

**Analysis II\*** von  
**Prof. PhD. A. Griewank**

14. Oktober 2009

Vorlesungsmitschrift von Paul Boeck

## Inhaltsverzeichnis

§28 Stetigkeit von Funktionen . . . . .	1
<b>VIII Differentialrechnung in mehreren Variablen</b>	<b>9</b>
§29 Differenzierbarkeit . . . . .	9
§30 Umkehrfunktion, Gleichungen, IFT . . . . .	16
§31 Unrestringierte und gleichheitsrestringierte Optimierung . . . . .	22
§32 Wege, Kurven und Weglängen . . . . .	28

## §28 Stetigkeit von Funktionen

### Definition 28.2

Eine Funktion /Abbildung  $f : X \rightarrow Y$  heißt stetig auf  $X$ , falls es an allen  $x_0 \in X$  stetig ist. Diese Eigenschaft gilt genau dann, wenn die folgenden äquivalenten Bedingungen erfüllt sind:

- (i)  $Y \supset V$  offen  $\implies U = f^{-1}(V) \subset X$  offen
- (ii)  $Y \supset V$  geschlossen  $\implies U = f^{-1}(V)$  geschlossen

Man schreibt dann  $f \in C(X, Y)$  oder  $f \in C(X)$  falls  $Y = \mathbb{R}, \mathbb{C}$  oder im Kontext klar.

### Beweis

Äquivalenz von (i) und (ii) folgt durch Komplementierung.

„ $\implies$ “ Sei  $f$  auf  $X$  stetig. Dann folgt für beliebiges offenes  $V \subset Y$  entweder

- $U = f^{-1}(V) = \emptyset$ , d.h. Aussage ist trivial
- oder es gibt für jedes  $x_0 \in U = f^{-1}(V)$ , ein  $\varepsilon > 0$  sodass  $B_\varepsilon(f(x_0)) \subset V$  (da  $V$  offen).

Dann folgt aus Stetigkeit auf  $x_0$  dass  $B_\delta(x_0) \subset f^{-1}(B_\varepsilon(f(x_0))) \subset f^{-1}(V)$ . Also enthält  $U$  Umgebung  $B_\delta(x_0)$  für beliebiges  $x_0$  und ist somit offen.

„ $\impliedby$ “ Betrachte beliebiges  $B_\varepsilon(f(x_0))$  für  $x_0 \in X, \varepsilon > 0$ . Dann ist  $f^{-1}(B_\varepsilon(f(x_0)))$  offen und enthält somit  $B_\delta(x_0)$  für ein  $\delta > 0$ . Das impliziert Stetigkeit an  $x_0$  und damit auf ganz  $X$ .

### Korollar 28.3 (Stetigkeit der Komposition)

$$f \in C(X, Y) \wedge g \in C(Y, Z) \implies g \circ f \in C(X, Z)$$

### Beweis

$V \subset Z$  offen  $\implies g^{-1}(V) \subset Y$  offen  $\implies f^{-1}(g^{-1}(V)) \subset X$  offen, wegen Stetigkeit von  $g$  und  $f$ .  
Bleibt zu zeigen:

$$\begin{aligned} x \in f^{-1}(g^{-1}(V)) &\iff x \in (g \circ f)^{-1}(V) \\ f(x) \in g^{-1}(V) &\iff g(f(x)) \in V \end{aligned}$$

□

### Lemma 28.4 (Vererbung von Stetigkeit)

$f \in C(X, Y), g \in C(X, Z)$  impliziert, dass

$$(i) (f, g) : X \rightarrow Y \times Z, (f, g)(x) = (f(x), g(x)) \quad (f, g) \in C(X, Y \times Z)$$

(ii)

$$\begin{aligned} \alpha f(x) + \beta g(x) : X &\rightarrow Y && \text{falls } Y = Z \text{ linearer Raum} \\ \alpha f + \beta g &\in C(X, Y) && \alpha, \beta \in \mathbb{K} = \mathbb{C} \text{ oder } \mathbb{R} \end{aligned}$$

(iii)

$$\begin{aligned} f(x) \cdot g(x) : X &\rightarrow Y && \text{falls } Z = \mathbb{K}, Y \text{ linearer Raum} \\ f \cdot g &\in C(X, Y) \end{aligned}$$

(iv) Sei  $f : X \rightarrow Y$  linearer Raum,  $g : X \rightarrow Z = \mathbb{K}$  und  $f$  stetig an  $x_0 \in X$  sowie  $g$  stetig und an Stelle  $x_0$  mit  $g(x_0) \neq 0$ . Dann gilt für Umgebung  $U \ni x_0$ , dass  $\frac{f(x)}{g(x)} : U \rightarrow Y$  ist stetig an Stelle  $x_0$ .

(Beweis als Zeitvertreib)

**Definition 28.5 (Lipschitzstetigkeit und Hölderstetigkeit)**

$f : X \rightarrow Y$  heißt (global) Lipschitzstetig auf  $X$ , falls für eine Konstante  $L \in \mathbb{R}$  und alle  $x, y \in X$  gilt:

$$d(f(x), f(y)) \leq L \cdot d(x, y)$$

$f$  heißt Hölderstetig vom Grad  $p \in (0, 1]$ , falls

$$d(f(x), f(y)) \leq L_p \cdot [d(x, y)]^p$$

Also ist Hölderstetigkeit vom Grad 1 äquivalent zur Lipschitzstetigkeit.

Falls  $L$  bzw.  $L_p$  jeweils nur in einer beschränkten Teilmenge von  $X$  gelten, so heißt  $f$  lokal Lipschitz- bzw. Hölderstetig.

**Beispiel** 1.)  $f(x) = |\sin x|$  für  $x \in \mathbb{R}$  ist global Lipschitzstetig mit  $L = 1$

2.)  $f(x) = x^2$  ist lokal Lipschitzstetig, da für

$$|x| \leq b \geq |y| \implies |x^2 - y^2| = |x - y||x + y| \leq |x - y|(|x| + |y|) \leq |x - y|2b$$

sodass  $L = 2b$  auf Intervall  $[-b, b]$  gültig.

3.)  $f(x) = \sqrt{|x|}$  dann gilt  $|\sqrt{|x|} - \sqrt{|y|}| \leq |x - y|^{\frac{1}{2}} \cdot L_{\frac{1}{2}}$ . Wird später nachgewiesen.

**Lemma 28.6 (Vererbung von Lipschitzstetigkeit)**

$f, g$  wie in Lemma 28.4 ergibt  $(f, g)$  und  $\alpha f + \beta g$  Lipschitz-/Hölderstetig, falls  $f$  und  $g$  Lipschitz-/Hölderstetig sind.

$f \cdot g$  ist lokal Lipschitz-/Hölderstetig, falls  $f$  und  $g$  das sind.

**Definition 28.7 (Lineare Operatoren)**

Funktion  $f : X \rightarrow Y$  zwischen linearen Vektorräumen schreibt man in Anwendung auf Argument  $x$  häufig in multiplikativen Form:  $f(x) = Ax$  wobei für  $X = \mathbb{C}^n, Y = \mathbb{C}^m$   $A$  als Matrix gedeutet werden kann.

$$Ae_j = \begin{pmatrix} \alpha_{1j} \\ \alpha_{2j} \\ \vdots \\ \alpha_{mj} \end{pmatrix} \text{ für } j = 1 \dots n \quad = \sum_{i=1}^m \alpha_{ij} e_i$$

mit  $e_j, e_i$  die Cartesischen Vektoren  $e_j = (0, \dots, 1, \dots, 0)$  in  $\mathbb{C}^n$  bzw.  $\mathbb{C}^m$ .

Es gilt dann als MatrixVektorProdukt für  $x = \sum x_j e_j$ :

$$Ax = \sum_{j=1}^n Ax_j e_j = \sum_{j=1}^n (Ae_j)x_j = \sum_{j=1}^n \left( \sum_{i=1}^m \alpha_{ij} e_i \right) x_j = \sum_{j=1}^n \left( \sum_{i=1}^m \alpha_{ij} x_j \right) e_i$$

M.a.W. die Koeffizienten von  $Ax$  bezüglich der Cartesischen Basis ergeben sich durch Multiplikation des Komponentenvektors von  $x$  mit  $A$ .

In unendlichdimensionalen Räumen nutzt man fast nie spezielle Basen und bezeichnet  $A$  gerne als „Linearen Operator“.

**Lemma 28.8**

Für eine lineare Abbildung/Operator  $A : X \rightarrow Y$  sind die folgenden Aussagen äquivalent:

(i)  $A$  Lipschitzstetig auf ganz  $X$

(ii)  $A$  stetig am Ursprung  $x = 0$  oder beliebigem anderen Punkt

(iii)  $A$  ist beschränkter Operator, d.h.  $\|A\| \equiv \sup_{0 \neq x \in X} \frac{\|Ax\|}{\|x\|} < \infty$

**Beweis**

(i)  $\implies$  (ii) offensichtlich

(ii)  $\implies$  (iii) Betrachte beliebiges  $\varepsilon > 0$  und entsprechend  $\delta > 0$  sodass

$$\|x - 0\| = \|x\| \leq \delta \implies \|Ax - A0\| = \|Ax\| \leq \varepsilon$$

Speziell erhalten wir für  $x \neq 0$

$$\frac{\|Ax\|}{\|x\|} = \left\| A \frac{x}{\|x\|} \right\| = \frac{\left\| A \frac{x\delta}{\|x\|} \right\|}{\delta} \leq \frac{\varepsilon}{\delta} \implies \|A\| \leq \frac{\varepsilon}{\delta} < \infty$$

(iii)  $\implies$  (i)

$$\|Ax - Ay\| = \|A(x - y)\| \leq \|A\| \|x - y\|$$

Also ist  $L = \|A\|$  die globale Lipschitzkonstante.

**Korollar 28.9**

Der Raum  $L(X, Y)$  aller beschränkten linearen Abbildungen zwischen zwei reellen oder komplexen Vektorräumen mit Normen  $\| \cdot \|_x$  und  $\| \cdot \|_y$  ist selbst ein normierter Raum bzgl. der induzierten Norm

$$\|A\| = \sup_{0 \neq x \in Y} \frac{\|Ax\|_y}{\|x\|_x}$$

**Beweis**

$$\|A\| = 0 \iff \|Ax\|_y = 0 \text{ für alle } x \iff Ax = 0 \text{ für } x \in X \iff A \equiv 0 \text{ Nulloperator}$$

$$\|\alpha A\| = \sup_{0 \neq x \in X} \frac{\|\alpha Ax\|_y}{\|x\|_x} = \alpha \sup \dots = \alpha \|A\| \implies \text{Homogenität}$$

$$\|A + B\| = \sup_{0 \neq x \in X} \frac{\|Ax + Bx\|_y}{\|x\|_x} \leq \sup_{0 \neq x \in X} \frac{\|Ax\|_y}{\|x\|_x} + \sup_{0 \neq x \in X} \frac{\|Bx\|_y}{\|x\|_x} = \|A\| + \|B\|$$

□

**Korollar 28.10**

$X, Y$  Banachräume  $\implies L(X, Y)$  auch Banach

**Beweis**

Für Normeigenschaften siehe Korollar 28.8 und Vollständigkeit siehe Übungsblatt 7.

**Bemerkung (Wozu das Ganze?)** In vielen Anwendungen sucht man

- i) Lösungen  $x \in X$  von linearen oder nichtlinearen Gleichungen  $F(x) = y$  für gegebenes  $y \in Y$
- ii) Optimalpunkte  $x$  von Variationsproblemen:

$$\min_{x \in X} f(x), \text{ sodass } F(x) = 0 \text{ mit } f : X \rightarrow \mathbb{R} \text{ und } F : X \rightarrow Y$$

Dazu erzeugt man Folgen  $(x_k)_{k=0}^\infty \subset X$  von approximativen Lösungen  $x_k$ , die hoffentlich gegen eine exakte Lösung bzw. einen Optimalpunkt  $x_* \in X$  konvergieren. Für Validierung der Existenz und Lösungseigenschaft von  $x_* \in X$  betrachtet man

- Stetigkeit von  $f$  und  $F$
- Abgeschlossenheit bzw. Vollständigkeit von  $X$
- Kontraktionseigenschaft von  $F$  bzw. im Banachraum  $F(x) - x$ . (siehe Fixpunktsätze)

**Lemma 28.11 (Umschreiben von Gleichung  $F(x) = y$  zu Fixpunktgleichung)**

Falls  $X, Y$  Banachräume und  $P \in L(X, Y)$  injektiv, d.h.  $Px = 0 \iff x = 0$ , gilt für festes  $y \in Y$ :

$$F(x) = y \iff G(x) := x - P(F(x) - y) = x$$

**Beweis**

Übungsaufgabe

**Satz 28.12 (Kontraktionssatz, Banach'scher Fixpunktsatz)**

Falls eine Abbildung  $G : X \rightarrow X$  auf vollständigem, metrischem Raum kontraktiv ist, d.h. für eine Lipschitzkonstante  $L < 1$  gilt

$$d(G(x), G(z)) \leq Ld(x, z) \text{ für } x, z \in X$$

dann gilt:

- (i) Es existiert genau ein Fixpunkt  $x_* \in X$ , d.h.  $G(x_*) = x_*$
- (ii) Für beliebigen Startwert  $x_0 \in X$  konvergiert die durch die Rekursion  $x_k = G(x_{k-1})$  für  $k = 1, 2, \dots$  erzeugte Folge  $(x_k)_{k=0}^\infty \subset X$  gegen  $x_* \in X$
- (iii)  $d(x_k, x_*) \leq L^k d(x_0, x_*)$  d.h. lineare Konvergenz sowie  $d(x_k, x_*) \leq \frac{L}{1-L} d(x_k, x_{k-1})$  (Größe des letzten Schrittes).

**Beweis**

(i)

$$\begin{aligned} d(x_k, x_{k+1}) &= d(G(x_{k-1}), G(x_k)) \\ &\leq Ld(x_{k-1}, x_k) \leq L^2 d(x_{k-2}, x_{k-1}) \quad \dots \text{ per Induktion} \end{aligned}$$

$\implies$  für  $m > k$

$$\begin{aligned} d(x_k, x_m) &\leq \sum_{j=k}^{m-1} d(x_j, x_{j+1}) \quad \text{nach Dreiecksungleichung} \\ &\leq \sum_{j=k}^{m-1} L^j d(x_0, x_1) = d(x_0, x_1) \sum_{j=k}^{m-1} L^j \\ &\leq d(x_0, x_1) L^k \sum_{j=0}^{\infty} L^j = \frac{d(x_0, x_1) L^k}{1-L} \end{aligned}$$

Also bilden  $x_k$  Cauchyfolge und haben wegen vorausgesetzter Vollständigkeit einen eindeutigen Grenzwert  $x_* \in X$ .

$$x_* = \lim_{k \rightarrow \infty} x_k = \lim_{k \rightarrow \infty} G(x_{k-1}) \stackrel{\text{da } G \text{ stetig}}{=} G\left(\lim_{k \rightarrow \infty} x_{k-1}\right) = G(x_*)$$

Daraus folgt  $x_*$  ist Fixpunkt wie in (i) behauptet. Eindeutigkeit folgt durch Widerspruch. Annahme:

$$\begin{aligned} x_* &= G(x_*), \quad \tilde{x}_* = G(\tilde{x}_*) \text{ mit } x_* \neq \tilde{x}_* \implies d(x_*, \tilde{x}_*) > 0 \\ d(x_*, \tilde{x}_*) &= d(G(x_*), G(\tilde{x}_*)) \leq Ld(x_*, \tilde{x}_*) \\ 1 &\leq L \text{ im Widerspruch zur Kontraktivität} \end{aligned}$$

- (ii)  $x_k \rightarrow x_*$  für beliebiges  $x_0$  bereits bewiesen

(iii)

$$\begin{aligned}d(x_k, x_*) &= d(G(x_{k-1}), G(x_*)) \leq Ld(x_{k-1}, x_*) \\ &\leq L^2d(x_{k-2}, x_*) \dots \leq L^k d(x_0, x_*) \\ d(x_k, x_*) &\leq d(x_k, x_{k+1}) + d(x_{k+1}, x_*) && \text{nach Dreiecksungleichung} \\ &\leq Ld(x_{k-1}, x_k) + Ld(x_k, x_*) \\ (1 - L)d(x_k, x_*) &\leq Ld(x_{k-1}, x_k) \implies \text{Behauptung nach Division durch } (1 - L)\end{aligned}$$

**Bemerkung** Anwendung für Banachraum  $X$  später zum Beweis von

- Umkehrfunktionentheorem
- implizite Funktionentheorem
- Theorem von Picard -Lindelöf garantiert Existenz von Lösungen für DGL's bzw. ODE's.

**Definition 28.13 (Kompakter metrischer Raum)**

Ein metrischer (Teil)Raum  $M \subset X$  heißt kompakt, wenn es die folgenden äquivalenten Bedingungen erfüllt:

- (i) **Folgenkompaktheit:** Jede Folge  $(x_k)_{k=1}^\infty \subset M$  hat eine konvergente Teilfolge mit Grenzwert  $x_* \in M$
- (ii) **Überdeckungskompaktheit:** Für jede Familie offener Mengen  $U_j \subset X$  mit  $j$  Element einer beliebigen Indexmenge  $J$  folgt aus  $M \subset \bigcup_{j \in J} U_j$  dass bereits für eine endliche Teilmenge  $J' \subset J$  gilt

$$M \subset \bigcup_{j \in J'} U_j$$

In Worten: Jede offene Überdeckung von  $M$  hat eine endliche Teilüberdeckung.

**Beweis**

- (ii)  $\implies$  (i) durch Widerspruch. Annahme  $(x_k) \subset M$  hat keinen Häufungspunkt. Dann gilt für jedes  $x \in M$  existiert eine Kugel  $B_r(x) \equiv U_x$ , die nur endlich viele Elemente von  $(x_k)_{k=1}^\infty$  enthält. Nach Definition von  $U_x$  mit  $J = M$  gilt

$$M = \{x \in M\} \subset \bigcup_{x \in M} U_x$$

Aus vorausgesetzter Überdeckungskompaktheit gibt es eine endliche Teilüberdeckung

$$M \subset \bigcup_{x \in \{z_1, \dots, z_m\}} U_x$$

Da in jeder dieser Umgebungen  $U_{z_j}$  nur endlich viele Elemente der Folge  $x_k$  liegen können, dürfte  $M$  selbst nur endlich viele Folgenglieder enthalten. Das steht im Widerspruch zur Definition von  $(x_k)_{k=1}^\infty$  als unendliche Folge.

**Bemerkung** Die in metrischen Räumen äquivalente Definition (ii) lässt sich auch auf topologische Räume übertragen, wo sie nicht mehr automatisch mit Konvergenz von Teilfolgen äquivalent ist.

**Satz 28.14**

- (i) Ein kompakter metrischer Raum  $X$  ist vollständig und beschränkt.
- (ii) Eine Teilmenge  $M$  eines endlichdimensionalen reellen oder komplexen Raumes ist kompakt gdw. sie abgeschlossen und beschränkt ist. (siehe Satz 13.4)

### Beweis

- (i) durch Widersprüche. Wäre  $X$  nicht vollständig, dann gäbe es eine Cauchyfolge  $(x_n) \subset X$  ohne Grenzwert in  $X$  und damit auch ohne konvergente Teilfolge. Wäre  $X$  unbeschränkt, so gäbe es eine Folge  $(x_n)_k$  mit  $d(x_k, x_*) \geq k$  für festes  $x_*$ . Daraus folgt für  $k$  fest und  $m \rightarrow \infty$

$$\liminf_{m \rightarrow \infty} d(x_m, x_n) \geq \lim_{m \rightarrow \infty} d(x_m, x_*) - d(x_k, x_*) = \infty$$

Also könnte es wiederum keine konvergente Teilfolge geben.

### Satz 28.15

Sei  $X, Y$  normierte Räume, falls  $f : X \rightarrow Y$  stetig auf kompakten Mengen  $M \subset X$  dann gilt:

- (i)  $f(M) \subset Y$  ist kompakt  
(ii)  $f$  ist auf  $M$  gleichmäßig stetig, d.h.  $\forall \varepsilon > 0 \exists \delta > 0 : d(x, y) < \delta \implies d(f(x), f(y)) < \varepsilon$  für  $x, y \in M$

### Beweis

- (i) Betrachte Folge  $(y_k)_{k=1}^\infty \subset f(M) \implies \exists x_k \in M : f(x_k) = y_k$ . Daraus folgt,  $(x_k)_{k=1}^\infty \subset M$  hat konvergente Teilfolge  $x_{k_j} \xrightarrow{j} x_* \in M$ . Daraus folgt,  $\lim f(x_{k_j}) = \lim y_{k_j} = f(x_*) \equiv y_*$  wegen Stetigkeit von  $f$ .

- (ii) Per Widerspruch. Annahme:

$$\exists \varepsilon > 0 \forall \delta > 0 \exists x, y \in M : d(x, y) < \delta \wedge d(f(x), f(y)) \geq \varepsilon$$

Insbesondere folgt für  $\delta = \frac{1}{k}$  Existenz von  $x_k, y_k \in M$  sodass  $d(x_k, y_k) < \frac{1}{k} \wedge d(f(x_k), f(y_k)) \geq \varepsilon$ . Wegen Kompaktheit, o.B.d.A  $x_k \rightarrow x_* \wedge y_k \rightarrow y_*$ .

$$d(x_k, y_k) \rightarrow 0, \quad d(x_*, y_*) \leq d(x_k, x_*) + d(x_k, y_k) + d(y_k, y_*) \implies d(x_*, y_*) = 0 \implies x_* = y_* \\ \implies \text{Stetigkeit } d(f(x_k), f(y_k)) \longrightarrow d(f(x_*), f(y_*)) = d(f(x_*), f(x_*)) = 0$$

im Widerspruch zur Annahme

□

### Korollar 28.16 (Weierstraß)

Falls  $f : X \rightarrow \mathbb{R}$  stetig und  $M \subset X$  kompakt existieren Punkte  $x_*, x^* \in M$  sodass

$$\underbrace{f_* = f(x_*)}_{\min\{f(x):x \in M\}} \leq f(x) \leq \underbrace{f^* = f(x^*)}_{\max\{f(x):x \in M\}} \quad \text{für } x \in M$$

### Beweis

Nach Satz 28.15 ist  $f(M) \subset \mathbb{R}$  kompakt und damit abgeschlossen und beschränkt. Folglich enthält  $f(M)$  sowohl sein Infimum als Minimum  $y_* = f(x_*)$  als auch sein Supremum als Maximum  $f(x^*) = y^*$ .

**Beispiel** Für kompakte Menge in unendlich dimensionalem Raum ergibt sich nach Arzela-Ascoli wie folgt:

### Satz 28.17

Im Banachraum  $C[a, b]$  mit  $\|f\|_\infty := \sup_{a \leq x \leq b} |f(x)|$  ist für jedes  $L > 0$  und  $l > 0$  die Menge

$$M_{L,l} := \{f \in C[a, b] : \|f\|_\infty \leq l \wedge |f(x) - f(y)| \leq L|x - y| \text{ für } x, y \in [a, b]\}$$

von Lipschitzstetigen Funktionen kompakt.

**Bemerkung** Satz erlaubt nach Peano Existenzbeweis für ODE's mit stetiger RHS.

### Beweis

Im Intervall  $[a, b]$  bilden die rationalen Punkte  $[a, b] \cap \mathbb{Q}$  eine dichte Teilmenge. Sie können wegen Abzählbarkeit von  $\mathbb{Q}$  durchnummeriert werden, d.h.  $\{q_1, q_2, \dots, q_j, \dots\} = [a, b] \cap \mathbb{Q}$ . Betrachte beliebige Folge  $(f_k)_{k=1}^\infty \subset M_{L,l}$ . Dann ist für jedes  $q_j$  die reelle Folge  $(f_k(q_j))_{k=1}^\infty \subset \mathbb{R}$  durch  $l$  beschränkt, hat also einen Häufungspunkt. Beginnend mit  $(f_k(q_j))_{k=1}^\infty$  definieren wir Teilfolge  $f_{k_i}$  von Funktionen, sodass  $f_{k_i}(q_1) \xrightarrow{j} f_*(q_1)$ . Aus den  $f_{k_i}$  entnehmen wir Teilfolge  $f_{k'_i}$  mit  $f_{k_1} = f_{k'_1}$  und  $f_{k'_i}(q_2) \xrightarrow{i} f_*(q_2)$ . Dieser Prozess kann abzählbar oft wiederholt werden und da man jeweils das erste Element dazunimmt ergibt sich eine Teilfolge  $f_{k_j^*}$  sodass

$$f_{k_j^*}(q_j) \xrightarrow{i} f_*(q_j) \text{ für } j \in \mathbb{N}$$

Zu zeigen bleibt, dass die  $f_{k_j^*}$  gleichmäßig gegen Grenzfunktion konvergieren, die für rationale  $q_j$  schon durch  $f_*(q_j)$  gegeben ist. O.b.D.A  $(f_{k_j^*})_{j=1}^\infty = (f_k)_{k=1}^\infty$ . Betrachte beliebiges  $\varepsilon > 0$  und setzen  $\delta = \frac{\varepsilon}{3L}$ . Dann hat die Überdeckung

$$[a, b] \subset \bigcup_{j \in \mathbb{N}} B_\delta(q_j) \implies [a, b] = \bigcup_{j=1}^n B_\delta(q_j)$$

hat wegen Kompaktheit endliche Teilüberdeckung. Wegen Konvergenz der  $f_k(q_j)$  an den endlich vielen Punkten  $q_1, \dots, q_n$  existiert ein  $k_0$ , sodass

$$k \geq k_0 \leq \tilde{k} \implies |f_k(q_j) - f_{\tilde{k}}(q_j)| \leq \frac{\varepsilon}{3}$$

Folglich ergibt sich aus der Dreiecksungleichung für beliebiges  $x \in B_\delta(q_j)$  dass

$$\begin{aligned} |f_k(x) - f_{\tilde{k}}(x)| &\leq |f_k(x) - f_k(q_j) + f_k(q_j) - f_{\tilde{k}}(x) + f_{\tilde{k}}(q_j) - f_{\tilde{k}}(q_j)| \\ &\leq \underbrace{|f_k(x) - f_k(q_j)|}_{\leq L\delta = \frac{\varepsilon}{3}} + \underbrace{|f_k(q_j) - f_{\tilde{k}}(q_j)|}_{\leq \frac{\varepsilon}{3}} + \underbrace{|f_{\tilde{k}}(q_j) - f_{\tilde{k}}(x)|}_{\leq L\delta = \frac{\varepsilon}{3}} \\ &\leq \varepsilon \end{aligned}$$

da  $f_k \in M_{L,l}$ . Da diese Aussage für alle  $x \in [a, b]$  gilt, bilden die  $(f_k)_{k=1}^\infty$  eine Cauchyfolge bzgl.  $\|\cdot\|_\infty$  und haben deshalb einen stetigen Grenzwert  $f_* \in C[a, b]$ .

$$\begin{aligned} f_* \in M_{L,l} \text{ folgt, da } |f_*(x) - f_*(y)| &= \left| \lim_{k \rightarrow \infty} f_k(x) - \lim_{l \rightarrow \infty} f_l(y) \right| \leq \lim_{n \rightarrow \infty} |f_n(x) - f_n(y)| \\ &\leq \lim_{n \rightarrow \infty} L|x - y| = L|x - y| \end{aligned}$$

□

**Bemerkung** Anwendung - Approximation nach Tschebyscheff = Chebyshev

**Gegeben:** Beliebige beschränkte Funktion  $f : [0, 1] \implies [0, l]$

**Gesucht:** Lipschitzstetige Annäherung  $g \in M_{L,l} \subset C[0, 1]$  für die der Abstand  $\varphi(g) = \|f - g\|_\infty$  minimiert.

**Zielfunktion**  $\varphi(g)$  ist auf  $C[a, b]$  Lipschitzstetig, da

$$|\varphi(g) - \varphi(\tilde{g})| = \left| \|f - g\|_\infty - \|f - \tilde{g}\|_\infty \right| \leq \tilde{L} \|g - \tilde{g}\|_\infty \text{ mit } \tilde{L} = 1$$

Nach Satz 28.15 erreicht  $\varphi$  auf  $M_{L,l}$  ein Minimum  $\varphi(g_*) = \min_{g \in M_{L,l}} \|f - g\|_\infty$ .

**Bemerkung** Die Voraussetzung von Lipschitzstetigkeit in Satz 28.16 kann abgeschwächt werden zu gleichgradiger Stetigkeit, d.h.  $M \subset C[a, b]$  muss so eingeschränkt sein, dass

$$\forall \varepsilon > 0 \exists \delta > 0 \forall f \in M; x, y \in [a, b] : |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon$$

M.a.W. Uniformität in  $x, y \in [a, b]$  und  $f \in M$ . Gilt mit  $\delta = \frac{\varepsilon}{L}$  im obigen Fall.

**Definition 28.18**

Eine Teilmenge  $M \subset X$  eines vollständigen metrischen Raumes heißt zusammenhängend, wenn es für alle  $x, y \in M$  eine stetige Funktion  $p : [0, 1] \rightarrow M$  mit  $p(0) = x$  und  $p(1) = y$  gibt.  $M$  heißt konvex, falls  $X$  Banachraum und  $p$  affin gewählt werden kann, d.h.

$$p(t) = (1 - t)x + ty = x + t(y - x) \in M$$

**Lemma 28.19**

Falls  $M \subset X$  zusammenhängend und  $f : X \rightarrow Y$  stetig dann ist auch  $f(M) \in Y$  zusammenhängend. Falls  $M \subset X$  konvexe Teilmenge eines Banachraumes und  $f : X \rightarrow Y$  affin, dann ist auch  $f(M) \in Y$  konvex.

**Beweis**

Für beliebiges  $\tilde{x}, \tilde{y} \in f(M)$  existieren Urbilder  $x, y \in M$  mit  $f(x) = \tilde{x}$  und  $f(y) = \tilde{y}$ . Jeder Pfad  $p(t)$  mit  $p(0) = x$  und  $p(1) = y$  ergibt ebenfalls stetigen Pfad  $f \circ p(t) = f(p(t)) \in f(M)$  mit  $(f \circ p)(0) = \tilde{x}$  und  $(f \circ p)(1) = \tilde{y}$ . Falls  $p$  und  $f$  linear sind, gilt dies auch für  $f \circ p$ , also vererbt sich Konvexität entsprechend.

**Bemerkung (Erinnerung)**  $f : X \rightarrow Y$  zwischen normierten Räumen:

**Linearität** Additivität + Homogenität  $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$

**Affinität**  $f(x) - f(0)$  ist linear  $\equiv$  Lineare Funktion + Konstante

## VIII Differentialrechnung in mehreren Variablen

### §29 Differenzierbarkeit

**Bemerkung** Die Idee eine glatte Funktion in der Nähe eines Punktes  $x$  durch eine affine „Tangentenfunktion“ anzunähern und in diesem Sinne zu „differenzieren“ überträgt sich direkt auf Abbildungen zwischen Banachräumen. Nach allgemeinen Einführen der Differenzierbarkeit konkrete Aussage in  $\mathbb{R}^n$ .

#### Definition 29.1

Eine Funktion  $f : X \rightarrow Y$  mit  $X, Y$  Banachräumen heißt am Punkt  $x \in X$  diffbar, wenn es einen linearen Operator  $A \in L(X, Y)$  gibt, sodass

$$\lim_{\|v\| \rightarrow 0} \frac{f(x+v) - f(x) - Av}{\|v\|} = 0 \quad \text{d.h. } \forall \varepsilon > 0 \exists \delta > 0 \implies \|f(x+v) - \underbrace{f(x) - Av}_{\text{affine in } v}\| \leq \varepsilon \|v\|$$

Dann nennt man  $F'(x) := A$  die Fréchet oder totale Ableitung von  $F$  an der Stelle  $x$ . Wenn  $F$  an allen  $x$  in offener Menge  $U \subset X$  diffbar ist und  $F'(x)$  als Funktion von  $x \rightarrow F'(x) \in L(X, Y)$  stetig ist, heißt  $F$  auf  $U$  Fréchet diffbar und man schreibt  $F \in C^1(U; Y)$ .

**Bemerkung** Da  $v \in X$  jetzt vektorwertig ist, können wir nicht durch  $v$  dividieren und  $\lim_{\|v\| \rightarrow 0} \frac{f(x+v) - f(x)}{v}$  macht keinen Sinn. Stattdessen schreibt man häufig  $f(x+v) - f(x) = Av + o(\|v\|)$  wobei  $o(\eta)$  wie folgt definiert ist.

#### Definition 29.2 (Landau-Symbolik)

Für  $p \geq 0$  und  $g, h : (-\delta, \delta) \rightarrow Y$  bedeuten die „Gleichungen“

$$g(\eta) = \mathcal{O}(\eta^p) \quad \text{und} \quad h(\eta) = o(\eta^p)$$

dass

$$\limsup_{\eta \rightarrow 0} \frac{\|g(\eta)\|}{\eta^p} < \infty \quad \text{bzw.} \quad \lim_{\eta \rightarrow 0} \frac{\|h(\eta)\|}{\eta^p} = 0$$

m.a.W.  $\frac{\|g(\eta)\|}{\eta^p}$  ist beschränkt im ersten Fall und  $\frac{\|h(\eta)\|}{\eta^p}$  geht gegen null wenn  $\eta \rightarrow 0$ .

**Beispiel**  $1007\eta^2 \sin(\frac{1}{\eta}) = \mathcal{O}(\eta^2)$  da

$$\overline{\lim}_{\eta \rightarrow 0} \frac{|\eta^2 \sin(\frac{1}{\eta})|}{\eta^2} = \overline{\lim}_{\eta \rightarrow 0} |\sin \frac{1}{\eta}| = 1$$

oder  $\frac{\sin(\eta^3)}{|\ln \eta|} = o(\eta^3)$  da

$$\lim_{\eta \rightarrow 0} \frac{\sin \eta^3}{\eta^3 |\ln \eta|} = \lim_{\eta \rightarrow 0} \frac{\sin \eta^3}{\eta^3} \cdot \lim_{\eta \rightarrow 0} \frac{1}{|\ln \eta|} = 1 \cdot 0 = 0$$

#### Lemma 29.3 (Rechenregeln für Landausymbole)

$$\mathcal{O}(\eta^p) \pm o(\eta^p) = \mathcal{O}(\eta^p)$$

$$\mathcal{O}(\eta^p) + \mathcal{O}(\eta^q) = \mathcal{O}(\eta^{\min(p,q)})$$

$$o(\eta^p) + o(\eta^q) = o(\eta^{\min(p,q)})$$

$$\mathcal{O}(\eta^p) \circ (\eta^q) = o(\eta^{p+q})$$

$$\frac{\mathcal{O}(\eta^p)}{\mathcal{O}(\eta^q)} \stackrel{?}{=} \mathcal{O}(\eta^{p-q})$$

NEIN

Linke Seite ist nicht definiert, da Nenner null sein kann. z.B.  $1 = \mathcal{O}(1)$ ,  $\sin(\frac{1}{\eta}) = \mathcal{O}(1) \implies \frac{1}{\sin(\frac{1}{\eta})}$  unbeschränkt durch beliebiges  $\eta^q$ .

**Bemerkung**  $\mathcal{O}(\eta^0) = \mathcal{O}(1)$  und  $o(\eta^0) = o(1)$  bezeichnen Größen die für  $\eta \rightarrow 0$  beschränkt sind bzw. gegen null gehen. Schreibweise  $\eta^0$  statt 1 ist eindeutig, wenn Ausdrücke verschiedene Variablen enthalten.

**Lemma 29.4**

$f : X \rightarrow Y$  ist an Stelle  $x$

stetig, falls  $f(x + v) = f(x) + o(\|v\|^0)$

Lipschitzstetig, falls  $f(x + v) = f(x) + \mathcal{O}(\|v\|)$

Fréchetdiffbar, falls  $f(x + v) = f(x) + Av + o(\|v\|)$

Daraus folgt, dass Diffbar  $\implies$  Lipschitzstetig  $\implies$  Stetigkeit.

**Satz 29.5 (Kettenregel)**

Sei  $f : X \rightarrow Y$  diffbar an Stelle  $x$  und  $g : Y \rightarrow Z$  diffbar an Stelle  $y = f(x)$ , dann ist  $h = g \circ f$  an der Stelle  $x$  diffbar mit

$$h'(x) = g'(f(x))f'(x) = g'(y)f'(x) \in L(X, Z) \text{ wobei } g'(y) \in L(Y, Z) \text{ und } f'(x) \in L(X, Y)$$

**Beweis**

$$\begin{aligned} h(x + v) - h(x) &= g(f(x + v)) - g(f(x)) = g(y + \underbrace{f(x + v) - f(x)}_{\equiv w}) - g(y) && \text{wegen Diffbarkeit von } g \\ &= g'(y) \cdot w + o(\|w\|) = g'(y) \left[ \underbrace{f(x + v) - f(x)}_{f'(x)v + o(\|v\|)} \right] + o(\| \underbrace{f(x + v) - f(x)}_{=O(\|v\|) \text{ wegen Lemma 29.4}} \|) \\ &= g'(y)f'(x)v + g'(y)o(\|v\|) + o(\mathcal{O}(\|v\|)) \\ &= g'(y)f'(x)v + o(\|v\|) \end{aligned}$$

Daraus folgt die Behauptung.

**Satz 29.6 (Differenzierungsregeln für  $X, Y, Z$  Banachräume)**

$f : X \rightarrow Y$  und  $g : X \rightarrow Z$  beide diffbar an  $x \in X$ , also sind auch diffbar in  $x$ :

- (i)  $h(x) = (f(x), g(x)) : X \rightarrow Y \times Z$  mit  $h'(x) = (f'(x), g'(x))$ , d.h.  $h'(x)v = (f'(x)v, g'(x)v)$
- (ii)  $h(x) = \alpha f(x) + \beta g(x)$  für  $Y = Z$  und  $\alpha, \beta \in \mathbb{R}$  mit  $h'(x) = \alpha f'(x) + \beta g'(x)$  (Linearität)
- (iii)  $h(x) = f(x) \cdot g(x) : X \rightarrow Y$  für  $Z = \mathbb{R}$  mit  $h'(x) = f'(x)g(x) + f(x)g'(x) \in L(X, Y)$
- (iv)  $h(x) = \frac{f(x)}{g(x)}$  für  $Z = \mathbb{R}$  mit  $g(x) \neq 0$  mit  $h'(x) = \frac{f'(x)g(x) - g'(x)f(x)}{g(x)^2}$

**Beweis**

- (i) 
$$\begin{aligned} h(x + v) - h(x) &= (f(x + v), g(x + v)) - (f(x), g(x)) = (f(x + v) - f(x), g(x + v) - g(x)) \\ &= (f'(x)v + o(\|v\|), g'(x)v + o(\|v\|)) = (f'(x)v, g'(x)v) + (o(\|v\|), o(\|v\|)) \\ &= o(\max\{\|v\|, \|v\|\}) \quad \text{in der Produkttopologie} \\ &= o(\|v\|) \end{aligned}$$
- (ii) *Hausaufgabe*
- (iii) 
$$\begin{aligned} h(x + v) - h(x) &= f(x + v)g(x + v) - f(x)g(x) \\ &= (f(x + v) - f(x))g(x + v) + f(x)(g(x + v) - g(x)) \\ &= (f'(x)v + o(\|v\|))g(x + v) + f(x)(g'(x)v + o(\|v\|)) \\ &= g(x)f'(x)v + o(\|v\|)g(x) + f(x)v\mathcal{O}(\|v\|) + o(\|v\|)\mathcal{O}(\|v\|) + o(\|v\|) \\ &= [g(x)f'(x)v + f(x)g'(x)v] + o(\|v\|) \end{aligned}$$
- (iv) *Übung*

**Bemerkung** (i) Aussage gilt auch umgekehrt, d.h.  $F = \begin{pmatrix} f \\ g \end{pmatrix} : X \rightarrow Y \times Z$  ist genau dann diffbar an  $x \in X$ , wenn sowohl  $f$  wie  $g$  an der Stelle  $x$  diffbar sind und es gilt  $F'(x) = \begin{pmatrix} f'(x) \\ g'(x) \end{pmatrix}$  entsprechend ist eine Vektorfunktion

$$F = \begin{pmatrix} F_1(x) \\ \vdots \\ F_m(x) \end{pmatrix} : X = \mathbb{R}^n \rightarrow Y = \mathbb{R}^m$$

genau dann an  $x \in \mathbb{R}^n$  diffbar, wenn dies für die Komponentenfunktionen  $F_i : \mathbb{R}^n \rightarrow \mathbb{R}$  gilt. Fürs erste betrachten wir deshalb „nur“ skalarwertige Funktionen  $f(x) = f(x_1, \dots, x_n) : \mathbb{R}^n \rightarrow \mathbb{R}$

**Bemerkung (Herleitung des Gradienten)** Die Abbildung  $f'(x) \in L(\mathbb{R}^n, \mathbb{R})$  muss erfüllen für  $v = e_j \cdot \nu(0, \dots, \nu, 0, \dots, 0)$

$$\begin{aligned} f(x+v) - f(x) &= f(x_1, \dots, x_{j-1}, x_j + \nu, x_{j+1}, \dots, x_n) - f(x_1, \dots, x_n) \\ &= f'(x)v + o(\|v\|) = f'(x)e_j \cdot \nu + o(|\nu|) \text{ da } \|v\| = |\nu| \\ f'(x)e_j &= \lim_{\nu \rightarrow 0} \frac{f(x + \nu e_j) - f(x)}{\nu} = \left. \frac{d}{d\nu} f(x + \nu e_j) \right|_{\nu=0} \end{aligned}$$

Partielle Ableitung nach  $x_j$  mit  $x_i$  konstant für  $i \neq j$

$$f'(x)e_j = \frac{\partial}{\partial x_j} f(x) = \partial_j f(x) = \left. \frac{d}{d\nu} f(x + \nu e_j) \right|_{\nu=0}$$

**Beispiel**

$$\begin{aligned} f(x_1, x_2) &= \sin(x_1 x_2) \exp(x_1 - x_2) \\ \frac{\partial f}{\partial x_1} &= \partial_1 f = \cos(x_1 x_2) x_2 \exp(x_1 - x_2) + \sin(x_1 x_2) \exp(x_1 - x_2) \\ \frac{\partial f}{\partial x_2} &= \partial_2 f = \cos(x_1 x_2) x_1 \exp(x_1 - x_2) - \sin(x_1 x_2) \exp(x_1 - x_2) \end{aligned}$$

**Bemerkung** Notwendige Bedingung für totale oder Fréchetdiffbarkeit von  $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$  ist partielle Diffbarkeit bezüglich aller  $n$  Variablen  $x_1, \dots, x_n$ .

**Bemerkung (Notation)** totale oder F-diffbar  $f(x+v) = f(x) + f'(x)v + o(\|v\|)$  wobei  $f'(x)v$  bedeutet Anwendung des linearen Operators, d.h. lineare Abbildung  $f'(x) \in L(X, Y)$  auf  $v = \Delta x \in X$ . Alternativ  $(f'(x))(v)$ . Bei Kettenregel angewandt auf  $h = g \circ f$  tritt auf  $h'(x) = g'(y)f'(x)$  mit  $y = f(x)$ . Hierbei bedeutet „Multiplikation“ der Operatoren  $f'(x) \in L(X, Y)$  und  $g'(y) \in L(Y, Z)$  einfach Hintereinanderausführung, d.h.

$$(g'(x)f'(x))v = g'(y)(f'(x))v = g'(y)w \quad \text{mit } w = f'(x)v$$

*Wichtig:* Multiplikation von linearen Operatoren ist assoziativ, aber nicht kommutativ. Addition ist beides, vorausgesetzt, Definitions- und Bildbereiche  $X, Y$  stimmen überein. Im endlichdimensionalen, d.h.  $X = \mathbb{R}^n, Y = \mathbb{R}^m$  und  $Z = \mathbb{R}^p$  kann man  $f'(x), g'(x), h'(x)$  bezüglich der kanonischen Basis  $(e_j)_{j=1, \dots, n}$  mit Matrizen  $\nabla f \in \mathbb{R}^{m \times n}, \nabla g \in \mathbb{R}^{p \times m}$  und  $\nabla h \in \mathbb{R}^{p \times n}$  identifizieren und es gilt  $\nabla h(x) = \nabla g(y)\nabla f(x)$  im Sinne der Matrizenmultiplikation.

**Bemerkung (Gradienten(transponierte Gradient) bzw. Nabla Operator)** Aus einer skalarwertigen Funktion  $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$  ergibt sich durch partielle Differentiation die vektorwertige Funktion

$$\nabla f(x) = (\partial_1 f(x), \partial_2 f(x), \dots, \partial_n f(x)) \in \mathbb{R}^{1 \times n}$$

vorausgesetzt die partiellen Ableitungen

$$\partial_j f(x) = \left. \frac{d}{d\nu} f(x + \nu e_j) \right|_{\nu=0} \quad \text{für } j = 1 \dots n$$

existieren. Nach Hölderungleichung folgt für  $v = \sum_{j=1}^n v_j e_j \in \mathbb{R}^n$ , dass

$$|f'(x)v| = \left| \sum_{j=1}^n f'(x)e_j v_j \right| = \left| \sum_{j=1}^n \partial_j f(x) v_j \right| \leq \left( \sum_{j=1}^n |\partial_j f(x)|^p \right)^{\frac{1}{p}} \left( \sum_{j=1}^n |v_j|^q \right)^{\frac{1}{q}} = \|\nabla f(x)\|_p \|v\|_q$$

mit  $\frac{1}{p} + \frac{1}{q} = 1$  Falls nicht anders definiert, benutzen wir immer  $p = q = 2$ .

**Beispiel** für das partielle Differenzierbarkeit an der Stelle  $x = 0 \in \mathbb{R}^2$ , aber die Funktion dort nicht F-diffbar ist:

$$f(x_1, x_2) = \min(|x_1|, |x_2|)$$

$$\begin{aligned} \partial_1 f(0, 0) &= \left. \frac{d}{d\nu} f(\nu, 0) \right|_{\nu=0} = \left. \frac{d}{d\nu} \min(|\nu|, 0) \right|_{\nu=0} = 0 = \partial_2 f(0, 0) \\ \implies f(v_1, v_2) - f(0, 0) - \nabla f(0) \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} &= \min(|v_1|, |v_2|) - 0 - 0 \neq o(\|v\|) \end{aligned}$$

da z.B. längs Diagonale  $v_1 = v_2 = \nu$

$$\frac{\min(|v_1|, |v_2|)}{\|v\|} = \frac{|\nu|}{\sqrt{2}|\nu|} = \frac{1}{\sqrt{2}} \text{ konstant}$$

Problem ist, dass die partiellen Ableitungen nur an Stelle  $x = 0$  existieren, aber nicht in Umgebung, z.B. Punkt  $x = (1, 1)$

### Satz 29.7

Falls  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  an allen Punkten  $x \in U$ ,  $U$  offen, partiell diffbar ist und die partielle Ableitung  $\partial_j f : U \rightarrow \mathbb{R}$  für  $j = 1, \dots, n$  auf  $U$  stetig sind, dann ist  $f$  selbst auf  $U$  Fréchet-diffbar, d.h.  $f \in C^1(U, \mathbb{R})$

#### Beweis

Betrachte festes  $x \in U$  mit  $B_r(x) \subset U$ . Dann folgt

$$\|v\|_2 < r \implies x + v \in B_r(x) \subset U$$

Es gilt sogar dass für  $i = 0, \dots, n$ :

$$x^{(i)} = (x_1 + v_1, \dots, x_i + v_i, x_{i+1}, \dots, x_n) \in B_r(x)$$

$$f(x + v) - f(x) = \sum_{i=1}^n f(x^{(i)}) - f(x^{(i-1)}) = f^{(n)} - f^{(n-1)} + f^{(n-1)} - f^{(n-2)} + \dots + f^{(1)} - f^{(0)}$$

Da alle  $f(x + ve_j)$  in  $B_r(x)$  nach  $v$  stetig diffbar sind, existieren nach Mittelwertsatz der Diffrechnung Werte  $\delta_i \in (0, 1)$  sodass

$$\begin{aligned} f(x^{(i)}) - f(x^{(i-1)}) &= \partial_i f(x^{(i-1)} + \delta_i v_i e_i) v_i \\ \implies f(x + v) - f(x) - \underbrace{\sum_{i=1}^n \partial_i f(x) v_i}_{=f'(x)v} \\ &= \sum_{i=1}^n [\partial_i f(x^{(i-1)} + \delta_i v_i e_i) - \partial_i f(x)] v_i \stackrel{?}{=} o(\|v\|) \end{aligned}$$

Wegen vorausgesetzter Stetigkeit der  $\partial_i f(x)$  existieren für  $\varepsilon > 0$  beliebig ein  $\delta$  sodass  $\|v\|_2 < \delta \implies |\partial_i f(x + v) - \partial_i f(x)| \leq \frac{\varepsilon}{n}$ . Da

$$\|x^{(i-1)} + \delta_i v_i e_i - x\|_2 = \left\| \sum_{j=1}^i \delta_j v_j e_j \right\|_2 \leq \|v\|_2$$

ergibt sich

$$|f(x+v) - f(x) - \nabla f(x)v| \leq \sum_{i=1}^n |\partial_i f(x^{(i-1)} + v_i \delta_i e_i) - \partial_i f(x)| |v_i| \leq \max_{1 \leq i \leq n} \{ |v_i| \} \cdot n \cdot \frac{\varepsilon}{n} \leq \|v\|_2 \cdot \varepsilon$$

### Beispiel

$$f(x_1, x_2) = \log(x_1 + x_2^2) \cdot \sin(x_1 x_2) : U \rightarrow \mathbb{R} \text{ mit } U := \{(x_1, x_2) \in \mathbb{R}^2 : x_1 > -x_2^2\}$$

$$\nabla f = (\partial_i f(x_1, x_2))_{i=1,2} = \left( \frac{1}{x_1 + x_2^2} \sin(x_1 x_2) + \log(x_1 + x_2^2) \cos(x_1 x_2) x_2, \frac{2x_2}{x_1 + x_2^2} \sin(x_1 x_2) + \log(x_1 + x_2^2) \cos(x_1 x_2) x_1 \right)$$

An allen  $x \in U$  ist  $\nabla f = (\partial_1 f, \partial_2 f)$  definiert, also ist  $f$  stetig diffbar

### Satz 29.8 (Mittelwertsatz für $f : \mathbb{R}^n \rightarrow \mathbb{R}$ )

Sei  $f \in C^1(U, \mathbb{R})$  mit  $U$  offen und  $X(t) = (1-t)x + ty \in U$  für  $t \in [0, 1]$  für gegebene  $x, y \in U$ . Dann existiert ein Wert  $t \in [0, 1]$ , sodass für  $z = p(t)$  gilt

$$f(y) - f(x) = \nabla f(z)(y - x)$$

### Beweis

Sowohl  $f : U \rightarrow \mathbb{R}$  wie  $p : (0, 1) \rightarrow U$  sind  $F$ -diffbar mit  $p'(t) = \frac{d}{dt} p(t) = (y - x)$ . Die Kettenregel ergibt dann für  $\varphi(t) = f(p(t))$  die Ableitung

$$\varphi'(t) = \nabla f(p(t)) \cdot p'(t) = \nabla f(p(t))(y - x)$$

Da  $\varphi(0) = f(x)$  und  $\varphi(1) = f(y)$  existiert nach Mittelwertsatz für reelle Funktionen eine Variable, sodass

$$f(y) - f(x) = \varphi(1) - \varphi(0) = \varphi'(t) = \nabla f(p(t))(y - x)$$

□

### Korollar 29.9 (Schrankensatz für $f : \mathbb{R}^n \rightarrow \mathbb{R}$ )

Unter Voraussetzung von Satz 29.8 gilt

$$|f(y) - f(x)| \leq L \|y - x\|_2 \text{ wobei } L = \max_{0 \leq t \leq 1} \|\nabla f((1-t)x + ty)\|_2$$

### Beweis

Nach 29.8 und Cauchy-Schwarz gilt

$$|f(y) - f(x)| = |\nabla f(z)(y - x)| \leq \|\nabla f(z)\| \cdot \|y - x\| \leq L \|y - x\|$$

da  $z = (1-t)x + ty$  für ein  $t \in [0, 1]$ .

### Satz 29.10

Sei  $G \subset \mathbb{R}^n$  offen und wegzusammenhängend und  $f : G \rightarrow \mathbb{R}$  stetig diffbar. Dann ist  $f \equiv c$  konstant gdw.  $\nabla f \equiv 0$ .

### Beweis

$\Rightarrow$  trivial

$\Leftarrow$  **konvexer Fall** wenn  $x \in G$ , dann gibt es ein  $\varepsilon > 0$  mit  $B(x, \varepsilon) \subset G$ . Für jedes  $y \in B(x, \varepsilon)$  gilt  $[x, y] \subset B(x, \varepsilon)$ .

$$\text{MWS: } f(y) - f(x) = \nabla f(z)(y - x) \text{ für ein } z \in [x, y]$$

Also ist  $f(y) - f(x) = 0$ , weil  $\nabla f(z) = 0$ .

**sonst** Seien  $x, y \in G$  beliebig. Dann gibt es ein  $\gamma : [0, 1] \rightarrow G$  stetig mit  $\gamma(0) = x, \gamma(1) = y$ .  
Für jedes  $t$  gibt es ein  $\varepsilon(t)$  mit  $B(\gamma(t), \varepsilon(t)) \subset G$ .

$$\bigcup_{t \in [0, 1]} B(\gamma(t), \varepsilon(t))$$

ist eine offene Überdeckung. Für jedes  $t \in [0, 1]$  gibt es ein  $\delta(t)$ , sodass

$$\gamma(B(t, \delta)) \subset B(\gamma(t), \frac{\varepsilon(t)}{2})$$

Da  $[0, 1]$  kompakt ist, gibt es

$$t_0 = 0 < t_1 < \dots < t_N = 1 \quad \delta_k = \delta(t_k), \varepsilon_k = \varepsilon(t_k) \text{ ,sodass } \bigcup B(t_k, \frac{\delta_k}{2}) \supset [0, 1]$$

$$t_k \in B(t_{k+1}, \delta_{k+1}) \text{ oder } t_{k+1} \in B(t_k, \delta_k) \\ \gamma(t_k) \in B(\gamma(t_{k+1}), \frac{\varepsilon_{k+1}}{2}) \text{ oder } \gamma(t_{k+1}) \in B(\gamma(t_k), \frac{\varepsilon_k}{2})$$

In beiden Fällen gilt  $f(\gamma(t_k)) = f(\gamma(t_{k+1}))$  da  $B(\gamma(t_i), \frac{\varepsilon_i}{2})$  konvex und  $B(\gamma(t_i), \frac{\varepsilon_i}{2}) \subseteq U$ , also MWS.

$$\implies f(x) = f(\gamma(t_0)) = f(\gamma(t_1)) = \dots = f(\gamma(t_N)) = f(y)$$

Da  $x, y \in G$  beliebig, gilt  $f(y) = f(x)$  für jedes  $y \in G$ , d.h.  $c = f(x)$

□

### Definition 29.11 (Vektorwertige Funktionen)

$G \subset \mathbb{R}^n$  sei ein Gebiet,  $f : G \rightarrow \mathbb{R}^m$ .  $f$  hat  $m$  Komponenten  $f_1, \dots, f_m : G \rightarrow \mathbb{R}$  mit

$$f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}$$

Sind alle Komponenten in  $x \in G$  partiell diffbar, dann kann man die Jacobi-Matrix definieren als

$$\nabla f(x) = \begin{pmatrix} - & \nabla f_1(x) & - \\ & \vdots & \\ - & \nabla f_m(x) & - \end{pmatrix} = \begin{pmatrix} | & & | \\ \partial_1 f(x) & \dots & \partial_n f(x) \\ | & & | \end{pmatrix} = \begin{pmatrix} \partial_1 f_1(x) & \dots & \partial_n f_1(x) \\ \vdots & & \vdots \\ \partial_1 f_m(x) & \dots & \partial_n f_m(x) \end{pmatrix}$$

$\partial_k f_l(x) = \frac{\partial f_l}{\partial x_k}(x)$  ist die partielle Ableitung von  $f_l$  nach  $x_k$  im Punkt  $x$ .

### Satz 29.12

$f$  ist auf  $G$  stetig diffbar, gdw. die partiellen Ableitungen der Komponenten existieren und stetig sind.

#### Beweis

Folgerung aus dem skalaren Fall.

**Beispiel (zum Mittelwertsatz)**  $f : [0, 1] \rightarrow \mathbb{R}^2, f(x) = \begin{pmatrix} x^2 \\ x^3 \end{pmatrix}$  Es ist  $\nabla f(x) = \begin{pmatrix} 2x \\ 3x^2 \end{pmatrix}$ .

$$f_1(1) - f_1(0) = 1 = 2z \cdot (1 - 0) \iff z = \frac{1}{2} \\ f_2(1) - f_2(0) = 1 = 3z^2 \cdot (1 - 0) \iff z = \sqrt{\frac{1}{3}}$$

Können keinen gemeinsamen Zwischenpunkt  $z \in [0, 1]$  mit  $f(1) - f(0) = \nabla f(1 - 0)$  finden.

**Satz 29.13 (Schranksatz)**

$G \subset \mathbb{R}^n$  offen,  $f : G \rightarrow \mathbb{R}^m$  stetig diffbar.  $x, y \in G$  mit  $[x, y] \subset G$ . Dann gilt

$$\|f(y) - f(x)\| \leq L\|y - x\| \quad \text{mit } L = \sup_{z \in [x, y]} \|\nabla f(z)\|$$

**Bemerkung** Wenn  $\|\cdot\|$  die euklidische Norm ist, dann ist  $\|A\| = \sqrt{\lambda_{\max}(A^T A)}$  mit  $\lambda_{\max}$  ist größter Eigenwert. Jede Matrix  $A$  lässt sich zerlegen in

$$A = U\Sigma V \quad \text{mit } U, V \text{ orthogonal}$$

Und  $\Sigma = \begin{pmatrix} \% \\ 0 \end{pmatrix}$  oder  $\Sigma = (\% \quad 0)$ .

**Beweis**

Betrachten Hilfsfunktion  $\varphi : G \rightarrow \mathbb{R}$ ,  $\varphi(x) = a^T f(x)$  für ein beliebiges  $a \in \mathbb{R}^m, a \neq 0$ . Nach skalarem MWS gibt es ein  $z \in [x, y]$  mit

$$\varphi(y) - \varphi(x) = \nabla \varphi(z) \cdot (y - x) = a^T \cdot \nabla f(z) \cdot (y - x)$$

Nach Cauchy-Schwarz gilt

$$\begin{aligned} |\varphi(y) - \varphi(x)| &\leq \|a\| \cdot \|\nabla f(z)(y - x)\| \leq \|a\| \cdot \|\nabla f(z)\| \cdot \|y - x\| \\ \frac{|a^T(f(y) - f(x))|}{\|a\|} &\leq \|\nabla f(z)\| \cdot \|y - x\| \\ &\leq L\|y - x\| \end{aligned}$$

$z$  hängt von  $a$  ab,  $L\|x - y\|$  ist von  $a$  unabhängig. Wenn  $f(x) = f(y)$ , dann ist nichts zu zeigen. Sonst setze  $a = f(y) - f(x)$ :

$$\frac{|a^T(f(y) - f(x))|}{\|a\|} = \frac{\|a\|^2}{\|a\|} = \|f(y) - f(x)\| \leq L\|y - x\|$$

□

**Korollar 29.14**

Jede stetig diffbare Funktion  $f : K \rightarrow \mathbb{R}^m$  auf einer konvexen, kompakten Menge  $K$  ist Lipschitzstetig.

$$L = \sup_{z \in K} \|\nabla f(z)\|$$

**Bemerkung (Newton-Verfahren)**

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

$f : \mathbb{R}^n \rightarrow \mathbb{R}$  stetig diffbar. Linearisiere  $f$  in  $x_k$ , suche Nullstelle der Linearisierung als  $x_{k+1}$ .

$$f(x + v) = f(x) + \nabla f(x) \cdot v + o(v) \quad \text{für „kleine“ } v$$

$f(x) + \nabla f(x)v$  ist die Linearisierung von  $f$  in  $x$ . Löse

$$0 = f(x) + \nabla f(x)v \quad \text{nach } v$$

$$v = -(\nabla f(x))^{-1} f(x)$$

Wenn  $f(x)$  schon „klein genug“, dann ist  $f(x + v) = o(f(x))$ . Oft auch  $f(x + v) = \mathcal{O}(\|f(x)\|^2)$ .

**Satz 29.15 (von Kantorowitsch, vereinfachte Version)**

$G \subset \mathbb{R}^n, f : G \rightarrow \mathbb{R}^n$ .  $D_0 \subset G$  sei offen, konvex,  $f$  auf  $D_0$  diffbar und  $\nabla f : D_0 \rightarrow \mathbb{R}^{n \times n}$  sei Lipschitzstetig mit Konstante  $L > 0$ . In  $x_0 \in D_0$  sei  $\nabla f(x_0)$  invertierbar mit

$$\gamma = L \cdot \|(\nabla f(x_0))^{-1}\|^2 \cdot \|f(x_0)\| \leq \frac{1}{2}$$

und

$$B(x_0, 2\|(\nabla f(x_0))^{-1} f(x_0)\|) \subset D_0$$

$$\implies x_{k+1} := x_k - \nabla f(x_k)^{-1} f(x_k) \text{ konvergiert gegen eine Nullstelle } x^* \in \overline{B}(x_0, 2\|(\nabla f(x_0))^{-1} f(x_0)\|)$$

## §30 Umkehrfunktion, Gleichungen, IFT

IFT = Implizite Funktionen Theroem

**Bemerkung (Motivation/Zusammenhang)** Als Verallgemeinerung von linearen Gleichungssystemen

$$Ax = y \text{ mit } y \in \mathbb{R}^n \text{ gegeben, } x \in \mathbb{R}^n \text{ gesucht}$$

betrachte nichtlineare Gleichungen

$$F(x) = y \text{ mit } F : \mathbb{R}^n \rightarrow \mathbb{R}^n$$

Aufgabe: Bestimmung von  $x$  für gegebenes  $y$ , d.h. Auswertung der (bzw. einer) Umkehrfunktion  $x = F^{-1}(y)$

**Beispiel (Polarkoordinaten)**

$$F(r, \varphi, \Theta) = \begin{pmatrix} r \sin(\Theta) \cos(\varphi) \\ r \sin(\Theta) \sin(\varphi) \\ r \cos(\Theta) \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

Hier bekommt man nach sukzessive auflösen

$$r = \sqrt{y_1^2 + y_2^2 + y_3^2} \quad \Theta = \arccos\left(\frac{y_3}{r}\right) \quad \varphi = \frac{x_2}{r \sin(\Theta)} \text{ falls } \Theta \neq 0$$

Im allgemeinen können Systeme nichtlinearer Gleichungen nicht in geschlossener Form, d.h. Formeln bzw. Ausdrücke gelöst werden. Dann ergeben sich zwei Aufgaben:

- (i) Beweise *Existenz* und gegebenenfalls *Eindeutigkeit* von Lösungen in bestimmten Umgebungen  $U \subset \mathbb{R}^n$  und für bestimmte  $y \in V \subset \mathbb{R}^n$ .
- (ii) Numerische *Approximation* von Lösungen mittels *iterativer* Verfahren( z.B. Newton und Varianten)

Umkehrbarkeit verlangt im allgemeinen Differenzierbarkeit und Regularität der Jacobimatrix

$$\nabla F(x_0) = \left( \frac{\partial F_i(x_1 \dots x_n)}{\partial x_j} \right)_{\substack{i=1 \dots n \\ j=1 \dots n}} \in \mathbb{R}^{n \times n}$$

$$\begin{aligned} \det(\nabla F(x_0)) \neq 0 &\iff \nabla F(x_0)^{-1} \text{ existiert} \\ &\iff \text{span}\{\partial_j F(x_0)\}_{j=1}^n = \mathbb{R}^n \\ &\iff \text{span}\{\nabla F_i(x_0)\}_{i=1}^n = \mathbb{R}^n \\ &\iff F(x_0)v = 0 \iff v = 0 \end{aligned}$$

Häufig sind Gleichungen abhangig bzw. variabel in Paramtern  $y \in \mathbb{R}^m$  und man sucht Losungen von

$$f(x, y) = 0 \in \mathbb{R}^n \text{ mit } x \in \mathbb{R}^n$$

Falls fur alle  $y \in V \subset \mathbb{R}^m$  Losung  $x = g(y)$  existieren und dieser Zusammenhang lokal eindeutig ist, heit  $x = g(y)$  eine *implizite Funktion*, die auch durch  $f(x, y) = 0$  definiert ist.

**Beispiel**

$$f(x, y) = x^2 + y^2 - 1 = 0 \quad x = g(y) = \sqrt{1 - y^2} \text{ fur } y \in [-1, 1]$$

Ruckfuhrung auf Umkehrfunktion durch Einbettung

$$F(x, y) = \begin{pmatrix} f(x, y) \\ y \end{pmatrix} : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}^{n \times m} \quad \text{mit } \nabla F(x, y) = \begin{pmatrix} \nabla_x f & \nabla_y f \\ 0 & I \end{pmatrix} \implies \det(\nabla F(x, y)) = \det(\nabla_x f)$$

**Beispiel (Gleichungslösung mit Newtonvarianten)** Annahme:  $F \in C^1(U, \mathbb{R}^n)$ ,  $U \subset \mathbb{R}^n$  offen und konvex.  $x_0 \in U$  annähernd Nullstelle  $F(x_0) \approx 0$ , d.h.  $\|F(x_0)\| \ll 1$

$$\text{Linearisierung: } F(x_0) + \nabla F(x_0)v + o(\|v\|)$$

Bedingungen an  $v$ :

$$\begin{aligned} F(x_0) + \nabla F(x_0)v &= 0 \\ \implies F(x_0 + v) &= F(x_1) = o(\|v\|) = o(\|F(x_0)\|) \\ \text{da } v &= -\nabla F(x_0)^{-1}F(x_0) = o(\|F(x_0)\|) \end{aligned}$$

$V \in \mathbb{R}^n$  ist eindeutig definiert gdw.  $\nabla F(x_0)$  regulär ist, so dass *algebraisch* gilt  $v = -[\nabla F(x_0)]^{-1}F(x_0)$ . In numerischer Praxis sollte man (fast) nie  $[\nabla F(x_0)]^{-1}$  berechnen, sondern nur das lineare System  $\nabla F(x_0)v = -F(x_0)$  lösen, zum Beispiel via *LU-Faktorisierung* = Gaußscher Elimination. Erste iterierte:

$$x_1 = x_0 + v = x_0 - [\nabla F(x_0)]^{-1}F(x_0)$$

Wiederholte Anwendung ergibt „unendliche“ Iterationen

$$\begin{aligned} x_{k+1} &= x_k - [\nabla F(x_k)]^{-1}F(x_k) && \text{volles Newtonverfahren} \\ x_{k+1} &= x_k - [\nabla F(x_0)]^{-1}F(x_k) && \text{vereinfachter Newton, „hard method“} \end{aligned}$$

Der vereinfachte Newton ist etwas langsamer, dafür aber billiger pro Schritt. Das erhoffte Ergebnis ist Konvergenz, d.h.

$$\begin{aligned} x_k \rightarrow x_* &\implies v_k = x_{k+1} - x_k \rightarrow 0 \\ &\implies [\nabla F(x_k)]^{-1}F(x_k) \rightarrow 0 \\ &\implies F(x_k) = -F'(x_k)v_k \rightarrow 0 \\ &\implies F(x_*) = 0 \iff x_* \text{ Nullstelle} \end{aligned}$$

Frage: Wie kann man die Konvergenz zeigen?

Antwort: Mit Hilfe des Banachschen Fixpunktsatzes angewandt auf  $G(x) = x - \nabla F(x_0)^{-1}F(x)$ .

$$\begin{aligned} &\nabla G(x) = I - [\nabla F(x_0)]^{-1}\nabla F(x) \\ \implies &\nabla G(x_0) = 0 \text{ wegen Stetigkeit } \nabla G(x) \text{ folgt} \\ \implies &\|\nabla G(x)\| \leq \frac{1}{2} \text{ für } x \in B_r(x_0) \\ \implies &\|G(x) - G(z)\| \leq \frac{1}{2}\|x - z\| \text{ für } x, y \in B_r(x_0) \text{ nach Schrankensatz} \end{aligned}$$

$$x_1 = x_0 - [\nabla F(x_0)]^{-1}F(x_0) \|x_1 - x_0\| = \|[\nabla F(x_0)]^{-1}F(x_0)\| \leq \|[\nabla F(x_k)]^{-1}\| \underbrace{\|F(x_0)\|}_{\approx 0}$$

Aus Kontraktion folgt  $\|x_{k+1} - x_n\| \leq (\frac{1}{2})^k \|x_1 - x_0\|$  Summation über  $j \leq k$  ergibt

$$\begin{aligned} \|x_k - x_0\| &\leq (1 + \frac{1}{2} + \frac{1}{4} + \dots)\|x_1 - x_0\| \leq 2\|x_1 - x_0\| \\ \implies &\{x_k\}_{k \in \mathbb{N}} \subset B_r(x_0) \text{ falls } \|[\nabla F(x_0)]^{-1}F(x_0)\| \leq \frac{r}{2} \\ \implies &x_k \rightarrow x_* \in B_r(x_0) \text{ mit } F(x_*) = 0 \end{aligned}$$

### Satz 30.1

Sei  $F \in C^1(U, \mathbb{R}^n)$ ,  $x \in U$  offen und konvex. Definiere  $r > 0$ , so dass  $\|\nabla G(x)\| \leq \frac{1}{2}$  für  $x \in B_r(x_0)$ . Dann folgt aus der Bedingung

$$\|[\nabla F(x_0)]^{-1}F(x_0)\| < \frac{r}{2}$$

dass die vereinfachte Newtoniteration  $x_{k+1} = x_k - [\nabla F(x_0)]^{-1}F(x_k)$  innerhalb  $B_r(x_0)$  bleibt und Konvergenz die einzige in  $B_r(x_0)$  existierende Nullstelle  $x_* = \lim_{k \rightarrow \infty} x_k$  erzielt.

**Beweis**

Wegen vorausgegangener Herleitung bleibt nur noch Eindeutigkeit zu überprüfen. Aus  $F(x_*) = 0 = F(\tilde{x}_*)$  mit  $x_*, \tilde{x}_* \in B_r(x_0)$  folgt:

$$\begin{aligned} \|x_* - \tilde{x}_*\| &= \|G(x_*) - G(\tilde{x}_*)\| \leq \frac{1}{2} \|x_* - \tilde{x}_*\| \\ &\implies \|\tilde{x}_* - x_*\| = 0 \end{aligned}$$

□

**Bemerkung** Asymptotisch ergibt sich die Konvergenzgeschwindigkeit

$$\overline{\lim}_{k \rightarrow \infty} \frac{\|x_{k+1} - x_*\|_2}{\|x_k - x_*\|_2} \leq \|I - [\nabla F(x_0)]^{-1} \nabla F(x_*)\| = \|\nabla G(x_*)\| \leq \frac{1}{2}$$

Das nennt man Q-lineare Konvergenz. Für volles Newtonverfahren ergibt schon für  $F \in C^1(U)$  superlineare Konvergenz d.h.

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} = 0$$

Schlüsselvoraussetzung:  $\nabla F(x)$  ist regulär und stetig nahe Nullstelle  $x_*$ . Als Korollar erhalten wir lokale Umkehrbarkeit im folgenden Sinne:

**Satz 30.2**

Falls  $F \in C^1(U, \mathbb{R}^n)$  mit  $U$  offen und an Stelle  $x_0 \in U$  mit  $y_0 = F(x_0)$  gilt

$$\det(\nabla F(x_0)) \neq 0$$

Dann existiert eine Kugel  $V = B_\rho(y_0)$  und eine Umkehrfunktion  $F^{-1} : V \rightarrow U$ , so dass

$$F \circ F^{-1}(y) = y \text{ für alle } y \in B_\rho(y_0)$$

Zudem ist  $F^{-1} \in C^1(V, U)$  mit der totalen Ableitung

$$\underbrace{\nabla F^{-1}(y_0)}_{\text{Ableitung der Umkehrfunktion}} = \underbrace{[\nabla F(x_0)]^{-1}}_{\text{Inverse Matrix der Jacobiatrix}}$$

**Bemerkung**

$F \in C^1(U, \mathbb{R}^n), U \subset \mathbb{R}^n$  offen und zusammenhängend

$$\det(\nabla F(x_0)) \neq 0$$

$$\implies \|I - (\nabla F(x_0))^{-1} \nabla F(x)\| < \frac{1}{2} \text{ für } x \in B_r(x_0)$$

$$\implies x_* \in B_{\tilde{r}} \text{ falls } \tilde{r} = 2\|\nabla F(x_0)^{-1} \nabla F(x_0)\| \leq r$$

Neue „Idee“: Löse  $F(x) = y$  für  $y \approx y_0$  führt zum Umkehrsatz 30.2. D.h. unter den Voraussetzungen existiert Umgebung  $V = B_\rho(y_0)$  und eine Umkehrfunktion  $F^{-1} : V \rightarrow U$  mit  $F^{-1} \in C^1(V, U)$ , so dass

$$F \circ F^{-1}(y) = y \iff x = F^{-1}(y) \iff y = F(x) \wedge x \in U$$

**Beweis (von Satz 30.2)**

Für festes  $y \approx y_0$  betrachte modifizierte Funktion

$$F_y(x) := F(x) - y \in C^1(U, \mathbb{R}^n),$$

so dass

$$F_y(x) = 0 \iff F(x) = y \iff G_y(x) = x,$$

wobei

$$G_y(x) := x - [\nabla F(x_0)]^{-1} \underbrace{(F(x) - y)}_{F_y(x)}$$

Differentiation ergibt unveränderte Jacobimatrix.

$$\nabla G_y(x) = I - [\nabla F(x_0)]^{-1} \nabla F(x) \text{ da } \nabla F_y(x) = \nabla F(x)$$

Also gilt  $\|\nabla G_y(x)\| \leq \frac{1}{2}$  für  $x \in B_r(x_0)$ . Anwendung von Satz 30.1 ergibt eindeutige Lösung  $x_* = x_*(y) \in B_r(x_0)$

$$\text{falls } \|\nabla F(x_0)^{-1} F_y(x_0)\| = \|\nabla F(x_0)^{-1} (y_0 - y)\| \leq \frac{r}{2}$$

Letzte Bedingung ist sicher erfüllt, wenn  $y \in B_\rho(y_0)$  mit  $\rho = \frac{r}{2} / \|\nabla F(x_0)^{-1}\|$  da dann nach Definition der Matrixnorm

$$\|\nabla F(x_0)^{-1} (y - y_0)\| \leq \|\nabla F(x_0)^{-1}\| \|y - y_0\| \leq \frac{r}{2}$$

Dann gilt zudem für  $F^{-1}(y) := x_*(y)$  mit  $y \in B_\rho(y_0)$

$$\|F^{-1}(y) - F^{-1}(y_0)\| = \|x_*(y) - x_0\| \leq \frac{2\|y_0 - y\|}{\|\nabla F(x_0)^{-1}\|}$$

m.a.W. die konstruierte Funktion  $F^{-1} : B_\rho(y_0) \rightarrow U$  ist an Stelle  $y_0$  Lipschitzstetig mit Konstante  $2/\|\nabla F(x_0)^{-1}\|$ . Daraus folgt, dass

$$\|x - x_0\| = \|F^{-1}(y) - F^{-1}(x_0)\| \leq \frac{2\|y - y_0\|}{\|\nabla F(x_0)^{-1}\|} \implies o(\|x - x_0\|) = o(\|y - y_0\|)$$

Nunmehr erhält man Diffbarkeit von  $F^{-1}$  an Stelle  $y$  wie folgt:

$$\begin{aligned} F(x) - F(x_0) &= \nabla F(x_0)(x - x_0) + o(\|x - x_0\|) \\ y - y_0 &= \nabla F(x_0)(F^{-1}(y) - F^{-1}(y_0)) + o(\|x - x_0\|) \\ [\nabla F(x_0)]^{-1}(y - y_0) - F^{-1}(y) + F^{-1}(y_0) &= \|\nabla F(x_0)^{-1}\| o(\|x - x_0\|) \\ \implies F^{-1}(y) &= F^{-1}(y_0) + [\nabla F(x_0)]^{-1}(y - y_0) + o(\|y - y_0\|) \end{aligned}$$

Also ist  $F^{-1}$  ist an Stelle  $y_0$  total diffbar und Jacobimatrix ist genau

$$\nabla F^{-1}(y_0) = [\nabla F(F^{-1}(y_0))]^{-1}$$

Zu beweisen bleibt, dass  $F^{-1}$  an allen Stellen  $\tilde{x} \in B_\rho(y_0)$  die Ableitung  $[\nabla F(F^{-1}(\tilde{y}))]^{-1} = \nabla F^{-1}(\tilde{y})$  besitzt. Ihre Stetigkeit folgt unmittelbar aus Lipschitzstetigkeit von  $F^{-1}$  selbst und vorausgesetzte Stetigkeit von  $\nabla F$ .

An Stelle  $\tilde{x} = F^{-1}(\tilde{y}) \in B_r(x_0)$  gilt

$$\|F - [\nabla F(x_0)]^{-1} \nabla F(\tilde{x})\| \leq \frac{1}{2} \implies \det(\nabla F(\tilde{x})) \neq 0$$

Andernfalls wäre  $\nabla F(\tilde{x})v = 0$  für  $v \neq 0$  und somit  $(I - [\nabla F(x_0)]^{-1} \nabla F(\tilde{x}))v = v$  was der Kontraktivität widerspricht. Wegen  $\det(\nabla F(\tilde{x})) \neq 0$  und weiterhin gültiger Stetigkeit in  $B_r(x_0)$  lässt sich bisheriges Resultat Satz 30.2 auch an Stelle  $\tilde{x}$  statt  $x_0$  anwenden und wir erhalten Diffbarkeit mit angenehmen Jacobimatrix.

□

## Beispiel

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r \sin(\theta) \cos(\varphi) \\ r \sin(\theta) \sin(\varphi) \\ r \cos(\theta) \end{pmatrix} \equiv F(r, \theta, \varphi) \quad \mathbb{R}^3 \rightarrow \mathbb{R}^3 = F(\mathbb{R}^3)$$

$F$  ist global diffbar und surjektiv, aber nicht injektiv.

$$-F(r, \theta, \varphi) = F(r, \theta + \pi, \varphi)$$

Es gibt keine Globale Umkehrung.

$$\begin{aligned} \nabla F = \nabla_{r, \theta, \varphi} F &= \begin{pmatrix} \sin(\theta) \cos(\varphi) & r \cos(\theta) \cos(\varphi) & -r \sin(\theta) \sin(\varphi) \\ \sin(\theta) \sin(\varphi) & r \cos(\theta) \sin(\varphi) & r \sin(\theta) \cos(\varphi) \\ \cos(\theta) & -r \sin(\theta) & 0 \\ \partial/\partial r & \partial/\partial \theta & \partial/\partial \varphi \end{pmatrix} \\ r^2 \sin(\theta) = \det &\begin{pmatrix} \sin(\theta) \cos(\varphi) & \cos(\theta) \cos(\varphi) & -\sin(\varphi) \\ \sin(\theta) \sin(\varphi) & \cos(\theta) \sin(\varphi) & \cos(\varphi) \\ \cos(\theta) & -\sin(\theta) & 0 \end{pmatrix} \\ \implies r^2 \sin(\theta) &= \det(\nabla F) \end{aligned}$$

Schlussfolgerung:

$$\det(\nabla F) \neq 0 \iff r \neq 0 \wedge \sin \theta \neq 0$$

**Bemerkung (Motivation von IFT)** Betrachte Gleichung  $f(x, y) = 0 \in \mathbb{R}^m$  mit  $y \in \mathbb{R}^m$  und  $x \in \mathbb{R}^n$ , d.h.  $m$  Gleichungen in denen  $m$  Unbekannte  $y$  und den  $n$  Parametern  $x$ . Vorausgesetzt, wir haben spezielle Lösungen  $(x_0, y_0) \in \mathbb{R}^{n+m}$  mit  $f(x_0, y_0) = 0$ .

Falls

$$\nabla_y f(x_0, y_0) = \left. \frac{d}{dy} f(x_0, y) \right|_{y=y_0}$$

nicht singular, ergibt Satz 30.2 Lösbarkeit von  $f(x_0, y) = z$  für hinreichend kleine  $z \in \mathbb{R}^m$ . Andererseits gilt für  $x \approx x_0$ , dass  $f(x, y_0) \approx f(x_0, y_0) = 0$ , vorausgesetzt  $f$  stetig. Man kann also erwarten, dass es für  $x \approx x_0$  Werte  $y = y(x) \approx y_0$  gibt, so dass  $f(x, y(x)) = 0$ . Man sagt, die Funktion  $y = y(x)$  ist implizit durch die Gleichung  $f(x, y) = 0$  nahe  $(x_0, y_0)$  definiert.

### Satz 30.3 (Implizite Funktionen Theorem IFT)

Sei  $f \in C^1(D, \mathbb{R}^m)$  mit  $D \subset \mathbb{R}^{n+m}$  offen. Falls  $f(x_0, y_0) = 0$  und  $\det(\nabla_y f(x_0, y_0)) \neq 0$  an  $(x_0, y_0) \in \mathbb{R}^{n+m}$  dann existiert eine Umgebung  $U = B_r(x_0)$  und eine Funktion  $g : U \rightarrow \mathbb{R}^m$ , so dass  $f(x, g(x)) = 0$  für  $x \in U$  und  $g$  ist diffbar mit Jacobimatrix

$$\nabla_x g(x) = -[\nabla_y f(x, g(x))]^{-1} \nabla_x f(x, g(x))$$

### Bemerkung (IFT umgangssprachlich)

$$f(x, y) = 0 \in \mathbb{R}^m \text{ mit } y \in \mathbb{R}^m \text{ und } \nabla_y f(x, y) \text{ regulär}$$

erlaubt Auflösen nach bzw. Rausschmeißen von  $y$  als Funktion  $g(x)$ . Es bleiben  $n = n+m-m$  Freiheitsgrade übrig. Typischerweise könnten auch andere  $m$  Komponenten von  $z = (x, y) \in \mathbb{R}^{n+m}$  als Funktion der verbleibenden  $n$  Komponenten dargestellt werden.

**Beweis**

Betrachte erweiterte Funktion

$$\begin{aligned}
 F \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} x \\ f(x, y) \end{pmatrix} : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{n+m} \\
 \nabla_{x,y} F \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} I & 0 \\ \nabla_x f(x, y) & \nabla_y f(x, y) \end{pmatrix} \\
 \text{wobei } \nabla_x f(x, y) &= \left( \frac{\partial f_i(x, y)}{\partial x_j} \right)_{j=1 \dots n}^{i=1 \dots m} \quad \text{und} \quad \nabla_y f = \left( \frac{\partial f_i(x, y)}{\partial y_j} \right)_{j=1 \dots n}^{i=1 \dots m} \\
 \nabla_y f(x, y) &= \begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \dots & \frac{\partial f_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial y_1} & \dots & \frac{\partial f_m}{\partial y_m} \end{pmatrix} \implies \det(\nabla_{x,y} F(x, y)) = \det(\nabla_y f(x, y)) \neq 0
 \end{aligned}$$

nach Voraussetzung für  $(x, y) \approx (x_0, y_0)$  da  $\nabla_y f$  wie  $\det()$  stetige Funktionen. Betrachte  $F(x_0, y_0) = \begin{pmatrix} x_0 \\ 0 \end{pmatrix}$  da  $f(x_0, y_0) = 0$ . Nach Satz 30.2 existiert Umkehrfunktion  $F^{-1} : B_\rho \begin{pmatrix} x_0 \\ 0 \end{pmatrix} \rightarrow \mathbb{R}^{n+m}$ , so dass

$$F \circ F^{-1} \begin{pmatrix} x \\ z \end{pmatrix} = \begin{pmatrix} x \\ z \end{pmatrix} \quad \text{für} \quad \begin{pmatrix} x \\ z \end{pmatrix} \in B_\rho \begin{pmatrix} x_0 \\ 0 \end{pmatrix}$$

d.h. für  $F^{-1} \begin{pmatrix} x \\ z \end{pmatrix} = \begin{pmatrix} x \\ G(x, z) \end{pmatrix}$  mit  $G \in C^1(B_\rho \begin{pmatrix} x_0 \\ 0 \end{pmatrix}, \mathbb{R}^m)$

$$F \begin{pmatrix} x \\ G(x, z) \end{pmatrix} = \begin{pmatrix} x \\ f(x, G(x, z)) \end{pmatrix} = \begin{pmatrix} x \\ z \end{pmatrix} \quad \text{für} \quad \begin{pmatrix} x \\ z \end{pmatrix} \in B_\rho \begin{pmatrix} x_0 \\ 0 \end{pmatrix}$$

speziell erhalten wir für  $z = 0 \in \mathbb{R}^m$  und  $g(x) := G(x, 0)$ , dass  $f(x, g(x)) = f(x, G(x, 0)) = 0 \implies f(x, g(x)) = 0$ . Das ist die Aussage für  $x \in B_\rho(x_0) \implies \begin{pmatrix} x \\ 0 \end{pmatrix} \in B_\rho \begin{pmatrix} x_0 \\ 0 \end{pmatrix}$ . Differenzierbarkeit von  $F^{-1}$  und  $g(x)$  mit

$$\begin{aligned}
 [\nabla F^{-1}(x_0, 0)] &= [\nabla F(x_0, y_0)]^{-1} = \begin{pmatrix} F & 0 \\ \nabla_x G & \nabla_y G \end{pmatrix}^{-1} \begin{pmatrix} I & 0 \\ \nabla_x f(x_0, y_0) & \nabla_y f(x_0, y_0) \end{pmatrix} \\
 &= \begin{pmatrix} I & 0 \\ -[\nabla_y f(x_0, y_0)]^{-1} \nabla_x f(x_0, y_0) & [\nabla_y f(x_0, y_0)]^{-1} \end{pmatrix} = \begin{pmatrix} I & 0 \\ \nabla_x G(x, z) & \nabla_y G(x, z) \end{pmatrix}
 \end{aligned}$$

Für spezielle Wahl  $z = 0$  ergibt sich schließlich

$$\nabla_x g(x_0) = \nabla_x G(x_0, 0) = -[\nabla_y f(x_0, y_0)]^{-1} \nabla_x f(x_0, y_0)$$

wie behauptet. Stetigkeit von  $\nabla_x g(x)$  auf  $B_\rho(x_0)$  folgt aus Stetigkeit von  $g(x)$  und  $\nabla_x f(x, y)$ . □

**Bemerkung** Nach Leibniznotation haben wir einfach

$$\frac{dg}{dx} = - \left( \frac{df}{dy} \right)^{-1} \left( \frac{df}{dx} \right) \approx - \frac{dy}{dx} \text{ Merkhilfe}$$

**Beispiel**

$$f(x, y_1, y_2) = \begin{pmatrix} x^3 + y_1^3 + y_2^3 - 7 \\ x y_1 + y_1 y_2 + y_2 x + 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} x_0 \\ y_{1_0} \\ y_{2_0} \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$$

Prüfung der Regularität:

$$\frac{\partial f}{\partial(y_1, y_2)} = \begin{pmatrix} 3y_1^2 & 3y_2^2 \\ x + y_2 & y_1 + x \end{pmatrix}$$

$$\nabla_y f(x_0, y_{10}, y_{20}) = \begin{pmatrix} 3 & 0 \\ 2 & 1 \end{pmatrix} \text{ hat Determinante } 3 \neq 0$$

Also ex.  $G : B_\rho(x_0) = (x_0 - \rho, x_0 + \rho) \rightarrow \mathbb{R}^2$   
 $g = (g_1, g_2)$ , so dass  $f(x, g_1(x), g_2(x)) = 0$

Ableitung nach IFT ist an Stelle  $(x_0, y_{10}, y_{20})$ :

$$\nabla_y g \begin{pmatrix} \frac{\partial g_1}{\partial x} \\ \frac{\partial g_2}{\partial x} \end{pmatrix} = -[\nabla_y f]^{-1} \nabla_x f = - \begin{pmatrix} 3 & 0 \\ 2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 3 \cdot 4 \\ -1 \end{pmatrix} = - \begin{pmatrix} \frac{1}{3} & 0 \\ -\frac{2}{3} & 1 \end{pmatrix} \begin{pmatrix} 12 \\ -1 \end{pmatrix} = - \begin{pmatrix} 4 \\ -8 - 1 \end{pmatrix} = \begin{pmatrix} -4 \\ 9 \end{pmatrix}$$

### §31 Unrestringierte und gleichheitsrestringierte Optimierung

**Gegeben:** Zielfunktion (objective funktion)  $\varphi \in C^1(\mathcal{D}, \mathbb{R})$  mit  $\mathcal{D} \subset \mathbb{R}^n$  offen und zusammenhängend.

**Aufgabe:** Unrestringierte Optimierung heißt Auffinden von lokalem Minimalpunkt  $z_* \in \mathcal{D}$ , so dass für ein  $r > 0$

$$z \in \mathcal{D} \cap B_\rho(z_*) \implies \varphi(z) \geq \varphi(z_*)$$

m.a.W.  $z_*$  soll Minimalpunkt von  $\varphi$  in hinreichend kleiner Umgebung sein.

#### Satz 31.1 (Optimierungsbedingungen 1. Ordnung für unrestr. Minimum)

$z_* \in \mathcal{D}$  kann nur dann lokales Minimum von  $\varphi \in C^1(\mathcal{D}, \mathbb{R})$  sein, wenn

$$\nabla_z \varphi(z_*) = \left( \frac{\partial \varphi(z_*)}{\partial z_j} \right)_{j=1 \dots n} = 0$$

#### Beweis

$z_*$  lokales Minimum von  $\varphi \implies \alpha = 0$  lokales Minimum von  $\varphi(z_* - \alpha \nabla \varphi^\top(z_*)) \implies$  Nach Optimalitätsbedingung 1. Ordnung in 1. Dimension:

$$0 = \frac{d}{d\alpha} \varphi(z_* - \alpha \nabla \varphi^\top(z_*)) \Big|_{\alpha=0} = +\nabla \varphi(z_*) (-\nabla \varphi^\top(z_*)) = -\|\nabla \varphi(z_*)\|_2^2 \implies \nabla \varphi(z_*) = 0$$

□

Zusätzlich sei gegeben  $F(z) = 0 \in \mathbb{R}^m$  mit  $F : \mathcal{D} \rightarrow \mathbb{R}^m$  und  $m \leq n$ .

**Aufgabe:** Optimierung mit Gleichheitsrestriktion (=Nebenbedingung). Finde Ideales Minimum  $z_* \in F^{-1}(0) = \{z \in \mathcal{D} : F(z) = 0\}$ , so dass für  $\rho > 0$

$$z \in F^{-1}(0) \cap B_\rho(z_*) \implies \varphi(z) \geq \varphi(z_*)$$

Schreibweise üblicherweise

$$\min_{z \in \mathcal{D}} \varphi(z) \text{ sodas } F(z) = 0$$

#### Definition 31.2

Lokales Minimum  $z_*$  erfüllt LICQ (linear independence constrained qualification), wenn  $\nabla_z F(z_*)$  regulär ist, d.h. äquivalenterweise

(i)  $\text{rang}(\nabla_z F(z_*)) = m$

(ii) Vektoren  $(\nabla_z F_i(z_*))$  für  $i = 1 \dots m$  sind linear unabhängig

(iii)  $\nabla_z F$  hat eine  $m \times m$  Teilmatrix mit nicht verschwindender Determinante.

**Bemerkung** Restringiertes Problem:  $\min_{\varphi}(z)$  für  $z \in F^{-1}(0) \equiv \{z \in \mathcal{D} : F(z) = 0\}$  wobei  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  mit  $m \leq n$ .

$$\begin{aligned} \text{LICQ: } \text{rang}(\nabla_z F(z)) &= m \\ \iff (\nabla_z F_i) &\text{ linear unabhängig} \\ \iff \det(A) \neq 0 &\text{ für ein } A \in \mathbb{R}^{m \times m} \text{ mit } A \subset \nabla_z F \\ A &= (\partial_j F)_{i \in J} \text{ mit } J \subset \{1 \dots n\} \end{aligned}$$

Praktische Berechnung von  $\text{rang}(\nabla_z F)$  durch Gaußfaktorisierung mit Zeilenpermutation. Der Rang ist die Anzahl der Zeilen, die dann nicht 0 sind.

**Lemma 31.3 (Lokale Reduktion auf unrestringiertes Problem)**

Angenommen  $z_*$  ist lokales Minimalpunkt von  $\varphi(z)$  für  $z \in F^{-1}(0)$  und  $\text{rang}(\nabla_z(z_*)) = m$ . Dann können die Komponenten von  $z$  sortiert und partitioniert werden, sodass  $z = (x, y)$  mit  $x \in \mathbb{R}^{n-m}$  und  $y \in \mathbb{R}^m$ ,  $\det(\nabla_y F(z_*)) \neq 0$  und mit der entsprechenden impliziten Funktion  $g : B_{\rho}(x_*) \rightarrow \mathbb{R}^m$  gilt:

$$\hat{\varphi}(x) = \varphi(x, g(x)) : B_{\rho}(x_*) \rightarrow \mathbb{R}$$

hat lokales Minimum an der Stelle  $x = x_*$ .

**Beweis**

Nach Umordnung, so dass  $\det(\nabla_x F(z_*)) \neq 0$  garantiert IFT Existenz von  $g \in C^1(B_{\rho}(x_*), \mathbb{R}^m)$ . Damit ist  $\hat{\varphi}$  auf  $B_{\rho}(x_*)$  wohldefiniert. Angenommen  $x_*$  ist nicht lokales Minimum von  $\hat{\varphi}(x)$ . Dann existiert Folge  $x_k \in B_{\rho}(x_*)$  mit  $x_k \rightarrow x_*$  und  $\hat{\varphi}(x_k) < \hat{\varphi}(x_*) = \varphi(z_*)$ . Dann konvergiert auch die Folge  $z_k = (x_k, g(x_k))$  gegen  $x_* = (x_*, y_*)$  und  $F(z_k) = F(x_k, g(x_k)) = 0$  sowie  $\varphi(z_k) = \hat{\varphi}(x_k) < \hat{\varphi}(x_*) = \varphi(z_*)$ . Also wäre  $z_*$  auch kein lokales Minimum von  $\varphi$  in  $F^{-1}(0)$  im Widerspruch zur Annahme. □

**Bemerkung** Falls  $\text{rang}(\nabla F) = m$ , lässt sich restringiertes Problem zunächst theoretisch auf unrestringiertes Problem reduzieren.

**Korollar 31.4 (Optimalitätsbedingung 1. Ordnung für restr. Probleme)**

Unter der Voraussetzung von Lemma 31.3 existiert genau ein Vektor  $\lambda \in \mathbb{R}^m$ , so dass

$$\nabla_z \varphi(z_*) + \lambda^T \nabla_z F(z_*) = 0 \in \mathbb{R}^n, F(z_*) = 0 \in \mathbb{R}^m$$

**Beweis**

Nach Lemma 31.3 können wir o.B.d.A. annehmen, dass  $z = (x, y) \in \mathbb{R}^{n-m} \times \mathbb{R}^m$  und in  $F^{-1}(0)$  gilt lokal  $y = g(x)$  mit  $g \in C^1(B_{\rho}(x_*), \mathbb{R}^m)$ . Die Ableitung von  $g(x)$  an  $x_*$  ist nach IFT

$$\nabla_x g(x_*) = -[\nabla_y F(z_*)]^{-1} \nabla_x F(z_*)$$

Die Kettenregel ergibt für  $\hat{\varphi}(x) = \varphi(x, g(x))$ :

$$\begin{aligned} \nabla_x \hat{\varphi}(x) &= \nabla_x \varphi(x, g(x)) + \nabla_y \varphi(x, g(x)) \nabla_x g(x) \\ &= \nabla_x \varphi(x, g(x)) - \underbrace{\nabla_y \varphi(x, g(x)) [\nabla_y F(x, g(x))]^{-1}}_{\equiv \lambda^T \text{ für } z=z_*} \nabla_x F(x, g(x)) \\ &\implies \text{für } z = z_* \text{ gilt:} \\ 0 &= \nabla \varphi(z_*) + \lambda^T \nabla_x F(z_*) \\ \lambda^T &= -\nabla_y \varphi(z_*) [\nabla_y F(z_*)]^{-1} \\ \lambda^T [\nabla_y F(z_*)] &= -\nabla_y \varphi(z_*) \\ &\implies \nabla_y \varphi(z_*) + \lambda^T \nabla_y F(z_*) = 0 \text{ wie behauptet} \\ \text{da } \nabla_z \varphi(z_*) &= (\nabla_x \varphi(z_*), \nabla_y \varphi(z_*)) = (\lambda^T \nabla_x F(z_*), \lambda^T \nabla_y F(z_*)) \end{aligned}$$

*Beweis der Eindeutigkeit von  $\lambda$  durch Widerspruch:*

Gelte auch

$$\nabla_{\varphi} + \tilde{\lambda}^{\top} \nabla F = 0 = \nabla_{\varphi} + \lambda^{\top} \nabla F$$

so folgt durch Subtraktion der rechten von linker Seite

$$\begin{aligned} & (\tilde{\lambda}^{\top} - \lambda^{\top}) \nabla_z F(z_*) = 0 \\ \iff & \sum_{i=1}^m (\tilde{\lambda}_i - \lambda_i) \nabla_z F_i(z_*) = 0 \\ \implies & \lambda_i - \tilde{\lambda}_i = 0 \end{aligned}$$

Da Zeilen  $\nabla_z F_i$  linear unabhängig sind.

□

**Bemerkung** Partitionierung des Vektors  $z$  in  $(x, y)$  fällt wieder raus und wir können direkt die beiden Gleichungen nach  $z$  und  $\lambda$  lösen. Diese lassen sich zusammenfassen zur Stationaritätsbedingung

$$\nabla_{z, \lambda} L(z, \lambda) = (\nabla_z L(z, \lambda), \nabla_{\lambda} L(z, \lambda)) = 0$$

wobei  $L(z, \lambda)$  die sogenannte Lagrangefunktion ist, d.h.

$$L(z, \lambda) = \varphi(z) + \lambda^{\top} F(z) = \varphi(z) + F(z)^t \text{ op } \lambda \quad \mathbb{R}^n \rightarrow \mathbb{R}$$

Entsprechend nennt man die Komponenten  $\lambda_i$  auch Lagrange-Multiplikatoren.

### Beispiel

$$\varphi(x, y, z) = 5x + y - 3z, \text{ so dass } 0 = F(x, y, z) = \begin{pmatrix} x + y + z \\ x^2 + y^2 + z^2 - 1 \end{pmatrix} \quad \lambda, \mu \text{ Lagrange-Mult.}$$

$$L(x, y, z, \lambda, \mu) = 5x + y - 3z + \lambda(x + y + z) + \mu(x^2 + y^2 + z^2 - 1)$$

$$0 = \nabla_x L = 5 + \lambda + 3\mu x \quad \nabla_{\lambda} L = x + y + z = 0$$

$$0 = \nabla_y L = 1 + \lambda + 2\mu y \quad \nabla_{\mu} L = x^2 + y^2 + z^2 - 1 = 0$$

$$\sum 0 = (3 + 3\lambda) + 2\mu(x + y + z) = 0 \implies 0 = 3(1 + \lambda) \implies \lambda = -1$$

$$4 + 2\mu x = 0, 2\mu y = 0 \implies y = 0 \implies z = -x - y = -x \implies x^2 + z^2 + 0 = 1 = 2x^2 \implies x = \pm \frac{1}{\sqrt{2}}, z = \mp \frac{1}{\sqrt{2}}$$

Da  $F^{-1}(0) \subset B_1(0) \subset \mathbb{R}^3$  kompakt und  $\varphi$  stetig ist, muss es mindestens einen globalen Minimalpunkt und einen globalen Maximalpunkt geben. Diese müssen auch lokale Minima sein und sind deshalb über die Optimalitätsbedingungen 1. Ordnung gegeben durch

$$\left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right)$$

Frage: Können wir lokale Minima und Maxima unterscheiden?

Antwort: Ja, durch Betrachtung der Optimalitätsbedingung 2. Ordnung!

Unrestringierter Fall:  $\min \varphi(z) \quad z \in \mathcal{D} \subseteq \mathbb{R}^n$ .

Annahme:  $\varphi \in C^2(\mathcal{D})$ , d.h.  $\nabla_z \varphi : \mathcal{D} \rightarrow \mathbb{R}^2$  mit  $\nabla \varphi \in C^1(\mathcal{D})$ .

$$\begin{aligned}
 v \in \mathbb{R}^n, t \in \mathbb{R} &\implies \frac{d}{dt} \varphi(z + tv) \stackrel{K.R.}{=} \nabla \varphi(z + tv) \cdot \frac{d}{dt} (z + tv) = \sum_{j=1}^n \partial_j \varphi(z + tv) v_j \\
 \frac{d^2}{dt^2} \varphi(z + tv) &= \frac{d}{dt} \sum_{j=1}^n \partial_j \varphi(z + tv) v_j \\
 &= \sum_{j=1}^n \frac{d}{dt} \partial_j \varphi(z + tv) v_j \\
 &= \sum_{j=1}^n \sum_{i=1}^n \partial_i \partial_j \varphi(z + tv) v_j v_i \\
 &= v^\top \nabla_z^2 \varphi v \text{ wobei } \nabla_z \varphi = \begin{pmatrix} \partial_1 \partial_1 \varphi & \dots & \partial_1 \partial_n \varphi \\ \vdots & & \vdots \\ \partial_n \partial_1 \varphi & \dots & \partial_n \partial_n \varphi \end{pmatrix} \in \mathbb{R}^{n \times n}
 \end{aligned}$$

$\nabla_z^2 \varphi$  heißt *Hessematrix* und ist fast immer symmetrisch gemäß folgendem Satz.

$$\partial_i \partial_j \varphi = \frac{\partial}{\partial z_i} \left[ \frac{\partial}{\partial z_j} \varphi \right] \stackrel{?}{=} \frac{\partial}{\partial z_j} \left[ \frac{\partial}{\partial z_i} \varphi(z) \right]$$

### Satz 31.5

Falls  $\partial_i, \partial_j \varphi \in C(\mathcal{D}, \mathbb{R})$  für alle  $1 \leq i, j \leq n$  dann ist

$$\partial_i \partial_j \varphi(z) = \partial_j \partial_i \varphi(z) \text{ für alle } z \in \mathcal{D}$$

und es gilt für  $v \in \mathbb{R}^n$

$$\varphi(z + v) = \underbrace{\varphi(z) + \nabla \varphi(z)v + \frac{1}{2} v^\top \nabla^2 \varphi(z)v}_{\text{lokale quadratische Approximation}} + o(\|v\|^2)$$

### Beweis

Mit  $e_i = (0, \dots, 1, \dots)^\top \in \mathbb{R}^b$  Basis und o.B.d.A.  $z = 0$  gilt

$$\begin{aligned}
 &\varphi(v_i e_i + v_j e_j) - \varphi(v_i e_i) - \varphi(v_j e_j) + \varphi(0) \\
 &= \varphi(v_i e_i + tv_j e_j) \Big|_0^1 - \varphi(tv_j e_j) \Big|_0^1 \\
 &= \varphi(v_i e_i + tv_j e_j) - \varphi(tv_j e_j) \Big|_0^1 \\
 &= v_j \left[ \partial_j \varphi(v_i e_i + \tau v_j e_j) - \frac{\partial}{\partial_j} \varphi(\tau v_j e_j) \right] \text{ für } \tau \in [0, 1] \text{ nach ZWS} \\
 &= v_j v_i \partial_i \partial_j \varphi(\tau_2 v_i e_i + \tau v_j e_j) \text{ für } \tau_2 \in [0, 1]
 \end{aligned}$$

Wegen der Austauschbarkeit von  $i$  und  $j$ , können wir MWS in umgekehrter Richtung angewandt werden und ergibt:

$$\begin{aligned}
 &\varphi(v_i e_i + v_j e_j) - \varphi(v_i e_i) - \varphi(v_j e_j) - \varphi(0) \\
 &= v_j v_i \partial_j \partial_i \varphi(\tilde{\tau}_2 v_i e_i + \tilde{\tau} v_j e_j) \text{ für } \tilde{\tau}_2, \tilde{\tau} \in [0, 1]
 \end{aligned}$$

Division durch  $v_j, v_i \neq 0$  ergibt

$$\begin{aligned}
 \partial_j \partial_i \varphi(\tilde{\tau}_2 v_i e_i + \tilde{\tau} v_j e_j) &= \partial_i \partial_j \varphi(\tau_2 v_i e_i + \tau v_j e_j) \\
 \partial_j \partial_i \varphi(0) &= \partial_i \partial_j \varphi(0)
 \end{aligned}$$

wegen vorausgesetzter Stetigkeit.

**Beweis (der quadratischen Annäherung (Taylorformel))**

Nach einer Taylorentwicklung in einer Variable( Satz 19.3) gilt

$$\begin{aligned}
 \varphi(z+v) &= \varphi(z) + \frac{d}{dt}\varphi(z+tv)\Big|_{t=0} + \frac{1}{2}\frac{d^2}{dt^2}\varphi(z+tv)\Big|_{t=\tau} \\
 \text{Kettenregel} \quad &= \varphi(z) + \sum_{j=1}^n \partial_j \varphi(z) v_j + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \partial_{ij}^2 \varphi(z+\tau v) v_i v_j \\
 &= \varphi(z+v) - \nabla \varphi(z)^\top v - \frac{1}{2} v^\top \nabla^2 \varphi(z) v \\
 &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n [\partial_{ij}^2 \varphi(z+\tau v) - \partial_{ij}^2 \varphi(z)] v_i v_j \\
 &= o(\|v\|^2)
 \end{aligned}$$

□

**Lemma 31.6 (Aussagen der LA zu  $H = H^\top \in \mathbb{R}^{n \times n}$ )**

$H$  heißt positiv semidefinit (SPD), gdw. eine der folgenden drei äquivalenten Aussagen gilt:

(i)  $v^\top H v \geq 0$  für  $0 \neq v \in \mathbb{R}^n$

(ii)  $H v = \lambda v \implies \lambda \geq 0$  für  $0 \neq v \in \mathbb{R}^n$  und  $\lambda \in \mathbb{R}$

(iii)  $H = C C^\top$  mit  $C \in \mathbb{R}^{n \times n}$  die o.B.d.A. dreiecksförmig gewählt werden kann.

$H$  heißt positiv definit, falls die Aussage (i) und (ii) mit strikten Ungleichungen gelten oder (iii) gilt mit  $\det(H) = \det(C)^2 > 0$ . Dann existiert ein  $\lambda_{\min} > 0$ , so dass  $v^\top H v \geq \lambda_{\min} \|v\|^2 = \lambda_{\min} v^\top v$  für  $v \in \mathbb{R}^n$ . Falls obige Aussage für  $-H$  gelten, heißt  $H$  negativ semidefinit bzw. negativ definit.

**Korollar 31.7 (Unrestringierte Optimalitätsbedingung 2. Ordnung)**

Ein Punkt  $z_* \in \mathcal{D}$  kann nur dann ein lokales Minimum von  $\varphi \in C^2(\mathcal{D})$  sein, wenn  $\nabla \varphi(z_*) = 0$  und  $\nabla^2 \varphi(z_*)$  SPD ist. Falls diese notwendige Optimalitätsbedingung 2. Ordnung an  $z_* \in \mathcal{D}$  gilt, und  $\det(\nabla^2 \varphi(z_*)) > 0$ , dann muss  $z_*$  sogar lokales Minimum. (Hinreichende Optimalitätsbedingung).

**Bemerkung**  $\varphi(z) = z^k$  erfüllt an  $z_* = 0$  notwendigen, aber nicht hinreichenden Bedingung 2. Ordnung für  $k \geq 2$ . Entsprechend ist  $\nabla \varphi(z_*) = 0$  und  $\nabla^2 \varphi(z_*)$  SND kann  $z_*$  lokales Maximum sein, wenn zudem  $\det(\nabla^2 \varphi(z_*)) \neq 0$ . ( $\det(-A) = (-1)^n \det(A)$ )

**Beweis**

Laut Lemma 31.1 muss  $\nabla \varphi(z_*)$  am Minimalpunkt  $z_*$  verschwinden. Dann gilt laut Satz 31.5

$$\varphi(z_* + v) - \varphi(z_*) = \frac{1}{2} v^\top \nabla^2 \varphi(z_*) v + o(\|v\|^2)$$

wäre nun  $v^\top \nabla^2 \varphi(z_*) v < 0$  für ein  $v \neq 0$ , so gälte für dieses feste  $v$

$$\varphi(z_* + tv) - \varphi(z_*) = \frac{1}{2} t^2 v^\top \nabla^2 \varphi(z_*) v + o(t\|v\|^2) = \frac{1}{2} t^2 (v^\top \nabla^2 \varphi(z_*) v + o(t^0)) < 0$$

für hinreichend kleine  $|t| \neq 0$ . Das widerspricht vorausgesetzter Minimalität. Falls zudem  $\nabla^2 \varphi(z_*)$  positiv definit, so gilt nach Lemma 31.6

$$\varphi(z_* + v) - \varphi(z_*) = \frac{1}{2} v^\top \nabla^2 \varphi(z_*) v + o(\|v\|^2) \geq \|v\|^2 \left( \frac{1}{2} \lambda_{\min} + o(\|v\|^2) \right) > 0$$

für  $\|v\|$  hinreichend klein. Also muss  $z_*$  lokales Minimum sein.

□

**Bemerkung** Im folgenden Verallgemeinerung auf restringierten Fall mittels Reduktion.

**Satz 31.8**

Unter Voraussetzung von 31.4 (Optimalitätsbedingung 1. Ordnung) kann  $z_* \in \mathcal{D}$  nur dann ein lokales Minimum von  $\varphi(z)$ ,  $z \in F^{-1}(0)$  sein, wenn

$$\nabla_{z,\lambda} L(z_*, \lambda) = 0 \text{ und } \nabla_z F(z_*)v = 0 \implies v^\top \nabla_z^2 L(z_*, \lambda)v \geq 0$$

Gilt zudem für alle  $0 \neq v \in \ker(\nabla_z F(z_*))$  dass  $v^\top \nabla_z^2 L(z_*, \lambda)v > 0$ , dann muss  $z_*$  ein lokales Minimum des restringierten Problems sein. Falls o.B.d.A.  $z = (x, y)$  mit  $\det(F_y(x, y)) \neq 0$  und  $y = g(x)$  dann ist notwendig bzw. hinreichend für Optimalität, dass die folgende projizierte Hessematrix positiv (semi-)definit ist.

$$\nabla_x^2 \hat{\varphi}(x_*) = \begin{pmatrix} I & \nabla_x g(x_*)^\top \end{pmatrix} \begin{pmatrix} \nabla_x^2 L(z_*) & \nabla_{xy} L(z_*) \\ \nabla_{yx} L(z_*) & \nabla_y^2 L(z_*) \end{pmatrix} \begin{pmatrix} I \\ \nabla_x g(x_*) \end{pmatrix} \in \mathbb{R}^{(n-m) \times (n-m)}$$

wobei laut IFT  $\nabla g(x_0) = -[\nabla_y F(z_*)]^{-1} \nabla_x F(z_*) \in \mathbb{R}^{m \times (n-m)}$ .

**Beweis**

Nur unter der Zusatzannahme  $z = (x, y)$ . Dann gilt für red. Fkt. für  $x \approx x_*$

$$\hat{\varphi}(x) = \varphi(x, g(x)) \equiv \varphi(x, g(x)) + \lambda^\top F(x, g(x)) = L(x, y, \lambda) = L(x, y)$$

Hier wie im folgenden wird  $\lambda$  an den durch die Optimalitätsbedingung 1. Ordnung eindeutig vorgegebene Werte festgehalten, kann also als Argument von  $L$  weggelassen werden. Differenziation nach  $x_j$  liefert mit  $y = g(x)$  nach Kettenregel

$$\frac{\partial}{\partial x_j} \hat{\varphi}(x) = \frac{\partial}{\partial x_i} L(x, y) + \nabla_y L(x, y) \frac{\partial}{\partial x_j} g(x)$$

Nach nochmaligem Differenzieren nach  $x_i$  liefert

$$\begin{aligned} \frac{\partial^2}{\partial x_i \partial x_j} \hat{\varphi}(x) &= \frac{\partial^2}{\partial x_i \partial x_j} L(x, y) + \frac{\partial}{\partial x_j} \nabla_y L(x, y) \frac{\partial}{\partial x_i} g(x) + \frac{\partial}{\partial x_i} \nabla_y L(x, y) \frac{\partial}{\partial x_j} g(x) + \\ &\frac{\partial}{\partial x_i} g(x) \nabla_y^2 L(x, y) \frac{\partial}{\partial x_j} g(x) + \nabla_y L(x, y) \frac{\partial^2}{\partial x_i \partial x_j} g(x) \end{aligned}$$

An Stelle  $z_* = (x_*, y_*)$  verschwindet der Gradient  $\nabla_y L(x_*, y_*)$  und damit der letzte Term in Summe und somit alle zweiten Ableitungen der impliziten Funktion  $g(x)$ . Umschreiben der letzten Gleichung an Stelle  $z_* = (x_*, y_*)$  in Matrixnotation ergibt genau den angegebenen Ausdruck für  $\nabla_x^2 \hat{\varphi}(x_*)$ . Somit folgt Behauptung bzgl. Optimalität aus entsprechenden Aussagen in Satz 31.7 für unrestringierten Fall.

**Bemerkung** Falls sogar  $\nabla_z^2 L(z_*) \in \mathbb{R}^{n \times n}$  positiv semidefinite oder positiv Definitheit ist dann gelten diese Aussagen auch für  $T^\top \nabla_z^2 L(z_*) T$  falls  $T \in \mathbb{R}^{n \times (n-m)}$  und  $\text{rang}(T) = n - m$ . Deswegen ist

$$\nabla_z L(z_*) = 0 \text{ und } \nabla_z^2 L(z_*) \text{ positiv definit}$$

auch hinreichend für lokale Optimalität.

**Beispiel**

$$\begin{aligned} \min 5x + y - 3z \text{ so dass } F &= \begin{pmatrix} x + y + z \\ x^2 + y^2 + z^2 - 1 \end{pmatrix} = 0 \\ \implies L(x, y, z, \lambda, \mu) &= (5x + y - 3z) + \lambda(x + y + z) + \mu(x^2 + y^2 + z^2 - 1) \\ \nabla L = 0 \implies (x, y, z, \lambda, \mu) &= \begin{cases} (\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}, -1, -2\sqrt{2}) \\ (-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}, -1, 2\sqrt{2}) \end{cases} \\ \mu &= -\frac{2}{x} \end{aligned}$$

$\mu > 0$ , d.h. 2. Lösung  $\implies \nabla_z^2 L$  positiv definit.  $\implies (-\frac{1}{\sqrt{2}}, \dots, 2\sqrt{2})$  ist globaler Minimalpunkt. Für  $\mu < 0$ , d.h. 1. Lösung  $\implies \nabla_z^2 L(z_*)$  negativ definite.  $(\frac{1}{\sqrt{2}}, \dots, -2\sqrt{2})$  globaler Maximalpunkt.

## §32 Wege, Kurven und Weglängen

$f(x, y) = c$  ist ein geometrisches Objekt in der Ebene.

- Im allgemeinen eine implizit gegebene Kurve. Zum Beispiel  $f(x, y) = x^2 + y^2 = 1 = c$ .
- Unter geeigneten Umständen existiert (partielle) explizite Darstellung  $y = g(x)$ . Hier  $y = \sqrt{1 - x^2}$  für  $x \in [-1, 1]$ .
- dritte Variante: Parameterdarstellung  $x = \gamma_1(t), y = \gamma_2(t)$  oder  $(x, y) = \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)$ .

### Definition 32.1

Sei  $I = [a, b]$ ,  $\gamma : I \rightarrow \mathbb{R}^n$  stetig. Dann heißt  $\gamma$  ein Weg in  $\mathbb{R}^n$ .  $\gamma$  heißt

- Jordan-Weg, falls  $\gamma$  injektiv ist.
- geschlossener Weg, falls  $\gamma(a) = \gamma(b)$
- geschlossener Jordan-Weg, falls  $\gamma(a) = \gamma(b)$  und  $\gamma|_{(a,b)}$  injektiv ist.

$C \subset \mathbb{R}^n$  heißt Kurve in  $\mathbb{R}^n$ , falls es einen Weg  $\gamma : [a, b] \rightarrow \mathbb{R}^n$  gibt, dessen Bildmenge  $C$  ist,  $C = \{\gamma(t) : t \in (a, b)\}$  analog Jordan-Kurve, geschlossene Kurve,...

**Beispiel (Peano-Kurve)** Die Folge der  $\gamma_k : [0, 1] \rightarrow [0, 1]^2$  konvergiert gleichmäßig gegen ein  $\gamma^*$ . Dieses ist stetig, also ein Weg, und verläuft durch jeden Punkt in  $[0, 1]^2$ .

### Satz 32.2

Jede geschlossene Jordan-Kurve im  $\mathbb{R}^2$  zerlegt den  $\mathbb{R}^2$  in zwei zusammenhängende Gebiete  $\mathbb{R}^2 \setminus C = G_1 \cup G_2$ , so dass  $C$  jeweils der Rand ist,  $\partial G_1 = C$ ,  $\partial G_2 = C$ .

### Definition 32.3 (Weglänge)

$I = [a, b]$ ,  $\gamma : I \rightarrow \mathbb{R}^n$  ein Weg.  $Z : a = t_0 < t_1 < \dots < t_N = b$  eine Zerlegung von  $I$ .

$$l(\gamma, Z) = \sum_{k=1}^N \|\gamma(t_k) - \gamma(t_{k-1})\|$$

Ist  $\tilde{Z}$  eine Verfeinerung der Zerlegung  $Z$ , dann gilt  $l(\gamma, \tilde{Z}) \geq l(\gamma, Z)$ .

$$L(\gamma) = \sup\{l(\gamma, Z), Z \text{ Zerlegung von } I\}$$

Ist  $L(\gamma) < \infty$ , dann heißt  $L$  die Weglänge von  $\gamma$ .

### Lemma 32.4 (Eigenschaften)

i) ist  $\gamma$  Lipschitz-stetig, insbesondere  $\gamma \in C^1(I, \mathbb{R}^n)$ , dann ist  $L(\gamma)$  endlich.

ii)  $L(\gamma) \geq \|\gamma(b) - \gamma(a)\|$

iii) Sind  $\gamma : [a, b] \rightarrow \mathbb{R}^n, \varphi : [b, c] \rightarrow \mathbb{R}^n$  Wege mit  $\gamma(b) = \varphi(b)$ , dann ist

$$\gamma \oplus \varphi : [a, c] \rightarrow \mathbb{R}^n, \quad (\gamma \oplus \varphi)(t) = \begin{cases} \gamma(t) & \text{falls } t \in [a, b] \\ \varphi(t) & \text{falls } t \in (b, c] \end{cases}$$

der verknüpfte Weg, und es gilt  $L(\gamma \oplus \varphi) = L(\gamma) + L(\varphi)$ .

iv)  $L(\gamma - \varphi) \geq |L(\gamma) - L(\varphi)|$

### Definition 32.5 (Weglängenfunktion)

$\gamma : [a, b] \rightarrow \mathbb{R}^n$  ein Weg,  $s : [a, b] \rightarrow \mathbb{R}^n$  definiert als

$$s(a) = 0 \quad s(t) = L(\gamma|_{[a,t]}) \quad t \in (a, b]$$

heißt Weglängenfunktion.

**Satz 32.6**

Es sei  $\gamma$  **rektifizierbar** = Weg endlicher Länge. Dann folgt

- $s$  ist monoton wachsend und stetig
- $\gamma$  zusätzlich Jordanweg, dann  $s$  streng monoton wachsend
- $\gamma \in C^1(I, \mathbb{R}^n)$ , dann  $s$  ist differenzierbar und

$$s(t) = \int_a^t \|\gamma'(\tau)\| d\tau \implies s'(t) = \|\gamma'(t)\|$$

**Beispiel (Kreis)**

$$\begin{aligned} \gamma(t) &= (\cos t, \sin t) \\ \implies \gamma'(t) &= (-\sin t, \cos t) \\ \|\gamma'(t)\| &= 1, \implies s(t) = t - a \end{aligned}$$

**Beispiel (Spirale)  $r = \varphi$ .**

$$\begin{aligned} \gamma(t) &= (t \cos t, t \sin t) && , t \in [a, b] \subset [0, \infty) \\ \gamma'(t) &= (\cos t - t \sin t, \sin t + t \cos t) \\ \|\gamma'(t)\| &= \sqrt{1 + t^2} \\ \int \sqrt{1 + t^2} dt &= \int \cosh(u) \cdot \cosh(u) du \end{aligned}$$