziert ein lineares Funktional $\tilde{x} \in V^{**}$ mittels

$$\xi \mapsto \langle \tilde{x}, \xi \rangle_{V^{**} \times V^*} := \langle \xi, x \rangle_{V^* \times V}, \qquad \xi \in V^*.$$

Im Folgenden und im Text identifizieren wir x und $\tilde{x}$.

[29] DAVID HILBERT (1863 – 1942) Deutscher Mathematiker, Professor in Königsberg und Göttingen. HILBERT war an der Entwicklung Göttingens zu einem führenden mathematischen Lehr- und Forschungszentrum maßgeblich beteiligt. Er war einer der bedeutendsten Mathematiker des 20. Jahrhunderts. Die auf einem Vortrag anlässlich des Mathematikerkongresses in Berlin von ihm formulierten 23 ungelösten Probleme wurden richtungsweisend für die Entwicklung der Mathematik in den folgenden Jahrzehnten.

[30] FRÉDÉRIC RIESZ (1880 – 1956) Ungarischer Mathematiker. Seine Hauptleistung besteht in der Übertragung von Begriffen und Resultaten der linearen Integrationstheorie auf allgemeine Räume und Operatoren. Neben BANACH ist er als Begründer der modernen Funkionalanalysis zu sehen. Weitere Beiträge betreffen Topologie, komplexe Funktionentheorie und subharmonische Funktionen.

(iv) Für $\dim V < \infty$ *gilt* $V \cong V^{**}$*; im Allgemeinen ist aber* $V \subsetneq V^{**}$*. Räume mit* $V \cong V^{**}$ *heißen* reflexiv.

Eigenschaften und Beispiele reflexiver Räume findet man in [Alt02], p. 218 ff., vgl. auch [VdM03], p. 50 – 52.

Definition A.9 [schwach∗-Topologie]
Auf dem Dualraum V^* eines topologischen Vektorraums V definieren wir die *schwach∗-Topologie* $\mathfrak{T}^*$ als die gröbste Topologie, bezüglich der alle $x \in V$ (aufgefasst als Abbildungen in V^{**}) stetig sind. Konkret geschieht dies durch beliebige Vereinigungen endlicher Schnitte von Urbildern offener Mengen unter Abbildungen $x \in V \subset V^{**}$.

Eigenschaften A.10 der schwach∗-Topologie

(i) V separiert V^*, *d. h. für* $\xi_1, \xi_2 \in V^*$, $\xi_1 \neq \xi_2$ *gibt es ein* $v \in V$ *mit* $\langle \xi_1, v\rangle_{V^* \times V} \neq \langle \xi_2, v\rangle_{V^* \times V}$.

(ii) Mit der schwach∗-Topologie $\mathfrak{T}^*$ *bildet* V^* *einen topologischen Vektorraum.*

Funktionalanalysis

Satz A.11 von Hahn[31]-Banach
Sei X *ein* $\mathbb{R}$*-Vektorraum; seien* $p : X \to \mathbb{R}$ *sublinear,* $f : Y \to \mathbb{R}$ *linear für einen Unterraum* $Y \subset X$ *und* $f \leq p$ *auf* Y*. DANN gibt es ein lineares* $F : X \to \mathbb{R}$ *mit* $F = f$ *auf* Y *und* $F \leq p$ *auf* X.

Ein Beweis ist in [Alt02], p. 163 zu finden.

Bemerkung. In unseren Anwendungen im Zusammenhang mit verallgemeinerten Richtungsableitungen erhalten wir aus Proposition 1.3 (iv) sogar Stetigkeit der linearen Fortsetzung.

Satz A.12 von Banach-Alaoglu
Sei V *ein* Banach*raum.*
DANN ist $\overline{B_1(0)} \subset V^*$ *schwach∗-kompakt bezüglich* $\mathfrak{T}^*$.

Einen Beweis findet man in [Rud91], 3.15, vgl. dort auch 3.16 und 3.17; für *separable* V wurde die Aussage in [VdM03], Satz 3.25 gezeigt.

Definition A.13 [Offene Abbildungen]
Eine Abbildung zwischen topologischen Räumen heißt *offen*, falls Bilder offener Mengen wieder offen sind.

Bemerkung. Sind X, Y normierte Räume und $f : X \to Y$ linear, so ist f genau dann offen, wenn es ein $\delta > 0$ mit $B_\delta(0) \subset f(B_1(0))$ gibt.

[31] Hans Hahn (1879 – 1934) Professor in Bonn und Wien. Hahn befasste sich mit Variationsrechnung, Fourieranalysis, mengentheoretischer Topologie und Grundlagentheorie. Vor allem lieferte er wichtige Beiträge zur linearen Funktionalanalysis. So führe er 1922 unabhängig von Banach den Begriff des *vollständigen normierten linearen Raums* ein und bewies 1927, zwei Jahre vor Banach, mit Hilfe der *transfiniten Induktion* (unter Verwendung des *Auswahlaxioms*) den hier zitierten Satz.

Satz A.14 von der offenen Abbildung
Seien X, Y BANACHräume. DANN ist jede lineare Abbildung aus $L(X, Y)$ surjektiv genau dann, wenn sie offen ist.

Ein Beweis ist in [Alt02], p. 206 f. zu finden.

Definition A.15 [Reguläre Maße]
Ein Maß μ heißt *regulär* , falls für alle messbaren Mengen E

$$\inf \{\mu(O \setminus K) \,|\, K \subset E \subset O,\, K \text{ kompakt},\, O \text{ offen}\} = 0.$$

Literaturverzeichnis

[Alt02] ALT, HANS WILHELM: *Lineare Funktionalanalysis. Eine anwendungsorientierte Einführung.* Springer, Berlin, 4., überarb. und erweiterte Auflage, 2002.

[Aub79] AUBIN, JEAN-PIERRE: *Applied functional analysis.* John Wiley & Sons, New York-Chichester-Brisbane, 1979. Translated from the French by Carole Labrousse, With exercises by Bernard Cornet and Jean-Michel Lasry.

[BF71] BONNESEN, T. und W. FENCHEL: *Theorie der konvexen Körper.* Chelsea Publishing Co., Bronx, N.Y., 1971. Reissue of the 1948 reprint of the 1934 original.

[BF89] BARNER, MARTIN und FRIEDRICH FLOHR: *Analysis II. 2., bearb. Aufl.* De Gruyter Lehrbuch. Berlin etc.: Walter de Gruyter. 449 S., 1989.

[BP63] BISHOP, ERRETT und R. R. PHELPS: *The support functionals of a convex set.* In: *Proc. Sympos. Pure Math., Vol. VII*, Seiten 27–35. Amer. Math. Soc., Providence, R.I., 1963.

[Cla83] CLARKE, FRANK H.: *Optimization and nonsmooth analysis.* Canadian Mathematical Society Series of Monographs and Advanced Texts. John Wiley & Sons Inc., New York, 1983. A Wiley-Interscience Publication.

[EG92] EVANS, LAWRENCE C. und RONALD F. GARIEPY: *Measure theory and fine properties of functions.* Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[Eva98] EVANS, LAWRENCE C.: *Partial differential equations*, Band 19 der Reihe *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 1998.

[GIS90] GOTTWALD, SIEGFRIED, HANS-JOACHIM ILGAUDS und KARL-HEINZ SCHLOTE (Herausgeber): *Lexikon bedeutender Mathematiker.* Bibliographisches Institut, Leipzig, 1990.

[GMSvdM02] GONZALEZ, O., J.H. MADDOCKS, F. SCHURICHT und H. VON DER MOSEL: *Global curvature and self-contact of nonlinearly elastic curves and rods.* Calc. Var. Partial Differ. Equ., 14(1):29–68, 2002.

[Hör55] HÖRMANDER, LARS: *Sur la fonction d'appui des ensembles convexes dans un espace localement convexe.* Ark. Mat., 3:181–186, 1955.

[HS71] Hirzebruch, Friedrich und Winfried Scharlau: *Einführung in die Funktionalanalysis.* Bibliographisches Institut, Mannheim, 1971. B. I.-Hochschultaschenbücher, No. 296*.

[IL74] Ioffe, A.D. und V.L. Levin: *Subdifferentials of convex functions.* Trans. Moscow Math. Soc., 26(1972):1–72, 1974.

[Leb75] Lebourg, Gérard: *Valeur moyenne pour gradient généralisé.* C. R. Acad. Sci. Paris Sér. A-B, 281(19):Ai, A795–A797, 1975.

[McL65] McLeod, R.M.: *Mean value theorems for vector valued functions.* Proc. Edinb. Math. Soc., II. Ser., 14:197–209, 1965.

[Roc70] Rockafellar, R. Tyrrell: *Convex analysis.* Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.

[Rud87] Rudin, Walter: *Real and complex analysis.* McGraw-Hill Book Co., New York, Dritte Auflage, 1987.

[Rud91] Rudin, Walter: *Functional analysis.* International Series in Pure and Applied Mathematics. McGraw-Hill Inc., New York, Zweite Auflage, 1991.

[RV74] Roberts, A. W. und D. E. Varberg: *Another proof that convex functions are locally Lipschitz.* Amer. Math. Monthly, 81:1014–1016, 1974.

[Spi70] Spivak, Michael: *A comprehensive introduction to differential geometry. Vol. One.* Published by M. Spivak, Brandeis Univ., Waltham, Mass., 1970.

[SvdM03a] Schuricht, Friedemann und Heiko von der Mosel: *Characterization of ideal knots.* erscheint in Calc. Var., DOI: 10.1007/s00526-003-0216-y, 2003.

[SvdM03b] Schuricht, Friedemann und Heiko von der Mosel: *Euler-Lagrange equations for nonlinearly elastic rods with self-contact.* Arch. Rat. Mech. Anal., 168(1):35–82, 2003.

[SvdM03c] Schuricht, Friedemann und Heiko von der Mosel: *Global curvature for rectifiable loops.* Math. Z., 243(1):37–77, 2003.

[VdM03] Mosel, Heiko Von der: *Variationsrechnung I.* `http://www.math.uni-bonn.de/people/heiko/VarI.pdf`, Mathematisches Institut der Universität Bonn, 2003.

[War81] Warga, J.: *Fat homeomorphisms and unbounded derivate containers.* J. Math. Anal. Appl., 81(2):545–560, 1981.

[War82] Warga, J.: *Errata: "Fat homeomorphisms and unbounded derivate containers" [J. Math. Anal. Appl. 81 (1981), no. 2, 545–560; MR 83f:58007].* J. Math. Anal. Appl., 90(2):582–583, 1982.

Index