

Analysis II - Vorlesungs-Script

Prof. Dr. Camillo De Lellis

Basisjahr 11 Semester I

Mitschrift:

Simon Hafner

Inhaltsverzeichnis

1	Metrik und Topologie des euklidischen Raumes	1
1.1	Konvergenz	3
1.2	Ein bisschen mehr Topologie	5
1.3	Stetigkeit	6
1.4	lineare Abbildungen	7
1.5	Mehr über stetige Funktionen	10
1.6	Kompakte Menge	12
1.7	Differenzierbare Funktionen	15
1.7.1	Das Differenzial	17
1.7.2	Richtungsableitung	17
1.7.3	Partielle Ableitung	18
1.8	Rechenregeln	19
1.9	Mittelwertsatz und Schrankensatz	22
1.10	Höhere (partielle) Ableitungen	23
1.11	Das Taylorpolynom zweiter Ordnung	27
1.12	Konvexität	30
1.13	Differentiation parameterabhängiger Integrale	30
2	Differenzierbare Abbildungen	36

1 Metrik und Topologie des euklidischen Raumes

$\mathbb{R}^n = \{(x_1, \dots, x_n), x \in \mathbb{R}\}$ In \mathbb{R}^n :

- Norm (Euklidische)
- Abstand (Euklidische)
- Topologie

“Abstrakte Theorie”

- Normierte Vektorräume
- Metrische Räume
- Topologische Räume

Definition 1.1. Sei $x \in \mathbb{R}^n$ ($x = (x_1, \dots, x_n)$, $x_i \in \mathbb{R}$)

$$\|x\| = \sqrt{x_1^2 + \dots + x_n^2} = \sqrt{\sum_{i=1}^n x_i^2}$$

Intuitiv: $\|x\|$ = ”der Abstand zwischen x und 0“

Lemma 1.2. $\|\cdot\|$ erfüllt die Regeln

1. $\|x\| \geq 0$ und $\|x\| = 0 \iff x = 0$
2. $\|\lambda x\| = |\lambda| \|x\| \quad \forall \lambda \in \mathbb{R}, \forall x \in \mathbb{R}$
3. $\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in \mathbb{R}$

Beweis. 1. ≥ 0 trivial

$$x = 0 \implies \sum x_i^2 = 0 \implies \|x\| = 0$$

$$x = 0 \iff x_i = 0 \forall i \iff \sum x_i^2 = 0 \iff \|x\| = 0$$

2.

$$\|\lambda x\| = \sqrt{\sum_{i=1}^n (\lambda x_i)^2} = \sqrt{\lambda^2 \sum_{i=1}^n x_i^2} = |\lambda| \sqrt{\sum_{i=1}^n x_i^2} = |\lambda| \|x\|$$

$$|\lambda| = \frac{\|x\| |\lambda|}{\|x\|}$$

3.

$$\iff \underbrace{\|x + y\|^2}_{=} \leq \|x\|^2 + \|y\|^2 + 2 \|x\| \|y\|$$

$$\begin{aligned} \sum_{i=1}^n (x_i + y_i)^2 &= \sum_{i=1}^n (x_i^2 + y_i^2 + 2x_i y_i) = \|x\|^2 + \|y\|^2 + \overbrace{2 \sum_{i=1}^n x_i y_i}^{\text{Skalarprodukt}} \\ &\iff \langle x, y \rangle \leq \|x\| \|y\| \end{aligned}$$

□

Satz 1.3. Cauchy-Schwartzsche Ungleichung

$$\sum_{i=1}^n x_i y_i \leq \sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}$$

Beweis. OBdA $y \neq 0$ ($y = 0$ trivial)

$$\begin{aligned} t \rightarrow g(t) &= \sum_{i=1}^n (x_i + ty_i)^2 \\ &= \left(\sum x_i^2 \right) + 2t \sum x_i y_i + t^2 \sum y_i^2 \\ &= \|x\|^2 + 2t \langle x, y \rangle + \|y\|^2 t^2 \end{aligned}$$

Sei $t_0 = \frac{\langle x, y \rangle}{\|y\|^2}$, dann $g(t_0) \geq 0$

$$\begin{aligned} 0 &\leq g(t_0) \\ &= \|x\|^2 - 2 \frac{\langle x, y \rangle^2}{\|y\|^2} + \|y\|^2 \frac{\langle x, y \rangle^2}{\|y\|^4} \\ &= \|x\|^2 - \frac{\langle x, y \rangle^2}{\|y\|^2} \\ \implies \langle x, y \rangle &\leq \|x\|^2 \|y\|^2 \\ \implies |\langle x, y \rangle| &\leq \|x\| \|y\| \end{aligned}$$

□

Definition 1.4. Ein normierter Vektorraum ist ein reeller Vektorraum V mit einer Abbildung $\|\cdot\| : V \rightarrow \mathbb{R}$ so dass:

1. $\|x\| \geq 0$ und $\|x\| = 0 \iff x = 0$ (Nullvektor)
2. $\|\lambda x\| = |\lambda| \|x\| \quad \forall \lambda \in \mathbb{R}, \forall x \in V$
3. $\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in V$

Beispiel 1.5. $V = \mathbb{R}^n$

$$\|x\|_p = \left(\sum |x_i|^p \right)^{\frac{1}{p}} \quad p \geq 1$$

$p = 2$ euklidische Norm

Definition 1.6. Seien $x, y \in \mathbb{R}^n$. Die euklidische Metrik $d(x, y) = \|x - y\|$

Lemma 1.7. 1. $d(x, y) \geq 0$ und $d(x, y) = 0 \iff x = y$

$$2. \quad d(x, y) = d(y, x)$$

$$3. \quad d(x, z) \leq d(x, y) + d(y, z) \quad (\text{Dreiecksungleichung})$$

Beweis.

$$\begin{aligned} \|x - z\| &\leq \underbrace{\|x - y\|}_v + \underbrace{\|y - z\|}_w \quad v + w = x - z \\ \|v + w\| &\leq \|v\| + \|w\| \end{aligned}$$

□

Definition 1.8. Ein metrischer Raum ist eine Menge X mit einer Abbildung

$$d : X \times X \rightarrow \mathbb{R} \quad (x, y) \mapsto d(x, y) \in \mathbb{R}$$

so dass

1. $d(x, y) \geq 0$ und $d(x, y) = 0 \iff x = y \quad \forall x, y \in X$
2. $d(x, y) = d(y, x) \quad \forall x, y \in X$
3. $d(x, z) \leq d(x, y) + d(y, z) \quad \forall x, y, z \in X$

Lemma 1.9. Sei $(V, \|\cdot\|)$ ein normierter Vektorraum. Dann sind V und $d(x, y) = \|x - y\|$ ein metrischer Raum.

Definition 1.10. Die offene Kugel mit Radius $r > 0$ und Mittelpunkt $x \in \mathbb{R}^n$ ist die Menge

$$K_r(x) = \{y \in \mathbb{R}^n, d(x, y) < r\}$$

Definition 1.11. Eine Menge heisst "Umgebung" von x , wenn V eine offene Kugel mit Mittelpunkt x enthält.

Definition 1.12. Eine Menge $U \subset \mathbb{R}^n$ heisst offen falls $\forall x \in U$ ist U eine Umgebung von x

$$\forall x \in U \exists \text{ eine Kugel } K_r(x) \subset U$$

Bemerkung 1.13. Eine offene Kugel ist offen.

Satz 1.14. 1. \emptyset und \mathbb{R}^n sind offen

2. Der Schnitt endlich vieler offener Mengen ist auch offen.

3. Die Vereinigung einer beliebigen Familie offener Mengen ist auch offen.

Beweis. 1. \mathbb{R}^n trivialerweise offen, auch \emptyset

2. Sei $x \in U \cap \dots \cap U_N$

$$\forall i \in \{1, \dots, N\} \quad K_r(x) \subset U_i$$

$$\text{Sei } r = \min \{r_1, \dots, r_N\}$$

$$\implies K_r(x) \subset U_i \forall i \implies K_r(x) \subset U_1 \cap \dots \cap U_N$$

3. $\{U_\lambda\}_{\lambda \in \Lambda}$. Sei $U = \bigcup_{\lambda \in \Lambda} U_\lambda$

$$x \in U \implies x \in U_\lambda \text{ für ein } \lambda \in \Lambda$$

$$\implies \exists K_r(x) \subset U_\lambda \subset U$$

□

Definition 1.15. Ein topologischer Raum ist eine Menge X und eine Menge \mathcal{O} von Teilmengen von X so dass:

1. $\emptyset, X \in \mathcal{O}$

2. $U_1 \cap \dots \cap U_N \in \mathcal{O}$ falls $U_i \in \mathcal{O}$

3. $\bigcap_{\lambda \in \Lambda} U_\lambda \in \mathcal{O}$ falls $U_i \in \mathcal{O}$

Satz 1.16. Sei (X, d) ein metrischer Raum

$$K_r(x) = \{y \in X : d(x, y) < r\}$$

Umgebungen und offene Mengen sind wie im euklidischen Fall. $\mathcal{O} = \{\text{offene Menge}\}$ definiert eine Topologie.

1.1 Konvergenz

Sei $\{x_k\}_{k \in \mathbb{N}} \quad x_k \in \mathbb{R} \quad x_k = (x_{k1}, \dots, x_{kn})$

Definition 1.17. Die Folge $\{x_k\}$ konvergiert gegen $x_\infty \in \mathbb{R}^n$ falls

$$\lim_{k \rightarrow \infty} d(x_k, x_\infty) = 0$$

$$\left(\lim_{k \rightarrow \infty} \|x_k, x_\infty\| = 0 \right)$$

Dann schreiben wir

$$x_\infty = \lim_{k \rightarrow \infty} x_k$$

Satz 1.18.

$$x_k \rightarrow x_\infty \iff x_{ki} \rightarrow x_{\infty_i} \quad \forall i \in \{1, \dots, n\}$$

Beweis.

$$\begin{aligned} \|x_k - x_\infty\| &= \sqrt{\sum_{i=1}^n (x_{ki} - x_{\infty_i})^2} \geq |x_{ki} - x_{\infty_i}| \geq 0 \\ \implies 0 &\leq \lim_{k \rightarrow \infty} |x_{ki} - x_{\infty_i}| \leq \lim_{k \rightarrow \infty} \|x_k - x_\infty\| = 0 \\ \|x_k - x_\infty\| &= \sqrt{\underbrace{\sum_{i=1}^n \underbrace{(x_{ki} - x_{\infty_i})^2}_{\rightarrow 0}}_{\rightarrow 0}} \leq \sum_{i=1}^n |x_{ki} - x_{\infty_i}| \\ \implies \|x_k - x_\infty\| &\rightarrow 0 \end{aligned}$$

Eine alternative Formulierung: $\lim_{k \rightarrow \infty} x_k = \left(\lim_{k \rightarrow \infty} x_{k1}, \dots, \lim_{k \rightarrow \infty} x_{kn} \right)$ □

Bemerkung 1.19.

$$\forall \varepsilon > 0 \exists N : \|x_k - x_\infty\| < \varepsilon \text{ falls } k \geq N$$

Für jede Umgebung U von x_∞ fast alle $x_k \in U$.

Definition 1.20. Eine Folge $\{x_k\} \subset \mathbb{R}^n$ heisst Cauchy falls:

$$\forall \varepsilon > 0 \exists N : m, k \geq N \implies \|x_k - x_m\| < \varepsilon$$

Lemma 1.21. $\{x_k\} \subset \mathbb{R}^n$ konvergiert genau dann, wenn $\{x_k\}$ Cauchy ist.

$$\text{Beweis. } \{x_k\} \text{ ist Cauchy} \implies \left\{ x_k \underbrace{i}_{\{\text{fixiert}\}} \right\} \text{ Cauchy!}$$

$$|x_{ki} - x_{mi}| \leq \|x_k - x_m\|$$

$\implies \{x_k\}$ ist eine Cauchyfolge $\xrightarrow{\text{Erstes Semester}} x_{ki}$ konvergiert $\xrightarrow{\text{Lemma 2}} x_k$ konvergiert. x_k konvergiert \implies Cauchyfolge

$$x_\infty = \lim_{k \rightarrow \infty} x_k \quad \forall \varepsilon > 0 \exists N : \|x_k - x_\infty\| < \frac{\varepsilon}{2} \quad \forall k \geq N$$

$$\begin{aligned} k, m \geq N \quad \|x_k - x_m\| &\leq \|x_k - x_\infty\| + \|x_\infty - x_m\| \leq d(x_k, x_\infty) + (x_\infty, x_m) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

□

Bemerkung 1.22. In einem metrischen Raum, Cauchy \Leftarrow Konvergenz. Aber allgemein: Cauchy $\not\implies$ Konvergenz. Falls Cauchy \implies Konvergenz, dann ist der metrische Raum vollständig.

Definition 1.23. Eine Folge $\{x_k\} \subset \mathbb{R}^n$ heisst beschränkt falls $\|x_k\|$ beschränkt ist.

Satz 1.24. 1. Eine konvergente Folge ist beschränkt

2. (Bolzano-Weierstrass) $\{x_k\}$ beschränkt $\implies \exists \{x_{k_j}\}$ die konvergiert.

Beweis.

$$\begin{aligned} \{x_k\} \text{ beschränkt} &\implies \{x_{k1}\}_{k \in \mathbb{N}} \text{ beschränkt} \\ &\implies \exists x_{k_j} : x_{k_j1} \rightarrow x_1 \end{aligned}$$

Ich definiere $y_j = x_{k_j}$ $y_{j1} \rightarrow x_1$

$$y_j \text{ beschränkt} \implies \exists j_l : y_{j_l2} \rightarrow x_2$$

$$z_l := y_{j_l} \text{ und } z_{l1} \rightarrow x_1, \quad z_{l2} \rightarrow x_2$$

... (n - 2) Schritte. w_r Teilfolge von x_k mit $w_{ri} \rightarrow x_i$

$$w_r \rightarrow (x_1, \dots, x_n)$$

□

1.2 Ein bisschen mehr Topologie

Definition 1.25. Eine Menge $G \subset \mathbb{R}^n$ heisst geschlossen falls $G^c := \mathbb{R}^n \setminus G$ eine offene Menge ist.

Bemerkung 1.26.

$$(A \cup B)^c = A^c \cap B^c$$

$$(A \cap B)^c = A^c \cup B^c$$

Satz 1.27. 1. \emptyset, \mathbb{R}^n sind abgeschlossen

2. G_1, \dots, G_N abgeschlossen $\implies G_1 \cup G_2 \cup \dots \cup G_N$ abgeschlossen

3. $\{G_\lambda\}_{\lambda \in \Lambda}$ abgeschlossen $\implies \bigcap_{\lambda \in \Lambda} G_\lambda$ abgeschlossen.

Satz 1.28. $G \subset \mathbb{R}^n$ G ist abgeschlossen $\iff \forall$ jede konvergente $\{x_k\} \subset G$ gehört der Grenzwert zu G (gilt auch für metrische Räume).

Beweis. \Leftarrow Die rechte Eigenschaft gilt. Ziel: G^c ist offen. Sei $x \in G^c$: das Ziel ist eine Kugel $K_r(x) \in G^c$ zu finden. Widerspruchsbeweis: $K_{\frac{1}{j}}(x) \not\subset G^c$, $j \in \mathbb{N} \setminus \{0\}$

$$\implies \exists x_j \in K_{\frac{1}{j}}(x) \cap G \implies \{x_j\} \subset G \text{ und } x_j \rightarrow x$$

$$\{x_j\} \subset G \quad x_j \rightarrow x \quad x \notin G$$

\implies d.h. G^c offen \implies falls $\{x_k\} \subset G$ und $x_k \rightarrow x$ dann $x \in G$
Widerspruch: G^c offen, aber $\exists \{x_k\} \subset G$ mit Grenzwert $x \notin G$, d.h. $x \in G^c$. Offenheit von G^c .

$$\implies \exists K_r(x) \subset G^c \implies K_r(x) \cap G = \emptyset$$

d.h. $\exists N$ mit

$$\|x_N - x\| < r \implies x_N \in K_r(x) \cap G$$

□

Beispiel 1.29. Eine offene Kugel ist nicht geschlossen.

$$K_r(x) = \{y : \|y - x\| < r\}$$

Sei $\{y_k\} \in K_r(x)$, (d.h. $\|y_k - x\| < r$) mit $y_k \rightarrow y$ und $\|y - x\| = r$.

Definition 1.30. Sei $\overline{K_r(x)} := \{y \in \mathbb{R}^n : \|y - x\| \leq r\}$.

Übung 1.31. $\overline{K_r(x)}$ ist abgeschlossen

Definition 1.32. $x \in \mathbb{R}^n$ ist ein Randpunkt von M falls

$$\forall K_r(x) \quad \exists y \in K_r(x) \cap M \quad \text{und} \quad \exists z \in K_r(x) \cap M^c$$

Definition 1.33. Sei M eine Menge in \mathbb{R}^n , dann ist der Rand von M

$$\partial M = \{x \in \mathbb{R}^n, \text{ Randpunkt von } M\}$$

Satz 1.34. $\partial M^c = \partial M$

1. $M \setminus \partial M$ ist die grösste offene Menge die in M enthalten ist.

2. $M \cup \partial M$ ist die kleinste geschlossene Menge die M enthält.

Beweis. $M \setminus \partial M$ ist offen.

$$\begin{aligned} x \in M \setminus \partial M &\implies x \in M \text{ und } \exists K_r(x) \text{ mit } K_r(x) \cap M^c = \emptyset \\ &\implies K_r(x) \subset M \end{aligned}$$

Sei $y \in K_r(x)$

$$\begin{aligned} &\implies |y - x| = \rho < r \\ &\implies K_{r-\rho}(y) \subset K_r(x) \subset M \implies y \in M, y \notin \partial M \\ &\quad K_r(x) \subset M \setminus \partial M \end{aligned}$$

x ist beliebig $\implies M \setminus \partial M$ ist offen.

Sei $A \subset M$ eine offene Menge. Das Ziel ist $A \subset M \setminus \partial M$. Sei $x \in A$. Ziel: $(x \in M \setminus \partial M) \implies x \notin \partial M$.

$$A \text{ offen} \implies \exists K_r(x) \subset A \subset M \implies x \notin \partial M \implies A \subset M \setminus \partial M$$

□

1.3 Stetigkeit

Definition 1.35. Sei $f : \Omega_{\mathbb{C}\mathbb{R}^n} \rightarrow \mathbb{R}^k$. f ist stetig an der Stelle $x \in \Omega$ falls $\forall \{x_k\} \subset \Omega$ mit $x_k \rightarrow x$.

$$\lim_{k \rightarrow \infty} f(x_k) = f(x)$$

Lemma 1.36. Eine äquivalente Definition:

$$\forall \varepsilon > 0 \exists \delta > 0 : f(K_\delta(x) \cap \Omega) \subset K_\varepsilon(f(x))$$

Beweis. ε - $\delta \implies$ Folgendefinition. Sei $x_k \rightarrow x$. Ziel: $f(x_k) \rightarrow f(x)$

$$\forall \varepsilon > 0 \exists \text{ mit } \underbrace{\|f(x_k) - f(x)\|}_{d(f(x_k), f(x))} < \varepsilon \quad \forall k \geq N$$

$$\exists \delta > 0 \quad \underbrace{f(K_\delta(x)) \subset K_\varepsilon(f(x))}$$

$$\exists \|x_k - x\| < \delta \quad k \geq N$$

$$x_k \in K_\delta(x) \implies f(x_k) \in K_\varepsilon(f(x))$$

Folgendefinition \implies $(\varepsilon$ - $\delta)$ -Defintion. Widerspruchsannahme:

$$\exists \varepsilon > 0 : f(K_\delta(x) \cap \Omega) \not\subset K_\varepsilon(f(x)) \quad \forall \delta > 0$$

$$\implies \forall \delta > 0 \exists y_\delta \in K_\delta(x) \text{ und } \|f(y_\delta) - f(x)\| \geq \varepsilon$$

Nehmen wir $\delta = \frac{1}{j}$ und $x_j = \frac{y_1}{j}$

$$\|x_j - x\| < \frac{1}{j} \quad (\text{weil } x_j \in K_{\frac{1}{j}}(x))$$

$$\|f(x_j) - f(x)\| = \|f(y_{\frac{1}{j}}) - f(x)\| \geq \varepsilon$$

$x_j \rightarrow x$ aber $f(x_j) \not\rightarrow f(x)$

□

Definition 1.37. Die allgemeine Defintion der Stetigkeit für metrische Räume: Seien (X, d) und (Y, \bar{d}) zwei metrische Räume. Sei $f : X \rightarrow Y$. f ist stetig an der Stelle x falls:

$$\forall \varepsilon > 0 \exists \delta > 0 \text{ mit } d(y, x) < \delta \implies \bar{d}(f(y), f(x)) < \varepsilon$$

$$\iff f(K_\delta(x)) \subset K_\varepsilon(f(x))$$

Definition 1.38. Eine $f : X \rightarrow Y$ heisst stetig falls f stetig an jeder Stelle $x \in X$ ist.

Satz 1.39. Sei $f : X \rightarrow Y$ $((X, d), (Y, \bar{d}))$ metrische Räume) Dann:

1. Die Stetigkeit in $x \iff \forall$ Umgebung U von $f(x)$ ist $f^{-1}(U)$ eine Umgebung von x .
2. Stetigkeit von $f \iff f^{-1}(U)$ ist offen $\forall U$ offen.

Beweis. 1. • Stetigkeit \implies Umgebung. U Umgebung von $f(x) \implies \exists \delta > 0$ mit $K_\delta(f(x)) \subset U$

$$\implies \exists \varepsilon > 0 : f(K_\varepsilon(x)) \subset K_\delta(f(x))$$

$$\implies f^{-1}(U) \supset f^{-1}(K_\delta(f(x))) \supset K_\varepsilon(x) \implies f^{-1}(U) \text{ Umgebung von } x$$

- Umgebung \implies Stetigkeit. Sei $\delta > 0$ $U = K_\delta(f(x))$. U Umgebung von $f(x)$. $f^{-1}(U)$ ist eine Umgebung von x .

$$\implies \exists \varepsilon > 0 : K_\varepsilon(x) \subset f^{-1}(U)$$

$$\implies f(K_\varepsilon(x)) \subset U = K_\delta(f(x))$$

2. • Stetigkeit \implies offen. Sei U offen $\iff \forall y \in U$ ist U eine Umgebung von y

$$f^{-1}U \ni x \implies f(x) \in U \xrightarrow{\text{Stetigkeit in } x} f^{-1}(U) \text{ ist eine Umgebung von } x$$

$$\implies f^{-1}(U) \text{ ist offen}$$

- offen \implies Stetigkeit an jedem $x \in X$. Sei $x \in X$, $\delta > 0$, $K_\delta(f(x))$ ist offen

$$f^{-1}(K_\delta(f(x))) \text{ ist offen} \implies x \in f^{-1}(K_\delta(f(x)))$$

$$\implies \exists \varepsilon > 0 : K_\varepsilon(x) \subset f^{-1}(K_\delta(f(x)))$$

$$f(K_\varepsilon(x)) \subset K_\delta(f(x))$$

□

1.4 lineare Abbildungen

Definition 1.40. Eine Abbildung $L : V \rightarrow W$ (V, W Vektoren) heisst linear, falls

$$L(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 L(v_1) + \lambda_2 L(v_2) \quad \forall v_1, v_2 \in V, \quad \forall \lambda_1, \lambda_2 \in \mathbb{R}$$

$$L : \mathbb{R}^n \rightarrow \mathbb{R}^k \iff \exists \text{ eine Matrix } L_{ij} :$$

$$L(x) = \left(\sum_{j=1}^n L_{1j} x_j, \sum_{j=1}^n L_{2j} x_j, \dots, \sum_{j=1}^n L_{kj} x_j \right)$$

Definition 1.41. Sei L_{ij} eine Matrix die die lineare Abbildung $L : \mathbb{R}^n \rightarrow \mathbb{R}^k$ darstellt. Die Hilbert-Schmidt Norm von L ist

$$\|L\|_{\text{HS}} = \sqrt{\sum_{i=1}^k \sum_{j=1}^n L_{ij}^2}$$

Bemerkung 1.42. $\{L : (L_{ij} n \times k \text{ Matrixen}) \sim \mathbb{R}^{nk} \|\cdot\|_{\text{HS}}$ ist die euklidische Norm.

Bemerkung 1.43. Sei $L : \mathbb{R}^n \rightarrow \mathbb{R}^k$ eine lineare Abbildung und $x \in \mathbb{R}^n$. Dann $\|L(x)\| \leq \|x\| \|L\|_{\text{HS}}$.

Korollar 1.44. Sei L wie oben, dann ist L stetig.

Beweis. Sei $x_k \rightarrow x$. Ziel $L(x_k) \rightarrow L(x)$

$$\begin{aligned}\|L(x_k) - L(x)\| &= \|L(x_k - x)\| \leq \|x_k - x\| \|L\|_{\text{HS}} \rightarrow 0 \\ \implies \|L(x_k) - L(x)\| &\rightarrow 0 \\ \implies \text{Stetigkeit}\end{aligned}$$

□

Beweis. Beweis von 1.36: $L(x) = y$

$$\begin{aligned}\|L(x)\|^2 &= \sum_{i=1}^k y_i^2 \\ &= \sum_{i=1}^k \left(\sum_{j=1}^n L_{ij} x_j \right)^2 \stackrel{\text{Cauchy-Schwartz}}{\leq} \sum_{i=1}^k \left(\sum_{j=1}^n L_{ij}^2 \right) \left(\sum_{j=1}^n x_j^2 \right) \\ &= \sum_{i=1}^k \sum_{j=1}^n L_{ij}^2 \|x\|^2 = \|x\|^2 \left(\sum_{i=1}^k \sum_{j=1}^n L_{ij}^2 \right) \\ \|x\|^2 \|L\|_{\text{HS}}^2 &\implies \|L(x)\| \leq \|x\| \|L\|_{\text{HS}}\end{aligned}$$

□

Definition 1.45. Sei $L : V \rightarrow W$ eine lineare Abbildung wobei $(V, \|\cdot\|_V)$ und $(W, \|\cdot\|_W)$ zwei endlich-dimensionalisierte Vektorräume sind. Die Operatornorm von L ist:

$$\|L\|_{L(V,W)} := \sup_{\|v\|_V \leq 1} \|L(v)\|_W$$

Satz 1.46. $\|\cdot\|_{L(V,W)}$ ist eine Norm und

$$\|L(v)\|_W \leq \|L\|_{L(V,W)} \|v\|_V$$

Deswegen: jede lineare Abbildung $L : V \rightarrow W$ ist stetig.

Beweis. Der Kern ist die folgende Eigenschaft:

$$\|L\|_{L(V,W)} < +\infty$$

Wenn das gilt dann:

1.

$$\underbrace{\|L\|_{L(V,W)}}_{\text{Kern}} \text{ und } \|L\|_{L(V,W)} = 0 \iff L = 0$$

\Leftarrow einfach. Sei $\|L\|_{L(V,W)} = 0$. Dann sei $v \in V$.

$$v = 0 \implies L(v) = 0$$

$$v \neq 0 \implies z \frac{v}{\|v\|_V} \implies \|z\|_V = 1$$

$$\|L(z)\|_W \leq \sup_{\|y\|_V \leq 1} \|L(y)\|_W = 0$$

$$\implies L(z) = 0 \implies L(v) = L(\|v\|_V z) = \|v\|_V L(z) = 0$$

2.

$$\begin{aligned}
\|\lambda L\|_{L(V,W)} &= |\lambda| \|L\|_{L(V,W)} \\
\|\lambda L\|_{L(V,W)} &= \sup_{\|y\|_V \leq 1} \|\lambda L(y)\|_W \\
&= \sup_{\|y\|_V \leq 1} |\lambda| \|L(y)\|_W \\
&= |\lambda| \sup_{\|y\|_V \leq 1} \|L(y)\|_W \\
&= |\lambda| \|L\|_{L(V,W)}
\end{aligned}$$

3.

$$\begin{aligned}
&\|L + L'\|_{L(V,W)} \\
&= \sup_{\|y\|_V \leq 1} \|(L + L')(y)\|_{L(V,W)} \\
&= \sup_{\|y\|_V \leq 1} \|L(y) + L'(y)\|_{L(V,W)} \\
&\leq \sup_{\|y\|_V \leq 1} (\|L(y)\|_W + \|L'(y)\|_W) \\
&\leq \sup_{\|y\|_V \leq 1} \|L(y)\|_W + \sup_{\|y\|_V \leq 1} \|L'(y)\|_W \\
&= \|L\|_{L(V,W)} + \|L'\|_{L(V,W)}
\end{aligned}$$

Wenn v_1, \dots, v_n Basis für V , w_1, \dots, w_k Basis für W . Die lineare Abbildung $E_{ij}(v_i) = w_j$, $E_{ij}(v_l) = 0$ falls $l \neq i$ ist eine Basis für $L(V, W) \implies L = \sum_{i,j} \lambda_{ij} E_{ij}$

□

$$(V, \|\cdot\|) \quad (W, \|\cdot\|) \quad L : V \rightarrow W \quad \|L\|_{L(V,W)} := \sup_{\|v\|_V \leq 1} \|L(v)\|_W \quad (1)$$

Satz 1.47. Falls $\dim(V), \dim(W) < +\infty$, $\|L\|_{L(V,W)} < +\infty$ Wahr ohne Beweis in V und deswegen $L(V, W), \|\cdot\|_{L(V,W)} \forall v \in V, \forall L \in L(V, W)$

$$\|L(v)\|_W \leq \|L\|_{L(V,W)} \|v\|_V \quad (2)$$

Aus 2 folgt dass L stetig ist wenn $\|L\|_{L(V,W)} < +\infty$.

Bemerkung 1.48. $\|L\|_{L(V,W)}$ ist die optimale Konstante in 2.

Beweis. Falls $\|v\|_V = 1$

$$\iff \|L(v)\|_W \leq \|L\|_{L(V,W)} = \sup_{\|v\|_V \leq 1} \|L(v)\|_W$$

Die Ungleichung ist eine direkte Folgerung von 1

$$\|v\|_V = 0 \implies L(v) = 0 \implies \|L(v)\|_W = 0 \implies 2$$

$\|v\|_V > 0$

$$\begin{aligned}
\tilde{v} &:= \frac{v}{\|v\|_V} \implies \|\tilde{v}\|_V = \frac{\|v\|_V}{\|v\|_V} = 1 \\
\|L(\tilde{v})\|_W &\leq \|L\|_{L(V,W)} \\
\left\| \frac{1}{\|v\|_V} L(v) \right\|_W &= \frac{1}{\|v\|_V} \|L(v)\|_W \\
\implies \frac{\|L(v)\|_W}{\|v\|_V} &\leq \|L\|_{L(V,W)}
\end{aligned}$$

□

Beweis. $\varepsilon - \delta$ Stetigkeit. $v, \varepsilon > 0$. Suche $\delta > 0$ mit

$$\|v' - v\|_V < \delta \implies \|L(v') - L(v)\|_W < \varepsilon$$

Linearität von L

$$\implies \|L(v') - L(v)\|_W = \|L(v' - v)\|_W$$

und aus 2

$$\begin{aligned} \|L(v' - v)\| &\leq \underbrace{\|L\|_{L(V,W)}}_{< \varepsilon} \overbrace{\|v' - v\|_V}^{< \delta} \\ \implies \delta &= \frac{\varepsilon}{\|L\|_{L(V,W)}} \end{aligned}$$

\implies Ungleichung erfüllt. □

Bemerkung 1.49. $V = \mathbb{R}^n$, $\|\cdot\|_V$ euklidische Norm. $W = \mathbb{R}^k$ mit euklidischer Norm.

$$\|L\|_{L(V,W)} \leq \|L\|_{\text{HS}}$$

$$L : \mathbb{R}^n \rightarrow \mathbb{R}^k \text{ linear}$$

$$\|L\|_{\text{HS}} = \sqrt{\sum_{i,j} L_{ij}^2}$$

$$\|L\|_{L(V,W)} := \sup_{\sum_{i=1}^n v_i^2 \leq 1} \sqrt{\sum_{j=1}^k \left(\sum_{i=1}^n L_{ji} v_i \right)^2}$$

1.5 Mehr über stetige Funktionen

Regeln für stetige Funktionen

Regel 1 Seien $f : X \rightarrow Y$, $g : X \rightarrow Y$. X : topologischer Raum, metrischer Raum, normierter Vektorraum, \mathbb{R}^n V ist ein normierter Vektorraum (\mathbb{R}^k). Falls f, g stetig sind, ist auch $f + g$ stetig.

$V = \mathbb{R}$ $fg, \frac{f}{g}$ ($g \neq 0$) stetig

$$V = \mathbb{R} \quad fg(x) = \sum_{i=1}^n f_i(x)g_i(x)$$

Beweis. Im Fall X Teilmenge von \mathbb{R}^n

$$\underbrace{\{x^k\}}_{\subset X} x^k \rightarrow x \in X$$

Stetigkeit von f und g : $g(x^k) \rightarrow g(x)$, $f(x^k) \rightarrow f(x)$.

$$g(x^k) = (g_1(x^k), \dots, g_m(x^k))$$

$$g(x) = (g_1(x), \dots, g_m(x))$$

$$f(x^k) = (f_1(x^k), \dots, f_m(x^k))$$

$$f(x) = (f_1(x), \dots, f_m(x))$$

$$(g + f)(x^k) = (g_1(x^k) + f_1(x^k), \dots, g_m(x^k) + f_m(x^k))$$

$$\rightarrow g_1(x) + f_1(x), \dots, g_m(x) + f_m(x) = (g + f)(x)$$

$$x^k \rightarrow x \in X \implies (f + g)(x^k) \rightarrow (f + g)(x).$$

□

Regel 2 Seien X, Y, Z topologische Räume. Seien $f : X \rightarrow Y$ und $g : Y \rightarrow Z$ stetig

$$g \circ f : \underbrace{X \rightarrow Z}_{x \mapsto g(f(x))}$$

Beweis. Sei U eine offene Menge in Z .

$$(g \circ f)^{-1}(U) = f^{-1}(\underbrace{g^{-1}(U)}_{\text{offen}})$$

□

Definition 1.50. Sei $f : X \rightarrow \mathbb{R}$.

$$\|f\| = \sup_{x \in X} \|f(x)\|$$

$f : X \rightarrow V$, $V, \|\cdot\|_V$ normierter Vektorraum

$$\|f\| = \sup_{x \in X} \|f(x)\|_V$$

Bemerkung 1.51. X Menge, $V, \|\cdot\|$ ein normierter Vektorraum.

$$F := \{f : X \rightarrow V\} \quad \text{mit} \quad \|f\|$$

Dann ist $F, \|\cdot\|$ ist ein normierter Vektorraum.

Definition 1.52. Eine Folge von Funktionen

$$f^k : X \rightarrow V$$

konvergiert gleichmässig gegen f falls

$$\|f^k - f\| \rightarrow 0$$

Bemerkung 1.53. $x \in X$

$$\|f^k(x) - f(x)\|_V \leq \|f^k - f\|$$

Folgerung f^k konvergiert gleichmässig

$$\implies f^k(x) \rightarrow f(x) \quad \forall x$$

Satz 1.54. Sei X ein metrischer Raum und $f^k : X \rightarrow V$ eine Folge die gleichmässig gegen f konvergiert. Dann ist f stetig.

Beweis. Seien $x \in X$ und $\varepsilon > 0$.

Ziel $\exists \delta > 0$ so dass

$$d(x, y) < \delta \implies \|f(x) - f(y)\| < \varepsilon$$

$\exists N$ so dass

$$\|f - f^k\| < \frac{\varepsilon}{3} \quad \text{falls} \quad k \geq N$$

f^N ist stetig: $\exists \delta > 0$:

$$d(x, y) < \delta \implies \|f^N(x) - f^N(y)\| < \frac{\varepsilon}{3}$$

$$d(x, y) < \delta$$

$$\|f(x) - f(y)\| = \|(f(x) - f^N(x)) + (f^N(x) - f^N(y)) + (f^N(y) - f(y))\|_V$$

$$\leq \|f(x) - f^N(x)\|_V + \|f^N(x) - f^N(y)\|_V + \|f^N(y) - f(y)\|_V$$

$$< \|f^N - f\| + \frac{\varepsilon}{3} + \|f^N - f\|$$

$$< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

□

1.6 Kompakte Menge

Definition 1.55. Eine Menge $K \subset \mathbb{R}^n$ heisst kompakt falls K abgeschlossen und beschränkt ($\iff \exists B_R(0) : K \subset B_R(0)$) ist.

Satz 1.56. Sei $k \subset \mathbb{R}^n$.

$$K \text{ kompakt} \iff \forall \{x^j\} \subset K \exists x^{j_l}$$

x^{j_l} ist eine Teilfolge, die gegen $x \in K$ konvergiert.

$K \implies$ Sei $\{x^j\}$ eine Folge

$$x^j \in K \subset B_R(0) \implies \|x^j\| < R$$

$\exists x^{j_l} \rightarrow x \in \mathbb{R}^n$, die Abgeschlossenheit von $K \implies x \in K$. Folgenkriterium \implies Abgeschlossenheit und Beschränktheit.

$$\text{nicht abgeschlossen} \implies \exists x^j \subset K \text{ mit } x^j \rightarrow \notin K$$

$$\text{Folgenkompaktheit} \implies \exists x^{j_l} \rightarrow y \in K$$

Widerspruch (weil x und y sind in derselben Menge)

Sei K nicht beschränkt.

$$\forall j \in \mathbb{N} \ B_j(0) \not\supset K$$

$$\exists x^j \in K \setminus B_j(0) \implies \|x^j\| \geq j$$

Wenn $x^{j_l} \rightarrow x$

$$\|x^{j_l}\| \leq \|x\| + \|x^{j_l} - x\|$$

$$\|x\| \leq \|x^{j_l}\| + \|x - x^{j_l}\|$$

$$\left| \|x\| - \|x^{j_l}\| \right| \leq \|x - x^{j_l}\|$$

$$\implies \|x^{j_l}\| \rightarrow \|x\|$$

$$\|x^{j_l}\| = j_l \rightarrow +\infty$$

\implies Widerspruch

Satz 1.57. $E \subset \mathbb{R}^n$

$$E \text{ kompakt} \iff E \text{ folgenkompakt}$$

d.h.

$$\forall \{x_k\} \subset E \exists \text{ Teilfolge } \{x_{k_l}\} \text{ die gegen } x \in E \text{ konvergiert}$$

Definition 1.58. (Überdeckungseigenschaft) Eine Teilmenge $E \subset \mathbb{R}^n$ besitzt die Überdeckungseigenschaft falls:

- \forall Überdeckung $\{U_\lambda\}_{\lambda \in \Lambda}$ von E mit offenen Mengen \exists endliche Teilüberdeckung.

$$\{U_\lambda\}_{\lambda \in \Lambda} \text{ Überdeckung} \iff \bigcup_{\lambda \in \Lambda} U_\lambda \supset E$$

Teilüberdeckung ist eine Teilfamilie von $\{U_\lambda\}$ die noch eine Überdeckung von E ist.

Beispiel 1.59. Eine offene Kugel hat diese Eigenschaft nicht.

$$\forall x \in K_r(0) \text{ sei } K_{\frac{r-\|x\|}{2}}(x) = U_x$$

1. $\{U_x\}_{x \in K_r(0)}$ ist eine Überdeckung von $K_r(0)$.

Einfach weil $x \in U_x$! Sei U_{x_1}, \dots, U_{x_N} eine beliebige endliche Teilfamilie. Sei

$$p := \max_{i \in \{1, \dots, N\}} \|x_i\| < r$$

\implies falls $\|y\| \geq \frac{\|x_i\| + r}{2}$ dann $y \notin U_{x_i}$. So, wenn $\|y\| \geq \frac{p+r}{2}$ dann

$$y \notin U_{x_1} \cup \dots \cup U_{x_N} \quad \frac{p+r}{2} < r$$

falls $\|y\| = \frac{p+r}{2}$, dann $y \in K_r(0)$. Mit einer geschlossenen Kugel ist das anders.

Satz 1.60. Sei $E \subset \mathbb{R}^n$

$$E \text{ kompakt} \iff E \text{ hat die Überdeckungseigenschaft}$$

Beispiel 1.61. $E = \mathbb{R}^n$, $U_n = K_{n+1}(0)$.

$$E \subset \bigcup_{n \in \mathbb{N}} U_n$$

Aber $\forall N \in \mathbb{N}$

$$\mathbb{R}^n = E \not\subset \bigcup_{n=0} U_n$$

Beweis. $\exists \{x_i\} \subset E$ ohne konvergente Teilfolge in $E \implies E$ ist nicht kompakt
 \implies Überdeckungseigenschaft gilt nicht. Zwei Möglichkeiten:

1. \exists eine Teilfolge $\{y_i\} \subset E$ $y_i \rightarrow y$ $y \notin E$
2. \exists eine Teilfolge $\{y_i\} \subset E$ $y_i \rightarrow +\infty$

Beim ersten ist die Menge offen.

$$U_0 := \mathbb{R}^n \setminus \underbrace{(\{y_1\} \cup \{y\})}_{E \text{ ist abgeschlossen}}$$

Beim zweiten gilt:

$$U_0 = \mathbb{R}^n \setminus \underbrace{\{x_i\}}_F \text{ ist offen}$$

$$U_n = U_0 \cup \{y_1, \dots, y_{n-1}\} \quad n \geq 0$$

U_n ist auch offen.

$$\bigcup_{n=0}^{\infty} U_n = \begin{cases} \mathbb{R}^n \setminus \{y\} & \text{im Fall 1} \\ \mathbb{R}^n \setminus \{x_i\} & \text{im Fall 2} \end{cases}$$

Aber jede endliche Familie

$$U_1 \cup \dots \cup U_n \not\supset E$$

in beiden Fällen lassen wir unendlich viele Punkte weg. E kompakt \implies Überdeckungseigenschaft. E ist beschränkt und abgeschlossen und sei $\{U_\lambda\}_{\lambda \in \Lambda}$ eine Familie von offenen Mengen mit $E \subset \{U_\lambda\}_{\lambda \in \Lambda}$. Wir decken die Menge U mit Würfeln:

$$[k_1, k_1 + 1] \times [k_2, k_2 + 1] \times \dots \times [k_n, k_n + 1] \\ W_1 \cup \dots \cup W_M$$

Falls jedes $E \cap W_i$ mit einer endlichen Familie von $\{U_\lambda\}$ überdeckt wird, dann finde ich eine endliche Überdeckung von E wenn N gross genug ist. So, angenommen dass die Überdeckungseigenschaft nicht gilt.

$$\exists E_i := E \cap W_i :$$

1. $\{U_\lambda\}_{\lambda \in \Lambda}$ eine Überdeckung von E_1
2. keine endliche Teilfamilie deckt E_1

Teilen wir W_i in 2^n Würfel mit Seite $\frac{1}{2}$

$$\tilde{W}_1, \dots, \tilde{W}_2$$

$\exists E_2 := E \cap \tilde{W}_i$: so dass die beiden Eigenschaften noch gelten

Induktiv

$$E \supset E_1 \supset E_2 \supset \dots$$

jede $E_i \subset W^i$ Würfel mit Seite 2^{-i+1} und die beiden Eigenschaften gelten mit E_j statt E_i .

$\{x_k\} \subset E$. $\{x_k\}$ ist eine Cauchy-Folge. $j, k > i$, $x_k, x_j \in W$ mit Seite 2^{-i+1}
 $\|x_j - x_k\| \leq \sqrt{n} 2^{-i+1}$

$$\implies x_j \rightarrow x \in E \rightarrow x \in U \in \{U_\lambda\}_{\lambda \in \Lambda} \implies K_r(x) \supset U$$

$$x \in E, x \in E^i \quad \forall i \implies x \in W^i$$

$$\implies W^i \subset B_r(x) \subset U$$

für i gross genug

$$\implies E_i \subset U$$

\implies wir haben eine endliche Teilüberdeckung $\{U\} \subset \{U_\lambda\}$ gefunden \implies Widerspruch mit den beiden Eigenschaften. \square

Bemerkung 1.62. f stetig $\implies f^{-1}(U)$ offen falls U offen.

Beweis. Sei $\{U_\lambda\}$ eine Überdeckung (mit offenen Mengen) von $f(E)$, dann ist $\{f^{-1}(U_\lambda)\}$ eine Überdeckung von E .

$$\exists f^{-1}(U_{\lambda_1}), \dots, f^{-1}(U_{\lambda_N}) \text{ Teilüberdeckung von } E$$

$U_{\lambda_1}, \dots, U_{\lambda_N}$ ist eine Überdeckung von $f(E) \implies f(E)$ ist kompakt \square

Korollar 1.63. Wenn $F : E \rightarrow \mathbb{R}$ stetig ist und $E \subset \mathbb{R}^n$ kompakt ist, besitzt f ein Maximum und ein Minimum.

Beweis. $f(E) \subset \mathbb{R}$ ist kompakt.

$$s = \sup f(E) < +\infty$$

$$\exists \{x_k\} \subset f(E) \text{ mit } x_k \rightarrow s \xrightarrow{\text{abgeschlossen}} s \in s \in f(E)$$

$$\left(s - \frac{1}{k} \implies \exists x_k \in f(E) \text{ mit } x_k > s - \frac{1}{k}, x_k \leq s \right)$$

$\implies s$ ist ein Maximum. \square

Definition 1.64. Das Intervallschachtelungsprinzip in \mathbb{R} . Sei I_j eine Intervallschachtelung:

1.

$$I_j = [a_j, b_j]$$

2.

$$I_0 \supset I_1 \supset \dots \supset I_j \supset I_{j+1}$$

3.

$$b_j - a_j \rightarrow 0$$

$$\implies \bigcap_{j=0}^{\infty} I_j \neq \emptyset$$

Satz 1.65. Sei E_j eine Folge von kompakten Mengen mit $E_j \supset E_{j+1} \forall j$ ($E_0 \subset \mathbb{R}^n$)

$$\bigcap_{j=1}^{\infty} E_j \neq \emptyset \text{ falls } E_j \neq \emptyset \forall j$$

Beweis. Sei E_j wie im Satz mit $E_j \neq \emptyset$, aber $\bigcap_{j=0}^{\infty} E_j = \emptyset$. Sei $U_j := \mathbb{R}^n \setminus E_j \implies U_j$ ist offen. $\bigcup_{j=1}^{\infty} U_j = \mathbb{R}^n \setminus \{E_j\}$ ist eine Überdeckung von E_0 . Aber $U_1 \cup \dots \cup U_N = U_N$ (weil $U_{j+1} \supset U_j$)

$$U_N \not\supset E_N \neq \emptyset \quad E_N \subset E_0$$

Keine endliche Teilfamilie von $\{U_j\}$ ist eine Überdeckung von E_0 . Widerspruch wegen Kompaktheit von E_0 . \square

1.7 Differenzierbare Funktionen

Erinnerung $f : \mathbb{R} \rightarrow \mathbb{R}$ heisst differenzierbar in $a \in \mathbb{R}$ falls

$$f'(a) = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h}$$

existiert. Was geschieht mit Funktionen von mehreren Variablen? Die “Tangentensteigung” hängt auch von der Richtung ab. D.h. Es gibt eine lineare Abbildung $L : \mathbb{R}^2 \rightarrow \mathbb{R}$

Definition 1.66. $f : U \rightarrow \mathbb{R}$, $U \subset \mathbb{R}^n$ offen, heisst differenzierbar in $a \in U$, falls

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - Lh}{\|h\|} = 0 \quad (3)$$

wobei $L : \mathbb{R}^n \rightarrow \mathbb{R}$ eine lineare Abbildung ist.

Bemerkung 1.67. $n = 1$: Es gilt $Lh = f'(a)h$

Bemerkung 1.68. Die lineare Abbildung L in 3 ist eindeutig definiert. Annahme $L' \neq L$ erfüllt die Bedingung. Sei $v \in \mathbb{R}^n$ mit $\|v\| = 1$. Es gilt:

$$(L - L')(v) \stackrel{\text{linear und } \|v\|=1}{=} \lim_{t \downarrow 0} \frac{(L - L')(tv)}{\|tv\|} \stackrel{3}{=} \lim_{h=tv} \frac{(L - L')(h)}{\|h\|} \stackrel{3}{=} 0 \implies L = L'$$

Bemerkung 1.69. Wir können 3 auch anders beschreiben:

$$f(a+h) - f(a) = Lh + \underbrace{R(h)}_{\text{Restglied}}$$

Dann gilt

$$3 \iff \lim_{h \rightarrow 0} \frac{R(h)}{\|h\|} = 0 \quad (4)$$

Definition 1.70. L heisst Differential von f in a . Man schreibt $df(a)$. Sei nun $\{e_1, \dots, e_n\}$ die Standardbasis \mathbb{R}^n , $h = (h_1, \dots, h_n) \in \mathbb{R}^n$

$$\implies df(a)h = df(a) \left(\sum_{i=1}^n h_i e_i \right) = \sum_{i=1}^n h_i df(a)e_i$$

Definition 1.71.

$$f'(a) = (df(a)e_1, \dots, df(a)e_n)$$

heisst Ableitung

Definition 1.72.

$$Tf(x, a) = f(a) + f'(a)(x - a) \quad (\text{Ebene (tangential)})$$

lineare Approximation

Satz 1.73. f differenzierbar in $a \implies f$ ist stetig in a

Beweis.

$$|f(a+b) - f(a)| = |df(a)h + R(h)| \leq |df(a)| + \underbrace{|R(h)|}_{\rightarrow 0}$$

□

Beispiel 1.74. $f(x) = Ax + b$, $A \in M_a(1, n, \mathbb{R})$, $b \in \mathbb{R}$

Behauptung 1.75. $Lh := ah$ ist linear

$$df(a)h = Ah, \quad f'(a) = A$$

Beweis.

$$f(a+h) - f(a) - Lh = R(h) = 0$$

□

Beispiel 1.76. $f(x) := x^T Ax$, $A = (a_{ij}) \in \text{Sym}(n, \mathbb{R})$

$$f(a+h) - f(a) - \underbrace{2a^T Ah}_{df(a)h} - \underbrace{h^T Ah}_{R(h)}$$

$Lh := 2a^T Ah$ ist linear (in h), $R(h) = h^T Ah (= \sum h_i a_{ik} h_k)$ z.z.: $|Rh| \leq \sum_{i,j=1}^n |a_{ij}| \|h\|_\infty^2$, d.h. $\frac{R(h)}{\|h\|} \rightarrow 0$ (falls $\|h\| \rightarrow 0$)

Ziel Wir wollen $df(a)h$ berechnen. sei $t \in \mathbb{R}$

$$\begin{aligned} f(a+th) &= f(a) + df(a)th + R(th) \\ \implies df(a)h &= \lim_{t \rightarrow 0} \frac{f(a+th) - f(a)}{t} \end{aligned} \quad (5)$$

Definition 1.77. $f : U \rightarrow \mathbb{R}$, $a \in U$. Die Richtungsableitung von f in Richtung $h \in \mathbb{R}^n$ ist der Grenzwert (falls er existiert)

$$\partial_n f(a) := \lim_{t \rightarrow 0} \frac{f(a+th) - f(a)}{t}$$

Die Ableitungen in Richtung e_1, \dots, e_n heissen partielle Ableitungen in a . Wir schreiben

$$\partial_{e_i} f(a) = \partial_i f(a) = \frac{\partial f}{\partial x_i}(a) = f_{x_i}(a)$$

Bemerkung 1.78. Wir haben nicht vorausgesetzt, dass f differenzierbar ist in a !

Satz 1.79. Sei f in a differenzierbar. Dann existieren die Richtungsableitungen in jede Richtung. Insbesondere existieren die partiellen Ableitungen. Es gelten:

$$df(a)h = f'(a)h = \partial_n f(a) = \sum_{i=1}^n \partial_i f(a) h_i \quad (6)$$

und

$$f'(a) = (\partial_1 f(a), \dots, \partial_n f(a))$$

Beweis. Existenz der Richtungsableitung oke (Herleitung von 5)

□

Frage Wie berechnet man die partielle Ableitung effizient? Es gilt:

$$\begin{aligned} \partial_i f(a) &= \lim_{t \rightarrow 0} \frac{f(a + t e_i) - f(a)}{t}, \quad a = (a_1, \dots, a_n) \\ g_i(x) &:= f(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_n) \\ \partial_i f(a) &= \lim_{t \rightarrow 0} \frac{g(a_i + t) - g(a_i)}{t} = g'(a_i) \end{aligned}$$

Beispiel 1.80.

$$\begin{aligned} f(x, y) &:= \sin(2x) e^{3y} \\ \partial_x f &= 2e^{3y} \cos(2x) \\ \partial_y f &= \sin(2x) e^{3y} 3 \end{aligned}$$

Frage Wann folgt aus der Existenz der partiellen Ableitung (Richtungsableitung) die Differenzierbarkeit?

Beispiel 1.81.

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

Es gilt: $f(tx, ty) = tf(x, y)$, d.h. der Graph von f besteht aus Geraden durch 0, für $h = (h_1, h_2) \in \mathbb{R}^2$

$$\implies \partial_h f(0, 0) = \lim_{t \rightarrow 0} \frac{f(th_1, th_2) - f(0, 0)}{t} = \lim_{t \rightarrow 0} \frac{t}{t} f(h_1, h_2) = f(h_1, h_2)$$

$$\implies \partial f(0, 0) = f(h_1, h_2)$$

$$\partial_{e_1} f(0, 0) = f(1, 0) = 0$$

$$\partial_{e_2} f(0, 0) = f(0, 1) = 0$$

Annahme f ist in $(0, 0)$ differenzierbar

$$\begin{aligned} \xrightarrow{\text{aus 6}} \underbrace{\partial_n f(0, 0)}_{=d f(a)h=0} &= \underbrace{\partial_1 f(a)}_0(h_1) + \underbrace{\partial_2 f(a)}_0(h_2) = 0 \\ \implies d f(a) &= 0 \end{aligned}$$

Test $L = 0$

$$\frac{f(h_1, h_1) - \overbrace{f(a_0) - L(h_1, h_1)}}{\|(h_1, h_1)\|_\infty} = \frac{h_1^3}{2h_1^2|h_1|} \rightarrow \pm \frac{1}{2}$$

$\implies f$ ist in $(0, 0)$ nicht differenzierbar.

1.7.1 Das Differenzial

$f : \Omega \rightarrow \mathbb{R}$, $\Omega \subset \mathbb{R}^n$, Umgebung von x .

f diff in $x \iff \exists L : \mathbb{R}^n \rightarrow \mathbb{R}$ linear s.d.

$$\lim_{h \downarrow 0} \frac{f(x+h) - f(x) - L(h)}{\|h\|} = 0 \quad (7)$$

$$\lim_{h \downarrow 0} G(h) = 0 \iff \forall \varepsilon > 0 \exists \delta > 0 \quad \|h\| < \delta \implies |G(h)| < \varepsilon$$

$$\iff \forall h_k = 0 \quad G(h_k) \rightarrow 0$$

Wenn f differenzierbar ist und 7 erfüllt, heisst L das Differential von f .

$$L = d f$$

$d f_x$ das Differential an der Stelle x

1.7.2 Richtungsableitung

$x \in \Omega$, $h \in \mathbb{R}^m$, $g(t) = f(x + th)$ (wohldefiniert für $|t|$ klein)

$$\partial_n f(x) = g'(0) = \lim_{t \rightarrow 0} \frac{f(x + th) - f(x)}{t}$$

1.7.3 Partielle Ableitung

(x_1, \dots, x_n) Kond. in \mathbb{R}^n $y \in \Omega$ so dass Ω eine Umgebung von y ist

$$\frac{\partial f}{\partial x_i}(y) (= \partial_{x_i} f(y)) = \lim_{t \rightarrow 0} \frac{y_1, \dots, y_i + t, \dots, y_n - f(y)}{t}$$

Falls $e_i = (0, \dots, 0, \underbrace{1}_i, 0, \dots, 0)$

$$= \lim_{t \rightarrow 0} \frac{f(y + te_i) - f(y)}{t} = \partial_{e_i} f(y)$$

Satz 1.82. (Hauptkriterium der Differenzierbarkeit) Sei $f : U \rightarrow \mathbb{R}$ und U eine Umgebung von y . Falls $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}$ in U existieren und stetig in y sind, dann ist f in y differenzierbar.

Beweis. $h = (h_1, \dots, h_n) \in \mathbb{R}^n$

$$L(h) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(y) h_i$$

Ziel L ist das Differential von f

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x) - L(h)}{\|h\|} = 0$$

$$f(x+h) - f(x) = f(x + (h_1, \dots, h_n)) - f(y + (h_1, \dots, h_{n-1}, 0)) + f(y + (h_1, \dots, h_{n-1}, 0)) - \dots$$

+ ... (ite Zeile)

$$+ f(y + (k, 0, \dots, 0)) - f(y) \quad (8)$$

$i \in \{1, \dots, n\}$

$$g(t) = f(y + (h_1, \dots, h_{i-1}, th_i, 0, \dots, 0))$$

$$\text{ite Zeile} = g_i(1) - g_i(0) = g'_i(\xi_i) \quad \xi \in [0, 1]$$

$$g'_i(t) = \lim_{\varepsilon \rightarrow 0} \frac{g_i(t + \varepsilon) - g_i(t)}{\varepsilon}$$

$$= h_i \lim_{\varepsilon \rightarrow 0} \frac{f(y_1 + h_1, \dots, y_{i-1}, y_i + (t + \varepsilon)h_i, y_{i+1}, \dots, y_n) - f(y_1 + h_1, \dots, y_i + th_i, \dots, y_n)}{\varepsilon h_i}$$

$$= h_i \frac{\partial f}{\partial x_i}(y_1 + h_1, \dots, y_i + th_i, y_{i+1}, \dots, y_n)$$

$$\text{ite Zeile} = h_i \frac{\partial f}{\partial x_i}(y_1 + h_1, \dots, y_{i-1}, h_{i-1}, y_i + \xi_i h_i, y_{i+1}, \dots, y_n)$$

$$\zeta_i = (h_1, \dots, h_{i-1}, \xi_i h_i, 0, \dots, 0)$$

$$= h_i \frac{\partial f}{\partial x_i}(y + \zeta_i) \quad (9)$$

9 in 8:

$$f(y+h) - f(y) = \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(y + \zeta_i) \quad (10)$$

$$f(x+h) - f(x) - L(h)$$

$$= \sum_{i=1}^n h_i \left(\frac{\partial f}{\partial x_i}(y + \zeta_i) - \frac{\partial f}{\partial x_i}(y) \right) \quad (11)$$

$$\frac{|f(x+h) - f(x) - L(h)|}{\|h\|}$$

$$\stackrel{11}{\leq} \sum_{i=1}^n \frac{|h_i| \left| \frac{\partial f}{\partial x_i}(y + \zeta_i) - \frac{\partial f}{\partial x_i}(y) \right|}{\|h\|} \quad (12)$$

Wenn $\|h\| \rightarrow 0$, $\|\zeta\| \rightarrow 0$. Die Stetigkeit von $\frac{\partial f}{\partial x_i}$ in y impliziert

$$\frac{\partial f}{\partial x_i}(y + \zeta_i) \rightarrow \frac{\partial f}{\partial x_i}$$

Die rechte Seite von 12 $\rightarrow 0$ wenn $h \rightarrow 0 \implies ??$. □

Definition 1.83. Der Gradient an der Stelle x_0 ist der Vektor

$$\left(\frac{\partial f}{\partial x_1}(x_0), \dots, \frac{\partial f}{\partial x_n}(x_0) \right) = \nabla f(x_0)$$

Bemerkung 1.84.

$$\begin{aligned} df|_{x_0}(h) (\partial_n f(x_0)) &= \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(x_0) \\ (\langle \nabla f(x_0), h \rangle) &= \nabla f(x_0) h \\ |\partial_n f(x_0)| &\stackrel{\text{Cauchy-Schwartz}}{\leq} \|\nabla f(x_0)\| \|h\| \end{aligned}$$

Falls $\|h\| = 1$, dann

$$|\partial_n f(x_0)| \leq \|\nabla f(x_0)\|$$

Fall $\|\nabla f(x_0)\| \neq 0$, wenn wir

$$K = \frac{\nabla f(x_0)}{\|\nabla f(x_0)\|}$$

bekommen wir $\|K\| = 1$ und

$$\partial_K f(x_0) = \|\nabla f(x_0)\|$$

Deswegen:

$$K = \frac{\nabla f(x_0)}{\|\nabla f(x_0)\|}$$

ist die Richtung der maximalen Steigung und

$$\|\nabla f(x_0)\|$$

ist die maximale Steigung.

1.8 Rechenregeln

Satz 1.85. Sei U eine Umgebung von $x \in \mathbb{R}^n$ und $f, g : U \rightarrow \mathbb{R}$ in x differenzierbar. Dann sind $f + g$ und fg auch differenzierbar in x und

$$d(f + g)|_x = df|_x + dg|_x$$

$$d(fg) = f(x) dg|_x + g(x) df|_x$$

Falls $f(x) \neq 0$ ist auch $\frac{1}{f}$ in x differenzierbar

$$d\left(\frac{1}{f}\right)|_x = -\frac{1}{(f(x))^2} df|_x$$

Korollar 1.86. $g(x) \neq 0$, dann

$$\begin{aligned} d\left(\frac{f}{g}\right)|_x &= \frac{1}{g(x)} df|_x - \frac{f(x)}{g(x)^2} dg|_x \\ &= \frac{g(x) df|_x - f(x) dg|_x}{g(x)^2} \end{aligned}$$

Beweis. Das Ziel ist eine lineare Abbildung L zu finden so dass

$$\lim_{h \rightarrow 0} \frac{\frac{1}{f(x+h)} - \frac{1}{f(x)} - L(h)}{\|h\|}$$

$$L = -\frac{1}{f(x)^2} \mathrm{d}f|_x$$

$$\lim_{h \rightarrow 0} \frac{\overbrace{\frac{1}{f(x+h)} - \frac{1}{f(x)} - \frac{1}{f(x)^2} \mathrm{d}f|_x(h)}^A}{\|h\|} = \frac{B+C}{\|h\|}$$

$$\frac{1}{f(x+h)} - \frac{1}{f(x)} = \frac{f(x) - f(x+h)}{f(x)f(x+h)}$$

$f(x+h) \neq 0$ falls $\|h\|$ klein genug

$$\frac{f(x+h) - f(x) - \mathrm{d}f|_x(h)}{\|h\|} \rightarrow 0$$

$$A = \left[\frac{-(-f(x) + f(x+h))}{f(x)f(x+h)} \frac{\mathrm{d}f|_x(h)}{f(x)f(x+h)} \right] = C$$

$$+ \frac{-\mathrm{d}f|_x(h)}{f(x)f(x+h)} + \frac{\mathrm{d}f|_x(h)}{f(x)^2} = B$$

$$\frac{B}{\|h\|} = -\frac{1}{f(x)f(x+h)} \underbrace{\frac{f(x+h) - f(x) - \mathrm{d}f|_x(h)}{\|h\|}}_{\rightarrow 0}$$

$$\lim_{h \rightarrow 0} f(x+h) = f(x) \neq 0$$

$$\lim_{h \rightarrow 0} \frac{B}{\|h\|} = 0$$

Diff von f für $\|h\| \rightarrow 0$

$$\frac{C}{\|h\|} = \underbrace{\frac{\mathrm{d}f|_x(h)}{\|h\|}}_{\text{ist beschränkt}} \frac{1}{f(x)} \underbrace{\left(\frac{1}{f(x)} - \frac{1}{f(x+h)} \right)}_{\rightarrow 0}$$

Sei $L = \mathrm{d}f|_x$ und $\|L\|_O$ ihre Operatornorm

$$|\mathrm{d}f|_x(h)| = |L(h)| \leq \|L\|_O \|h\|$$

$$\Rightarrow \frac{|\mathrm{d}f|_x(h)|}{\|h\|} \leq \|L\|$$

□

Definition 1.87. Eine Kurve ist eine Abbildung $\gamma : [a, b] \rightarrow \mathbb{R}^n$ (d.h. $\forall t \ \gamma(t) \in \mathbb{R}^n$)

$$\gamma(t) = (\gamma_1(t), \dots, \gamma_n(t))$$

deswegen $t \rightarrow \gamma_i(t) \in \mathbb{R}$. Die Kurve γ heisst differenzierbar wenn jede γ_i differenzierbar ist.

$$\gamma' = (\gamma'_1(t), \dots, \gamma'_n(t))$$

Satz 1.88. (Kettenregel 1. Version) Sei $f : U \rightarrow \mathbb{R}$ mit U Umgebung von x und f differenzierbar in x . Sei $\gamma : [a, b] \rightarrow U$ eine differenzierbare Kurve mit $\gamma(t_0) = x$. Sei $g = f \circ \gamma$

$$g(t) = f(\gamma(t))$$

Sei g in t_0 differenzierbar. Dann

$$g'(t_0) = \mathrm{d}|_{\gamma(t_0)}(\dot{\gamma}(t_0)) = \langle \nabla f(\gamma(t_0)), \dot{\gamma}(t_0) \rangle$$

Beweis. Das Ziel:

$$\lim_{h \rightarrow 0} \frac{g(t_0 + h) - g(t_0) - h [\mathrm{d} f|_{\gamma(t_0)}(\dot{\gamma}(t_0))]}{h} = 0$$

$$R(h) = g(t_0 + h) - g(t_0) - g(t_0) - h [\mathrm{d} f|_{\gamma(t_0)}(\dot{\gamma}(t_0))]$$
 (13)

$$\lim_{h \rightarrow 0} \frac{R(h)}{h} = 0$$
 (14)

Neue Notation

$$14 \iff R(h) = o(h)$$

$$x_0 = \gamma(t_0)$$

Annahmen: Differenzierbarkeit von f

$$\lim_{k \rightarrow 0} \frac{f(x_0 + k) - f(x_0) - \mathrm{d} f|_{x_0}(k)}{\|k\|} \left(= \frac{r(k)}{\|k\|} \right) = 0$$

$$(r(k) = o(\|k\|))$$

Differenzierbarkeit von γ :

$$\lim_{k \rightarrow 0} \frac{\gamma(x_0 + k) - \gamma(x_0) - \mathrm{d} h \gamma'|_{x_0}(k)}{h} \left(= \frac{p(k)}{\|k\|} \right) = 0$$

$$p(h) = o(h)$$

$$\gamma(t_0 + h) = \gamma(t_0) + k \left(= \gamma(t_0 + h) - \underbrace{\gamma(t_0)}_{x_0} \right)$$

$$g(t_0 + h) - g(t_0) = f(\gamma(t_0 + h)) - g(\underbrace{\gamma(t_0)}_{x_0})$$

$$= f(\gamma(t_0) + k) - f(\gamma(t_0)) = \mathrm{d} f|_{\gamma(t_0)}(k) + r(k)$$

$$= \mathrm{d} f|_{\gamma(t_0)}(\gamma(t_0 + h) - \gamma(t_0)) + r(k)$$

$$= \mathrm{d} f|_{\gamma(t_0)}(h \dot{\gamma}(t_0) + p(h)) + r(k)$$

$$\stackrel{\text{Linearit\"at von } \mathrm{d} f}{=} h \mathrm{d} f|_{\gamma(t_0)}(\dot{\gamma}(t_0)) + \mathrm{d} f|_{\gamma(t_0)}(p(h)) + r(k)$$

$$g(t_0 + h) - g(t_0) - h \mathrm{d} f|_{\gamma(t_0)}(\dot{\gamma}(t_0))$$

$$= f|_{\gamma(t_0)}(p(h)) + r(\gamma(t_0 + h) - \gamma(t_0)) = R(h)$$

$$|R(h)| \leq \frac{\overbrace{|f|_{\gamma(t_0)}(p(h))|}^L + r(\gamma(t_0 + h) - \gamma(t_0))}{\|h\|}$$

$$\leq \|L\| \frac{p(h)}{\|h\|} + \frac{r(\gamma(t_0 + h) - \gamma(t_0))}{\|h\|}$$

Ziel

$$\lim_{h \rightarrow 0} \frac{r(\gamma(t_0 + h) - \gamma(t_0))}{|h|}$$

Falls

$$r(\gamma(t_0 + h) - \gamma(t_0)) = 0$$

dann $r(0) = 0$. Wenn

$$r(\gamma(t_0 + h) - \gamma(t_0)) \neq 0$$

$$= \frac{r(\gamma(t_0 + h) - \gamma(t_0))}{\|\gamma(t_0 + h) - \gamma(t_0)\|} \frac{\|\gamma(t_0 + h) - \gamma(t_0)\|}{|h|}$$

$$\frac{r(\gamma(t_0 + h) - \gamma(t_0))}{\|\gamma(t_0 + h) - \gamma(t_0)\|} = \frac{r(k)}{\|k\|} \rightarrow 0$$

... wenn $\|k\| \rightarrow 0$ und $h \rightarrow 0$. Es fehlt die Beschränktheit von

$$\begin{aligned} & \frac{\|\gamma(t_0 + h) - \gamma(t_0)\|}{|h|} \\ & \frac{t_0 + h - \gamma(t_0)}{h} - \frac{h\dot{\gamma}(t_0)}{h} = \frac{p(h)}{h} \\ & \frac{\gamma(t_0 + h) - \gamma(t_0)}{h} = \underbrace{\dot{\gamma}(t_0)}_{\text{konstant}} + \underbrace{\frac{p(h)}{h}}_{\rightarrow 0} \end{aligned}$$

Deswegen

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{\|\gamma(t_0 + h) - \gamma(t_0)\|}{|h|} &= \|\dot{\gamma}(t_0)\| \\ \implies \frac{|R(h)|}{\|h\|} &\rightarrow 0 \end{aligned}$$

\implies Differenzierbarkeit und Kettenregel! \square

Bemerkung 1.89. Der Gradient ist orthogonal zur Niveaumenge (Höhenlinien).

Definition 1.90. Sei $\gamma : [a, b] \rightarrow U$ eine differenzierbare Kurve, U offen. Sei $f : U \rightarrow \mathbb{R}$ differenzierbar. Wenn $f(\gamma(t)) = c_0$ (c_0 hängt nicht von t ab). Dann

$$\nabla f(\gamma(t)) \perp \dot{\gamma}(t)$$

d.h.

$$\begin{aligned} \langle \nabla f(\gamma(t)), \dot{\gamma}(t) \rangle &= 0 \\ 0 = g'(t) = (f(\gamma(t)))' &\stackrel{\text{Kettenregel}}{=} \langle \nabla f(\gamma(t)), \dot{\gamma}(t) \rangle \end{aligned}$$

1.9 Mittelwertsatz und Schrankensatz

$f : [a, b] \rightarrow \mathbb{R}$, $\xi \in]a, b[$

$$f(b) - f(a) = f'(\xi)(b - a)$$

Sei nun:

$f : U \mapsto \mathbb{R}$ differenzierbar auf U

$x, y \in U$ so dass das Segment $[x, y] \subset U$

Was ist ein Segment? Gerade durch x und y

$$\{x + t(y - x) | t \in \mathbb{R}\}$$

$$[[x, y]] = \{x + t(y - x) | t \in [0, 1]\}$$

$$\gamma(t) := x + t(y - x)$$

$$f(y) - f(x) = (x_1 + t(y_1 - x_1), \dots, x_n + t(y_n - x_n))$$

γ ist differenzierbar.

$$g = f \circ \gamma \quad g(t) = f(\gamma(t))$$

$$g(1) - g(0) = g'(\tau) \quad \text{für } \tau \in]0, 1[$$

$$f(y) - f(x) = \mathrm{d}f|_{\gamma(\tau)}(\dot{\gamma}(\tau))$$

$$\dot{\gamma}(\tau) = (\gamma'_1(\tau), \dots, \gamma'_n(\tau))$$

$$= (y_1 - x_1, \dots, y_n - x_n) = y - x$$

$$\gamma(\tau) = \xi$$

$$f(y) - f(x) = \mathrm{d}f|_{\xi}(y - x) = \partial_{y-x}f(\xi) \quad (15)$$

Satz 1.91. (Mittelwertsatz) U offen, $[x, y] \subset U$ und $f : U \rightarrow \mathbb{R}$ differenzierbar. Dann $\exists \xi \in]x, y[$ so dass 15 gilt.

Definition 1.92. U sternförmig: wenn $0 \in U$ und $[x, 0] \subset U \forall x \in U$. Sternförmig mit Zentrum x_0 wenn $x_0 \in U$ $[x, x_0] \subset U \forall x \in U$

Satz 1.93. (Schränkensatz) Sei U eine offene Menge, die sternförmig ist und $f : U \rightarrow \mathbb{R}$ eine differenzierbare Funktion mit

$$\sup_{x \in U} \|df|_x\|_O = S < \infty \left(= \sup_{x \in U} \|\nabla f(x)\| \right)$$

Dann

$$|f(x) - f(0)| \leq S \|x\|$$

Wenn U konvex ist, d.h. das Segment $[x, y] \subset U \forall x, y \in U$, dann

$$|f(x) - f(y)| \leq S \|y - x\|$$

Definition 1.94. $f : \underbrace{K}_{\in \mathbb{R}^n} \rightarrow \mathbb{R}$ heisst Lipschitz wenn $\exists L[0, +\infty[$ so dass

$$|f(y) - f(x)| \leq L \|y - x\| \quad \forall x, y \in K$$

Wenn $f : (X, d) \rightarrow \mathbb{R}$ Lipschitz bedeutet die Existenz eines L so dass

$$|f(y) - f(x)| \leq Ld(y - x) \quad \forall x, y \in K$$

1.10 Höhere (partielle) Ableitungen

Sei

$$f : \Omega_{\mathbb{C}\mathbb{R}} \rightarrow \mathbb{R}$$

Die partiellen Ableitungen von f :

$$\frac{\partial f}{\partial x_i}(x) = \lim_{\varepsilon \rightarrow 0} \frac{f(x + \varepsilon e_i) - f(x)}{\varepsilon}$$

$$e_i = (0, \dots, i, \dots, 0)$$

$$\frac{\partial f}{\partial x_i} : \Omega \rightarrow \mathbb{R}$$

$$\frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right) (x) = \frac{\partial^2 f}{\partial x_j \partial x_i} (x)$$

$$= \lim_{\varepsilon \downarrow 0} \frac{\frac{\partial f}{\partial x_i}(x + \varepsilon e_j) - \frac{\partial f}{\partial x_i}(x)}{\varepsilon}$$

$$\frac{\partial^3 f}{\partial x_k \partial x_j \partial x_i} (x)$$

$$= \lim_{\varepsilon \downarrow 0} \frac{\frac{\partial^2 f}{\partial x_i}(x + \varepsilon e_j) - \frac{\partial^2 f}{\partial x_j \partial x_i}(x)}{\varepsilon}$$

$$\left(\frac{\partial f}{\partial x_i \partial x_i} \right) = \frac{\partial^2 f}{\partial x_i^2}$$

$$\frac{\partial^3 f}{\partial x_i \partial x_i \partial x_i} = \frac{\partial^3 f}{\partial x_i^3}$$

Satz 1.95. (von Schwarz) Sei $f : \Omega \rightarrow \mathbb{R}$ eine Funktion die in einer Umgebung von $p \in \Omega$ die partielle Ableitungen $\frac{\partial f}{\partial x_i}$, $\frac{\partial f}{\partial x_j}$ und $\frac{\partial^2 f}{\partial x_i \partial x_j}$ besitzt. Falls $\frac{\partial^2 f}{\partial x_i \partial x_j}$ stetig in p ist, dann existiert $\frac{\partial^2 f}{\partial x_j \partial x_i}(p)$ und

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(p) = \frac{\partial^2 f}{\partial x_j \partial x_i}(p)$$

Beispiel 1.96.

$$\begin{aligned}
f(x_1, x_2) &= \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} a_{ij} x_1^i x_2^j \\
\frac{\partial f}{\partial x_1} &= \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} i a_{ij} x_1^{i-1} x_2^j \\
\frac{\partial^2 f}{\partial x_2 \partial x_1} &= \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} i j a_{ij} x_1^{i-1} x_2^{j-1} \\
\frac{\partial f}{\partial x_2} &= \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} j a_{ij} x_1^i x_2^{j-1} \\
\frac{\partial^2 f}{\partial x_1 \partial x_2} &= \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} i j a_{ij} x_1^{i-1} x_2^{j-1}
\end{aligned}$$

Beispiel 1.97. Sei $V : \mathbb{R} \rightarrow \mathbb{R}$.

$$\begin{aligned}
v : \mathbb{R}^2 &\rightarrow \mathbb{R} \\
v(x_1, x_2) &= V(x_2) \\
\frac{\partial f}{\partial x_1} &= 0 \\
\frac{\partial^2 f}{\partial x_2 \partial x_+} &= 0
\end{aligned}$$

Beweis. Die Idee ist den Mittelwertsatz zu benutzen.**Schritt 1** Von Dimension $n \rightarrow 2$

$$\begin{aligned}
&f(x_1, \dots, x_i, \dots, x_j, \dots, x_n) \\
&p = (p_1, \dots, p_i, \dots, p_j, \dots, p_n) \\
&g : U_{\mathbb{C}\mathbb{R}} \rightarrow \mathbb{R} \\
&g(y, z) = g(p_1, \dots, p_{i-1}, y, p_{i+1}, \dots, p_{j-1}, z, p_{j+1}, \dots, p_n) \\
&\frac{\partial f}{\partial x_i}(p) = \frac{\partial g}{\partial y}(p_i, p_j) \\
&\frac{\partial f}{\partial x_j}(p) = \frac{\partial g}{\partial y}(p_i, p_j) \\
&\frac{\partial f}{\partial x_j \partial x_i} = \frac{\partial^2 g}{\partial z \partial y}(p_i, p_j) \\
&\frac{\partial f}{\partial x_i \partial x_j}(p) = \lim_{\varepsilon \downarrow 0} \frac{\frac{\partial f}{\partial x_j}(p_1, \dots, p_i + \varepsilon, \dots, p_j, \dots, p_n) - \frac{\partial f}{\partial x_j}(p)}{\varepsilon} \\
&\frac{\partial g}{\partial z}(p) = \lim_{\varepsilon \downarrow 0} \frac{\frac{\partial f}{\partial x_j}(p_i + \varepsilon p_j) - \frac{\partial f}{\partial x_j}(p)}{\varepsilon} \\
&= \frac{\partial g}{\partial y \partial z}(p_i, p_j)
\end{aligned}$$

Falls

$$\frac{\partial g}{\partial y \partial z}(p_i, p_j)$$

existiert und

$$\frac{\partial g}{\partial z \partial y}(p_i, p_j)$$

gleich, dann ist das Theorem bewiesen.

Deswegen Nun,

$$f : \Omega_{\mathbb{C}\mathbb{R}^2} \rightarrow \mathbb{R}$$

$\frac{\partial f}{\partial x_1}$, $\frac{\partial f}{\partial x_2}$ und $\frac{\partial^2 f}{\partial x_2 \partial x_1}$ existieren in einer Umgebung von $p = (a, b)$, dann ist $\frac{\partial^2 f}{\partial x_2 \partial x_1}$ stetig auf p . Zu beweisen: $\frac{\partial^2 f}{\partial x_2 \partial x_1}(p)$ existiert und

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(p) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(p)$$

$p = (a, b)$, $h, k \in \mathbb{R} \setminus \{0\}$, $Q =$ Rechteck mit Ecken (a, b) , $(a + h, b)$, $(a, b + k)$, $(a + h, b + k)$.

$$D_Q f = f(a + h, b + k) - f(a + h, b) - f(a, b + k) + f(a, b)$$

$$\begin{aligned} & \lim_{k \rightarrow 0} \lim_{h \rightarrow 0} \frac{D_Q f}{hk} \\ &= \lim_{k \rightarrow 0} \lim_{h \rightarrow 0} \frac{f(a + h, b + k) - f(a, b + k)}{hk} - \frac{f(a + h, b) - f(a, b)}{hk} \\ &= \lim_{k \rightarrow 0} \frac{\frac{\partial f}{\partial x_1}(a, b + k) - \frac{\partial f}{\partial x_1}(a, b)}{k} \\ &= \frac{\partial^2 f}{\partial x_2 \partial x_1}(a, b) \\ & \lim_{h \rightarrow 0} \left(\lim_{k \rightarrow 0} \frac{D_Q f}{hk} \right) \\ &= \lim_{h \rightarrow 0} \frac{\frac{\partial f}{\partial x_2}(a + h, b) - \frac{\partial f}{\partial x_2}(a, b)}{h} =? \end{aligned}$$

wenn der Limes existiert

$$= \frac{\partial^2 f}{\partial x_1 \partial x_2}(a, b)$$

Ziel $\lim_{h \rightarrow 0} \lim_{k \rightarrow 0} \frac{D_Q f}{hk}$ existiert und gleicht $\lim_{k \rightarrow 0} \lim_{h \rightarrow 0} \frac{D_Q f}{hk}$. \implies Satz von Schwarz

Zuerst Wir behaupten ($\forall h, k$ klein genug) die Existenz von einer Stelle $(\xi, \zeta) \in Q$ so dass

$$\frac{D_Q f}{hk} = \frac{\partial^2 f}{\partial x_2 \partial x_1}(\xi, \zeta) \quad (16)$$

$$\begin{aligned} \frac{D_Q f}{hk} &\neq \frac{1}{h} \left\{ \frac{f(a + h, b + k) - f(a + h, b)}{k} - \frac{f(a, b + k) - f(a, b)}{k} \right\} \\ &= \frac{1}{h} \{g(a + h) - g(a)\} \stackrel{\text{Mittelwertsatz}}{=} g'(\xi) \end{aligned}$$

$x \in]a, a + h[, \zeta \in]b, b + k[$ OBdA: $h, k > 0$

$$g(z) = \frac{f(z, b + k) - f(z, b)}{k}$$

$$g'(z) = \left(\frac{\partial f}{\partial x_1}(z, b + k) - \frac{\partial f}{\partial x_1}(z, b) \right) \frac{1}{k}$$

...

$$\begin{aligned} &= \frac{1}{k} \left(\frac{\partial f}{\partial x_1}(\xi, b + k) - \frac{\partial f}{\partial x_1}(\xi, b) \right) \\ &= \frac{\partial f}{\partial x_2} \left(\frac{\partial f}{\partial x_1} \right) (\xi, \zeta) \end{aligned}$$

Womit wir beim zweiten Teil von 16 wären.

$$\begin{aligned} \frac{D_Q f}{hk} - \frac{\partial^2 f}{\partial x_2 \partial x_1}(a, b) &= \frac{\partial^2 f}{\partial x_2 \partial x_1}(\xi, \zeta) - \frac{\partial^2 f}{\partial x_2 \partial x_1}(a, b) \\ \lim_{h \rightarrow 0} \lim_{k \rightarrow 0} \left(\frac{D_Q f}{hk} - \frac{\partial^2 f}{\partial x_2 \partial x_1}(a, b) \right) \\ &= \lim_{h \rightarrow 0} \lim_{k \rightarrow 0} \left(\frac{\partial^2 f}{\partial x_2 \partial x_1}(\xi, \zeta) - \frac{\partial^2 f}{\partial x_2 \partial x_1}(a, b) \right) \end{aligned}$$

$\forall \varepsilon \exists \delta$ so dass wenn $\sqrt{h^2 + k^2} < \delta$

$$\implies \left| \frac{\partial^2 f}{\partial x_2 \partial x_1}(\xi, \zeta) - \frac{\partial^2 f}{\partial x_2 \partial x_1}(a, b) \right| < \varepsilon \quad (17)$$

$$\begin{aligned} &\limsup_{h \rightarrow 0} \left| \lim_{k \rightarrow 0} \frac{D_Q f}{hk} - \frac{\partial^2 f}{\partial x_2 \partial x_1}(a, b) \right| \\ &\leq \sup_{h \in]0, \frac{\delta}{2}[} \left| \lim_{k \rightarrow 0} \frac{D_Q f}{hk} - \frac{\partial^2 f}{\partial x_2 \partial x_1}(a, b) \right| \\ &\leq \sup_{h \in]0, \frac{\delta}{2}[} \sup_{k \in]0, \frac{\delta}{2}[} \left| \frac{D_Q f}{hk} - \frac{\partial^2 f}{\partial x_2 \partial x_1}(a, b) \right| \stackrel{17}{\leq} \varepsilon \\ &\implies \limsup_{k \rightarrow 0} \dots = \lim_{k \rightarrow 0} \dots = 0 \\ &\implies \lim_{h \rightarrow 0} \lim_{k \rightarrow 0} \frac{D_Q f}{hk} \\ &= \frac{\partial^2 f}{\partial x_2 \partial x_1}(a, b) \\ &\left(= \lim_{k \rightarrow 0} \lim_{h \rightarrow 0} \frac{D_Q f}{hk} \right) \end{aligned} \quad (18)$$

$$\begin{aligned} &\lim_{h \rightarrow 0} \lim_{k \rightarrow 0} \frac{1}{hk} \{ f(a+h, b+k) - f(a+h, b) - f(a, b+h) + f(a, b) \} \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \left\{ \lim_{k \rightarrow 0} \frac{f(a+h, b+k) - f(a+h, b) - f(a, b+h) + f(a, b)}{k} \right\} \\ &\quad \lim_{h \rightarrow 0} \frac{1}{h} \left\{ \frac{\partial f}{\partial x_2}(a+h, b) - \frac{\partial f}{\partial x_2}(a, b) \right\} \\ &= \frac{\partial}{\partial x_1} \left(\frac{\partial f}{\partial x_2} \right)(a, b) \\ &= \frac{\partial^2 f}{\partial x_1 \partial x_2}(a, b) \end{aligned} \quad (19)$$

18 = 19

□

Sei $a \in \Omega$ und $w \in \mathbb{R}$. Dann

$$\begin{aligned} d^{(k)} f(a) w^k &= \sum_{i_1=1}^n \dots \sum_{i_k=1}^n \frac{\partial^k f(a)}{\partial x_{i_1} \dots \partial x_{i_k}} w_{i_1} \dots w_{i_k} \\ T_x^k f(z) &= f(x) + d f|_x(z-x) + \dots + \frac{1}{k!} d f^{(k)}|_x(z-x)^k \\ R_x^k f(z) &= \frac{1}{(k+1)!} d f^{(k+1)}|_x(z-x)^{k+1} \end{aligned}$$

Falls f beliebig mal differenzierbar ist ($f \in C^\infty(\Omega)$ d.h. die ganze partielle Ableitung existieren und sind stetig) können wir die Taylorreihe schreiben.

$$\sum_{k=0}^{\infty} \frac{1}{k!} d f^{(k)}|_x(z-x)^k$$

Konvention:

$$\frac{1}{0!} d f^{(0)}|_x(z-x)^0 = f(x)$$

Definition 1.98. Eine Funktion $f \in C^\infty(\Omega)$ heisst analytisch wenn $\forall x \in \Omega$
 $\exists B_r(x) \subset \Omega$ mit der Eigenschaft dass:

$$T_x(z) = f(z) \quad \forall z \in B_r(x)$$

$$(f \in C^\omega(\Omega))$$

1.11 Das Taylorpolynom zweiter Ordnung

$$f(z) = f(x) + \underbrace{\sum_{i=1}^n \frac{\partial f}{\partial x_i}(x)(z_i - x_i)}_{\langle \nabla f(x), z-x \rangle}$$

$$+ \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x)(z_i - x_i)(z_j - x_j) \quad (20)$$

$$+ \text{Fehler} = \sum \sum \sum \cdots (z_{i_1} - x_{i_1})(z_{i_2} - x_{i_2})(z_{i_3} - x_{i_3})$$

Die Hesse Matrix

$$Hf(x) = \left(\frac{\partial f}{\partial x_i \partial x_j}(x) \right)$$

$$\begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \frac{\partial^2 f}{\partial x_3^2} \end{pmatrix}$$

Bemerkung 1.99. Schwarz $\implies Hf(x)$ ist symmetrisch wenn alle Ableitungen zweiter Ordnung stetig sind.

$$\underbrace{\sum_i \frac{\partial^2 f}{\partial x_1 \partial x_j}(x)(z_i - x_i), \cdots, \sum_i \frac{\partial^2 f}{\partial x_n \partial x_i}(x)(z_i - x_i)}_{=Hf(x)(z-x)}$$

Deswegen

$$\sum_{j=1}^n (z_j - x_j) \sum_{i_1}^n \frac{\partial^2 f}{\partial x_j \partial x_{i_1}}(x)(z_j - x_j) = 20$$

$$= \langle z - x, Hf(x)(z - x) \rangle$$

$$= (z - x)^T Hf(x)(z - x)$$

H $n \times n$ Matrix, die Abbildung

$$w \mapsto w^T A w (= \langle w, A w \rangle)$$

ist eine "quadratische Form".

Das Taylorpolynom zweiter Ordnung

$$T_x^2 f(z) = f(x) + \langle \nabla f(x), z - x \rangle + \frac{1}{2} (z - x)^T Hf(x)(z - x)$$

Korollar 1.100. Falls $f \in C^3(\Omega)$ und $B_r(x) \subset \Omega$

$$f(z) = T_x^2 + O(\|z - x\|^3)$$

d.h.

$$|f(z) - T_x^2 f(z)| \leq C \|z - x\|^3$$

Korollar 1.101. Falls $f \in C^2(\Omega)$ und $B_r(x) \subset \Omega$, dann

$$f(z) = T_x^2 f(z) + o(\|z - x\|^2)$$

d.h.

$$\lim_{z \rightarrow x} \frac{f(z) - T_x^2 f(z)}{\|z - x\|^2} = 0$$

Beweis. Taylorapproximation mit Ordnung 1

$$\begin{aligned}
 f(z) &= T_x^1 f(z) + \frac{1}{2}(z-x)^T Hf(\zeta)(z-x) \\
 f(z) - T_x^2 f(z) &= \frac{1}{2}(z-x)^T Hf(\zeta)(z-x) - \frac{1}{2}(z-x)^T Hf(x)(z-x) \\
 &= \frac{1}{2}(z-x)^T (Hf(\zeta) - Hf(x))(z-x) \\
 &\leq \frac{1}{2} \|z-x\| \|Hf(\zeta) - Hf(x)\| \|z-x\| \\
 &\leq \frac{1}{2} \|z-x\| \|Hf(\zeta) - Hf(x)\|_O \|z-x\| \\
 &= \frac{1}{2} \|z-x\|^2 \|Hf(\zeta) - Hf(x)\|_O \\
 \frac{|f(z) - T_x^2 f(z)|}{\|z-x\|^2} &\leq \frac{1}{2} \|Hf(\zeta) - Hf(x)\|_O \\
 \|\zeta - x\| &\leq \|z-x\|^x
 \end{aligned}$$

Stetigkeit der Ableitungen 2. Ordnung

$$\implies \lim_{\zeta \rightarrow x} \|Hf(\zeta) - Hf(x)\|_O = 0$$

□

Definition 1.102. $X \subset \mathbb{R}^n$, $\exists : X \rightarrow \mathbb{R}$. f hat in $a \in X$ ein lokales Minimum/Maximum

$$\iff \exists a \in V \text{ (Umgebung) } . f(a) \leq f(x) \text{ (bzw. } \geq f(x)) \forall x \in V$$

Man sagt das Minimum/Maximum ist strikt (oder isoliert)

$$\iff f(a) < f(x) \text{ (bzw. } > f(x)) \forall x \in V \setminus \{a\}$$

Satz 1.103. (Notwendiges Kriterium für lokale Extrema). Sei $U \subset \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}$ haben ein lokales Extremum in $a \in U$ und sei partiell differenzierbar. Dann gilt

$$\partial_1 f(a) = \dots = \partial_n f(a) = 0$$

D.h. wenn f differenzierbar ist, dann gilt $df|_a = 0$

Beweis. $F(t) = f(a + t e_i)$ (für t sehr klein, so dass $a + t e_i \in U$) F hat lokale Extrema in 0, d.h. $F'(0) = \partial_1 f(a) = 0$ □

Definition 1.104. f differenzierbar, dann heisst a mit $df|_a = 0$ kritischer Punkt. Man sagt auch f ist stationär in a .

Bemerkung 1.105. lokale Extremum $\implies \nRightarrow$ kritischer Punkt.

Satz 1.106. (Hinreichendes Kriterium für lokale Extrema) $U \subset \mathbb{R}^n$ offen, $f \in C^2(U, \mathbb{R})$ it $df|_a = 0$. Dann

$$\begin{aligned}
 H_f(a) &> 0 \implies a \text{ lokales Minimum} \\
 H_f(a) &< 0 \implies a \text{ lokales Maximum} \\
 H_f(a) &\text{ indefinit} \implies a \text{ kein Extremum}
 \end{aligned}$$

Im indefiniten Fall gilt: \exists Geraden G_1, G_2 durch a so dass $f|_{G_1 \cap U}$ in a ein lokales Minimum und $f|_{G_2 \cap U}$ in a ein lokales Maximum hat, d.h. a ist ein Sattelpunkt.

Bemerkung 1.107. • $H_f(a) > 0$ bedeutet $H_f(a)$ positiv definit, d.h.

$$v^T H_f(a) v > 0 \quad \forall v \in \mathbb{R} \setminus \{0\}$$

- $H_f(a)$ indefinit, $\exists v, w \in \mathbb{R}^n \setminus \{0\}$ mit

$$v^T H_f(a) v > 0$$

$$w^T H_f(a) w < 0$$

Beh. $H_f(a) > 0$

$$df|_a = 0 \xrightarrow{\text{Taylor}} f(a+h) = f(a) + \frac{1}{2} h^T H_f(a) h + R(h)$$

mit

$$\frac{R(h)}{\|h\|^2} \rightarrow 0 \quad (\|h\| \rightarrow 0)$$

$f \in C^2$

- $\implies h \mapsto h^T H_f(a) h$ ist stetig
- $\implies \dots$ hat ein Minimum auf $\{\|h\| = 1\}$ (kompakt), $m > 0$ (da $H_f(a) > 0$).
- $\implies h^T H_f(a) h \geq m \|h\|^2$ (da $h = \|h\| \frac{h}{\|h\|}$, $h \neq 0$)

Wähle $\varepsilon > 0$ so klein, dass $B_\varepsilon(a) \subset U$

$$|R(h)| \leq \frac{m}{4} \|h\|^2 \quad \forall h \in B_\varepsilon(a)$$

$$\implies f(a+h) \geq f(a) + \frac{m}{4} \|h\|^2 > f(a) \quad \forall h \in B_\varepsilon(a)$$

d.h. f hat in a ein lokales Minimum

$H_f(a) < 0$ Betrachte $-f$ wie oben.

$H_f(a) < 0$

$$\exists v, w : v^T H_f(a) v > 0, w^T H_f(a) w < 0$$

$$F_v(t) := f(a+tv), F_w(t) = f(a+tw)$$

$$\implies F_v''(0) > 0 \implies \text{lokales Maximum}$$

$$\implies F_w''(0) < 0 \implies \text{lokales Minimum}$$

\implies Beh

□

Bemerkung 1.108. Mit diesem Satz lässt sich keine Aussage machen, falls $H_f(a)$ semidefinitiv ist, d.h. $H_f(a) \geq 0$, $H_f(a) \leq 0$.

Beispiel 1.109. $f(x, y) = y^2(x-1) + x^2(x+1)$

$$df|_{(x,y)} = (y^2 + 3x^2 + 2x, 2(x-1)y)$$

$\implies df|_{(x,y)} = (0, 0) \implies$ kritische Punkte:

$$P_1 = (0, 0), P_2 = \left(-\frac{2}{3}, 0\right)$$

$$\implies H_f(x, y) = \begin{pmatrix} 6x+2 & 2y \\ 2y & 2(x-1) \end{pmatrix}$$

d.h.

$$\implies H_f(P_1) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

indefinit, d.h. Sattelpunkt.

$$\implies H_f(P_2) = \begin{pmatrix} -2 & 0 \\ 0 & -\frac{10}{3} \end{pmatrix} < 0$$

d.h. lokales Maximum

Beispiel 1.110. $f(x, y) = x^2 + y^3$, $g(x, y) = x^2 + y^4$ Beim Punkt 0 ist die Hesse-Matrix in beiden Fällen $\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$. Daraus kann man nichts schliessen (siehe Graphen (Freiwilliger gesucht))

1.12 Konvexität

Definition 1.111. $U \subset \mathbb{R}^n$ heisst konvex

$$\iff \forall x, y \in U : [x, y] \subset U$$

Definition 1.112. $f : U \rightarrow \mathbb{R}$ heisst konvex

$$\iff \forall x, y \in U : f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$$

- Falls $\forall x, y \in U \forall t \in (0, 1)$ “<”, heisst die Funktion strikt konvex.
- f heisst (streng) konkav, falls $-f$ (streng) konvex

Bemerkung 1.113. f ist konvex

$$\iff \forall x \neq y \in U : F_{x,y}(t) = f(x + t(y-x)) \text{ konvex (auf } [x, y])$$

Satz 1.114. (Konvexitätskriterium) Sei $f : U \rightarrow \mathbb{R}, C^2$ $U \subset \mathbb{R}^n$ offen, konvex. Es gilt:

1. f konvex $\iff H_f(x) \geq 0 \forall x \in U$
2. $H_f(x) > 0 \forall x \in U \implies f$ streng konvex

Bemerkung 1.115. Umkehrung von 1 gilt nicht, z.B. $f(x, y) = x^4 + y^4$

Beweis. 1. f konvex: $\forall x \in U$ wähle $r > 0: B_r(x) \subset U$

$$\implies F_{x, x+h}(t) \text{ konvex } \forall h \in B_r(0)$$

$$\implies h^T H_f(x) h = F''_{x, x+h}(0) \underbrace{\geq}_{\text{Konvexität in 1-Dim}} 0 \forall h \in B_r(0)$$

$$\xrightarrow{\text{homogenität}} h^T H_f(x) h \geq 0 \forall h \in \mathbb{R}^n, \text{ d.h. } H_f(x) \text{ positiv semidefinit}$$

$$H_f(x) \geq 0 \forall x \in U:$$

$$a, b \in U \implies F''_{a,b}(t) = (b-a)^T H_f(a + t(b-a))(b-a) \geq 0$$

$$\implies F_{a,b} \text{ konvex } \forall a, b \in U \implies \text{Behauptung}$$

2. Analog wie die zweite Richtung im Ersten.

□

1.13 Differentiation parameterabhängiger Integrale

$f : U \times [a, b] \rightarrow \mathbb{R}$, $U \subset \mathbb{R}^n$ offen. Sei $t \rightarrow f(x, t)$ stetig. $\forall x \in U$. Definiere

$$F(x) := \int_a^b f(x, t) dt \quad x \in U$$

Satz 1.116. Sei f wie oben und es gelte:

1. $\forall t \in [a, b]: x \mapsto f(x, t)$ nach x_i partiell differenzierbar
2. $(x, t) \mapsto \partial_i f(x, t)$ ist stetig auf $U \times [a, b]$

$\implies F$ ist nach x_i stetig partiell differenzierbar, und es gilt:

$$\frac{\partial F}{\partial x_i}(x) = \int_a^b \frac{\partial f}{\partial x_i}(x, t) dt$$

Sei $f : \underbrace{U}_{\subset \mathbb{R}} \times [a, b] \rightarrow \mathbb{R}$ stetig (U offen) $\forall x \in U$ sei

$$F(x) = \int_a^b f(x, t) dt$$

Satz 1.117. (*Differentiationssatz*) Falls

1. $\forall t \in [a, b]$ ist $x \mapsto f(x, t)$ nach x_i partiell differenzierbar

$$\exists \frac{\partial f}{\partial x_i}(x, t) \quad \forall (x, t) \in U \times [a, b]$$

2. und $\frac{\partial f}{\partial x_i}$ ist stetig

dann \exists auch $\frac{\partial F}{\partial x_i}(x)$ und

$$\begin{aligned} \frac{\partial F}{\partial x_i}(x) &= \int_a^b \frac{\partial f}{\partial x_i}(x, t) \, dt \\ \frac{\partial}{\partial x_i} \int_a^b f(x, t) \, dt &= \int_a^b \frac{\partial}{\partial x_i} f(x, t) \, dt \end{aligned}$$

Beweis. Sei $x \in U$ und $e_i = (0, \dots, \underbrace{1}_{i\text{te Stelle}}, \dots, 0)$

$$\begin{aligned} \frac{\partial F}{\partial x_i} &= \lim_{\varepsilon \rightarrow 0} \frac{F(x + \varepsilon e_i) - F(x)}{\varepsilon} \\ &= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \left\{ \int_a^b f(x + \varepsilon e_i, t) \, dt - \int_a^b f(x, t) \, dt \right\} \\ &= \lim_{\varepsilon \rightarrow 0} \int_a^b \frac{f(x + \varepsilon e_i, t) - f(x, t)}{\varepsilon} \, dt \\ &= \int_a^b \frac{\partial f}{\partial x_i}(x, t) \, dt \iff \\ &\iff \lim_{\varepsilon \rightarrow 0} \left\{ \int_a^b \frac{f(x + \varepsilon e_i, t) - f(x, t)}{\varepsilon} \, dt - \int_a^b \frac{\partial f}{\partial x_i}(x, t) \, dt \right\} = 0 \\ &\iff \lim_{\varepsilon \rightarrow 0} \left\{ \int_a^b \left[\frac{f(x + \varepsilon e_i, t) - f(x, t)}{\varepsilon} - \frac{\partial f}{\partial x_i}(x, t) \right] \, dt \right\} = 0 \end{aligned}$$

Wir behaupten mehr, d.h.

$$\int_a^b \left| \underbrace{\frac{f(x + \varepsilon e_i, t) - f(x, t)}{\varepsilon}}_{\frac{\partial f}{\partial x_i}(\xi_\varepsilon, t)} - \frac{\partial f}{\partial x_i}(x, t) \right| \, dt \xrightarrow{\varepsilon \rightarrow 0} 0$$

wobei $\xi_\varepsilon(t) \in [x, x + \varepsilon e_i]$

$$\begin{aligned} \int_a^b |\dots| &= \int_a^b \left| \frac{\partial f}{\partial x_i}(\xi_\varepsilon(t), t) - \frac{\partial f}{\partial x_i}(x, t) \right| \, dt \\ \lim_{\varepsilon \rightarrow 0} \xi_\varepsilon(t) &= x \end{aligned}$$

und (wegen der Stetigkeit von $\frac{\partial f}{\partial x_i}$)

$$\frac{\partial f}{\partial x_i}(\xi_\varepsilon(t), t) \rightarrow \frac{\partial f}{\partial x_i}(x, t)$$

Behauptung 1.118. $\forall \varepsilon > 0 \exists \varepsilon_0 > 0$ so dass

$$\begin{aligned} |\varepsilon| \leq \varepsilon_0 &\implies \sup_{t \in [a, b]} \left| \frac{\partial f}{\partial x_i}(\xi(t), t) - \frac{\partial f}{\partial x_i}(x, t) \right| < \delta \\ \implies & \\ \limsup_{\varepsilon \rightarrow 0} A(\varepsilon) &\leq \sup_{|\varepsilon| < \varepsilon_0} A(\varepsilon) \\ &\leq \int_a^b \delta \, dt = \delta(b-a) \end{aligned}$$

δ ist beliebig

$$\lim_{\varepsilon \rightarrow 0} A(\varepsilon) = 0$$

Lemma 1.119. Sei $g : U \times [a, b] \rightarrow \mathbb{R}$ stetig (wobei $U \subset \mathbb{R}^n$ offen ist). Sei $x \in U$. Dann $\forall \delta > 0 \exists \varepsilon > 0$ mit

$$\sup_{y \in B_\varepsilon(x)} |g(y, t) - g(x, t)| < \delta \quad \forall t \in [a, b]$$

Betrachte x als "Parameter" $\forall y$ sei $t \mapsto g(y, t) = g_y(t)$. Dann $g_y \rightarrow g_x$ gleichmässig für $x \rightarrow x$.

Bemerkung 1.120. Das Lemma nutzt nur die Kompaktheit von $[a, b]$ (in der Behauptung können wir $[a, b]$ durch eine beliebige kompakte Menge $K \subset \mathbb{R}$ ersetzen)

Beweis. Sei $\varepsilon > 0$ gegeben $\forall (x, t) \exists \delta(x, t) > 0$ so dass

$$|g(\xi, \tau) - g(x, t)| < \frac{\varepsilon}{10} \quad \forall (\xi, \tau)$$

mit

$$\begin{aligned} \left\| \underbrace{(\xi, \tau)}_{\in \mathbb{R}^n} - \underbrace{(x, t)}_{\in \mathbb{R}^n} \right\| &< \delta(x, t) \\ \|(\xi, \tau) - (x, t) \| &= \sqrt{\| \xi - x \|^2 + (t - \tau)^2} \end{aligned}$$

Nun $\forall t \in [a, b]$

$$\{B_{\delta(x, t)}(x, t) : t \in [a, b]\}$$

ist eine Überdeckung von

$$K = \{(x, t) : t \in [a, b]\}$$

kompakt weil

$$[a, b] \ni t \mapsto (x, t)$$

stetig von $[a, b]$ nach $\mathbb{R}^n \times \mathbb{R}$. K ist das Bild von $[a, b]$ durch diese Abbildung. \square

	K kompakt	\implies	$f(K)$ kompakt
<i>Bemerkung 1.121.</i>	f stetig	\implies	$f^{-1}(A)$ abgeschlossen
	A abgeschlossen	\implies	$f^{-1}(A)$ abgeschlossen
	O offen	\implies	$f^{-1}(A)$ offen

Alle anderen Implikationen stimmen NICHT.

$\forall (x, t)$ Sei

$$U_{x, t} = \underbrace{B_{\frac{\sqrt{2}}{2}\delta(x, t)}(x)}_{\subset \mathbb{R}^n} \times \left[t - \frac{\sqrt{2}}{2}\delta(x, t), t + \frac{\sqrt{2}}{2}\delta(x, t) \right]$$

$(y, \tau) \in U_{x, t}$

$$\implies \|y - x\| \leq \frac{\sqrt{2}}{2}\delta(x, t) \quad \text{und} \quad |t - \tau| < \frac{\sqrt{2}}{2}\delta(x, t)$$

$$\begin{aligned}\|(y, t) - (x, \tau)\| &< \sqrt{\frac{1}{2}\delta^2(x, t) + \frac{1}{2}\delta^2(x, t)} = \delta(x, t) \\ \implies (y, t) &\in B_{\delta(x, t)}(x, t) \\ \implies U_{x, t} &\subset B_{\delta(x, t)}(x, t)\end{aligned}$$

$\{U_x, t : t \in [a, b]\}$ ist eine offene Überdeckung von K . Kompaktheit $\implies \exists \{U_{x_i, t_i} : i \in \{1, \dots, N\}\}$ Überdeckung von K . Sei

$$\delta = \min \left\{ \frac{\sqrt{2}}{2} \delta(x_i, t_i) : i \in \{1, \dots, N\} \right\} > 0$$

Sei $t \in [a, b]$, $(x, t) \in U_{x_i, t_i}$ für mindestens ein $i \in \{1, \dots, N\}$. Sei y so dass $y - x < \delta$

$$\begin{aligned}(x, t), (y, t) &\in U_{x_i, t_i} \subset B_{\delta(x_i, t_i)}(x_i, t_i) \\ \implies |g(y, t) - g(x_i, t_i)| &< \frac{\varepsilon}{10}\end{aligned}$$

und

$$\begin{aligned}\implies |g(x, t) - g(x_i, t_i)| &< \frac{\varepsilon}{10} \\ \implies |g(x, t) - g(y, t)| &< \frac{\varepsilon}{5} \\ \implies \sup_{y \in B_\delta(x)} |g(x, t) - g(y, t)| &\leq \frac{\varepsilon}{5} < \varepsilon \quad \forall t \in [a, b]\end{aligned}$$

□

Korollar 1.122. Sei $g : U \times [a, b] \rightarrow \mathbb{R}$ stetig. Dann

$$F(x) = \int_a^b g(x, t) \, dt$$

ist eine stetige Funktion

Beweis. Seien $x \in U$ und $\varepsilon > 0$. Das letzte Lemma $\implies \exists \delta > 0$ so dass

$$|g(x, t) - g(y, t)| < \frac{\varepsilon}{b-a}$$

$\forall t$ und $\forall y, x$ mit $\|y - x\| < \delta$. Deswegen für $\|y - x\| < \delta$

$$\begin{aligned}|F(y) - F(x)| &= \left| \int_a^b (g(x, t) - g(y, t)) \, dt \right| \\ &\leq \int_a^b |g(x, t) - g(y, t)| \, dt \\ &< \int_a^b \frac{\varepsilon}{b-a} \, dt = \varepsilon\end{aligned}$$

□

Bemerkung 1.123. Im Differentiationssatz ist $\frac{\partial f}{\partial x_i}$ eine stetige Funktion. Da

$$\frac{\partial f}{\partial x_i}(x) = \int_a^b \frac{\partial f}{\partial x_i}(x, t) \, dt$$

ist $\frac{\partial F}{\partial x_i}$ stetig.

Bemerkung 1.124. Eine sehr wichtige Konsequenz: Sei $f : \underbrace{U}_{\subset \mathbb{R}^2} \rightarrow \mathbb{R}$ eine stetige Funktion. Sei $\underbrace{[a, b] \times [c, d]}_R \subset U$

$$\begin{aligned} s &\mapsto F(s) = \int_a^b f(t, s) \, dt \\ \int_c^d F(s) \, ds &= \int_c^d \left(\int_a^b f(t, s) \, ds \right) dt \\ \int_c^d \int_a^b f(t, s) \, dt \, ds & \\ t &\mapsto G(t) = \int_c^d f(t, s) \, ds \\ \int_a^b G(t) \, dt &= \int_a^b \int_c^d f(t, s) \, ds \, dt \end{aligned}$$

Satz 1.125. f stetig \implies

$$\int_a^b \int_c^d f(s, t) \, ds \, dt = \int_c^d \int_a^b f(s, t) \, dt \, ds$$

Beweis. $(x, y) \in [a, b] \times [c, d]$

$$\begin{aligned} F(x, y) &= \int_a^x \int_c^y f(s, t) \, ds \, dt \\ G(x, y) &= \int_c^y \int_a^x f(t, s) \, dt \, ds \end{aligned}$$

□

Satz 1.126. Sei $f : \underbrace{U}_{\mathbb{R}^2} \rightarrow \mathbb{R}$ eine stetige Funktion. Sei $R = [a, b] \times [c, d] \subset U$.

Dann:

$$\int_a^b \int_c^d f(s, t) \, dt \, ds = \int_c^d \int_a^b f(s, t) \, ds \, dt$$

Beweis. Wir definieren

$$\begin{aligned} \Phi(x, y) &= \int_a^x \int_c^y f(s, t) \, dt \, ds \\ \Psi(x, y) &= \int_c^y \int_a^x f(s, t) \, ds \, dt \end{aligned}$$

Konvention: $\int_\alpha^\beta = -\int_\beta^\alpha$ falls $\beta < \alpha$ und $\int_\alpha^\alpha = 0$

Φ und Ψ sind stetig differenzierbar und $\nabla \Phi = \nabla \Psi$

$\Phi = \Psi$ (Kein Problem mit Definition. Die Funktionen sind wohldefiniert für $(x, y) \in]a - \varepsilon, b + \varepsilon[\times]c - \varepsilon, d + \varepsilon[$ wobei $\varepsilon > 0$ klein genug ist) Sei y fixiert

$$\frac{\partial \Phi}{\partial x}(x, y) = ?$$

$$\phi(x) = \int_c^y f(x, t) \, dt$$

ϕ stetig wegen der letzten Vorlesung. Fundamentalsatz der Int.:

$$\frac{\partial \Phi}{\partial x}(x, y) = \phi(x) = \int_c^y f(x, t) \, dt$$

$\frac{\partial \Phi}{\partial x}$ ist eine stetige Funktion. Sei (x_0, y_0) , $\varepsilon > 0$. Dann (aus der letzten Vorlesung stetig in x) $\exists \delta$

$$\left| \frac{\partial \Psi}{\partial x}(x, y_0) - \frac{\partial \Psi}{\partial x}(x_0, y_0) \right| < \frac{\varepsilon}{2}$$

Sei x fixiert:

$$\begin{aligned} & \left| \frac{\partial \Psi}{\partial x}(x, y_0) - \frac{\partial \Psi}{\partial x}(x, y) \right| \\ &= \left| \int_c^y f(x, t) \, dt - \int_c^{y_0} f(x, t) \, dt \right| \\ &= \left| \int_{y_0}^y f(x, t) \, dt \right| \\ &\leq \int_{y_0}^y |f(x, t)| \, dt \\ &\leq M |y - y_0| \end{aligned}$$

Deswegen für $\bar{\delta} \leq \frac{\varepsilon}{2H}$

$$\begin{aligned} & |y - y_0| < \bar{\delta} \\ \implies & \left| \frac{\partial \Psi}{\partial x}(x, y_0) - \frac{\partial \Psi}{\partial x}(x, y) \right| < \frac{\varepsilon}{2} \end{aligned}$$

Wenn

$$\begin{aligned} & \|(x, y) - (x_0, y_0)\| < \min\{\delta, \bar{\delta}\} \\ \implies & |x - x_0| < \delta \text{ und } |y - y_0| < \bar{\delta} \end{aligned}$$

$$\begin{aligned} & \left| \frac{\partial \Phi}{\partial x}(x, y) - \frac{\partial \Phi}{\partial x}(x_0, y_0) \right| \\ &\leq \left| \frac{\partial \Phi}{\partial x}(x, y) - \frac{\partial \Phi}{\partial x}(x, y_0) \right| + \left| \frac{\partial \Phi}{\partial x}(x, y_0) - \frac{\partial \Phi}{\partial x}(x_0, y_0) \right| < \frac{\varepsilon}{2} \end{aligned}$$

Das gleiche Argument: $\frac{\partial \Psi}{\partial y}$ existiert und ist stetig.

$$\begin{aligned} \psi(x, y) &:= \int_a^x f(s, y) \, ds \\ \frac{\partial \Psi}{\partial x} &= \frac{\partial}{\partial x} \int_c^y \psi(x, t) \, dt \\ &\stackrel{?}{=} \int_c^y \frac{\partial \psi}{\partial x}(x, t) \, dt \end{aligned}$$

Wir brauchen hier die Stetigkeit von ψ . Das haben wir mit dem letzten Argument!

$$\begin{aligned} \frac{\partial \psi}{\partial x}(x, t) &= \frac{\partial}{\partial x} \int_a^x f(s, t) \, ds \stackrel{\text{Fundamentalsatz}}{=} f(x, t) \\ \frac{\partial \Psi}{\partial x} &= \int_c^y f(x, t) \, dt \stackrel{!}{=} \frac{\partial \Phi}{\partial x} \end{aligned} \quad (21)$$

Das gleiche Argument $\frac{\partial \Psi}{\partial x} = \frac{\partial \Phi}{\partial x}$ sind stetig. Sei $\alpha := \Phi - \psi \implies \alpha$ ist differenzierbar und $d\alpha = 0$

$$= [a - \varepsilon, b + \varepsilon] \times [c - \varepsilon, d + \varepsilon]$$

$$|\alpha(x_0, y_0) - \alpha(x_1, y_1)| \leq \|(x_1, y_1) - (x_0, y_0)\| \max \|\nabla \alpha\| = 0$$

Schranksatz? da $[(x_0, y_0)(x_1, y_1)]$ ist im Definitionsbereich

$$\Phi - \Psi = \alpha = \text{konstant} = \Phi(a, c) - \Psi(a, c) = 0 - 0 = 0$$

$$\implies \Phi(x, y) = \Psi(x, y) \quad \forall (x, y) \in [a - \varepsilon, b + \varepsilon] \times [c - \varepsilon, d + \varepsilon]$$

$y = d, x = b \implies$ den Satz. □

2 Differenzierbare Abbildungen

$$f : \underbrace{\mathbb{R}^n}_{\text{Vektorraum}} \rightarrow \mathbb{R}^m$$

Definition 2.1. f ist in x_0 differenzierbar falls $\exists L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ lineare Abbildung:

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0) - L(h)}{\|h\|} = 0$$

d.h. wenn

$$R(h) := f(x_0 + h) - f(x_0) - L(h)$$

dann

$$\lim_{h \rightarrow 0} \frac{\|R(h)\|}{\|h\|} = 0$$

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \text{so dass} \quad 0 < \|h\| < \delta \implies \frac{\|R(h)\|}{\|h\|} < \varepsilon$$

oder auch “ $\|R(h)\| \rightarrow 0$ schneller als $\|h\|$ ” (in “klein-o-Notation”: $R(h) = o(\|h\|)$) Deswegen

$$f \text{ diff in } x_0 \iff \exists L \text{ mit } f(x_0 + h) = f(x_0) + L(h) + o(\|h\|) \quad (22)$$

Bemerkung 2.2. f differenzierbar in $x_0 \implies$ stetig in x_0

f differenzierbar in $x_0 \implies \exists!$ lineare Abbildung die 22 erfüllt. Wir nennen L das Differential von f . $df|_{x_0}$

Bemerkung 2.3. $f : U \rightarrow \mathbb{R}^m$

$$f(x) = \underbrace{(f_1(x), \dots, f_m(x))}_{m \text{ Funktionen}}$$

$\forall i \frac{\partial f_i}{\partial x_j}$ n partielle Ableitungen

$$L : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

$$L = \begin{pmatrix} L_{11} & \cdots & L_{1n} \\ L_{21} & \cdots & L_{2n} \\ \vdots & & \vdots \\ L_{m1} & \cdots & L_{mn} \end{pmatrix} = \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_m \end{pmatrix}$$

$$L(x) = \begin{pmatrix} L_{11} + L_{12} + \cdots + L_{1n}x_n \\ L_{21} + \cdots + L_{2n}x_n \\ \vdots \\ L_{m1} + \cdots + L_{mn}x_n \end{pmatrix} = \begin{pmatrix} L_1 x \\ L_2 x \\ \vdots \\ L_m x \end{pmatrix}$$

$\exists m$ lineare Abbildungen $\mathbb{L} : \mathbb{R}^n \rightarrow \mathbb{R}$

$$L(x) = \begin{pmatrix} \mathbb{L}_1(x) \\ \mathbb{L}_2(x) \\ \vdots \\ \mathbb{L}_m(x) \end{pmatrix}$$

$$\mathbb{L}_i(x) = L_i x$$

Bemerkung 2.4. Sei $f : U \rightarrow \mathbb{R}^m$ differenzierbar in x_0 und sei $L = df|_{x_0}$. Dann:

$$\overbrace{\frac{f(x_0 + h) - f(x_0) - L(h)}{\|h\|}}^A \rightarrow 0 \quad (23)$$

$$A := \begin{pmatrix} f_1(x_0 + h) \\ \vdots \\ f_m(x_0 + h) \end{pmatrix} - \begin{pmatrix} f_1(x_0) \\ \vdots \\ f_m(x_0) \end{pmatrix} - \begin{pmatrix} \mathbb{L}_1(h) \\ \vdots \\ \mathbb{L}_m(h) \end{pmatrix}$$

$$= \begin{pmatrix} f_1(x_0 + h) - f_1(x_0) - \mathbb{L}_1(h) \\ \vdots \\ f_m(x_0 + h) - f_m(x_0) - \mathbb{L}_m(h) \end{pmatrix}$$

$$\frac{A}{\|h\|} = \begin{pmatrix} \frac{f_1(x_0 + h) - f_1(x_0) - \mathbb{L}_1(h)}{\|h\|} \\ \vdots \\ \frac{f_m(x_0 + h) - f_m(x_0) - \mathbb{L}_m(h)}{\|h\|} \end{pmatrix}$$

Deswegen

$$23 \iff \lim_{h \rightarrow 0} \frac{f_i(x_0 + h) - f_i(x_0) - \mathbb{L}_i(h)}{\|h\|} = 0 \quad \forall i \in \{1, \dots, m\}$$

$$\iff f_i \text{ ist differenzierbar in } x_0 \text{ und } \mathbb{L}_i = d f_i|_{x_0}$$

Satz 2.5. Sei $f : \underbrace{U}_{\subset \mathbb{R}^n} \rightarrow \mathbb{R}^m$ mit U offen und $f = (f_1, \dots, f_m)$

1. f ist differenzierbar in $x_0 \iff f_i$ differenzierbar in $x_0 \quad \forall i \in \{1, \dots, m\}$

2.

$$d f|_{x_0}(h) = \begin{pmatrix} d f_1|_{x_0}(h) \\ \vdots \\ d f_m|_{x_0}(h) \end{pmatrix}$$

3.

$$d f|_{x_0}(h) = \begin{pmatrix} \nabla f_1 + (x_0)h \\ \vdots \\ \nabla f_m + (x_0)h \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \frac{\partial f_i}{\partial x_j} & \vdots \\ \frac{\partial f_m}{\partial x_1} & \dots & \dots & \frac{\partial f_m}{\partial x_n} \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}$$

Das ist die Jacobi Matrix.

Bemerkung 2.6. $f, g : U \rightarrow \mathbb{R}^m$ beide differenzierbar in x_0 , dann

$$f + g \left(= \begin{pmatrix} f_1 + g_1 \\ \vdots \\ f_m + g_m \end{pmatrix} \right)$$

ist differenzierbar in x_0 und $d f|_{x_0} + d g|_{x_0}$

Bemerkung 2.7. $f : U \rightarrow \mathbb{R}^m$ und $g : U \rightarrow \mathbb{R}^m$ differenzierbar in x_0

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

$$\stackrel{?}{d(gf)} = d \begin{pmatrix} g f_1 \\ \vdots \\ g f_m \end{pmatrix} \Big|_{x_0} (h) = \begin{pmatrix} d(g f_1)|_{x_0}(h) \\ \vdots \\ d(g f_m)|_{x_0}(h) \end{pmatrix}$$

$$= \begin{pmatrix} d g|_{x_0}(h) f_1(x_0) + g(x_0) d f_1|_{x_0}(h) \\ \vdots \\ d g|_{x_0}(h) f_m(x_0) + g(x_0) d f_m|_{x_0}(h) \end{pmatrix}$$

Jacobi-Matrix

$$\begin{aligned}
 & \begin{pmatrix} \frac{\partial g}{\partial x_1}(x_0)f_1(x_0) + g(x_0)\frac{\partial f_1}{\partial x_1}(x_0) & \cdots & \frac{\partial g}{\partial x_n}(x_0)f_1(x_0) + g(x_0)\frac{\partial f_1}{\partial x_n}(x_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial g}{\partial x_1}(x_0)f_m(x_0) + g(x_0)\frac{\partial f_m}{\partial x_1}(x_0) & \cdots & \frac{\partial g}{\partial x_n}(x_0)f_m(x_0) + g(x_0)\frac{\partial f_m}{\partial x_n}(x_0) \end{pmatrix} \\
 &= \begin{pmatrix} \frac{\partial g}{\partial x_1}(x_0)f_1(x_0) & \cdots & \frac{\partial g}{\partial x_n}(x_0)f_1(x_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial g}{\partial x_1}(x_0)f_m(x_0) & \cdots & \frac{\partial g}{\partial x_n}(x_0)f_m(x_0) \end{pmatrix} \\
 &+ \begin{pmatrix} g(x_0)\frac{\partial f_1}{\partial x_1}(x_0) & \cdots & g(x_0)\frac{\partial f_1}{\partial x_n}(x_0) \\ \vdots & \ddots & \vdots \\ g(x_0)\frac{\partial f_m}{\partial x_1}(x_0) & \cdots & g(x_0)\frac{\partial f_m}{\partial x_n}(x_0) \end{pmatrix} \\
 &= \underbrace{\begin{pmatrix} \frac{\partial g}{\partial x_j}(x_0)f_i(x_0) \end{pmatrix}}_{=f(x_0) \otimes \nabla g(x_0)} + \overbrace{\begin{pmatrix} \frac{\partial f_i}{\partial x_j}(x_0) \end{pmatrix}}^A \\
 &\quad \text{lineare Abbildung mit Rang 1} \\
 d(gf)|_{x_0} &= \underbrace{g(x_0) df|_{x_0}}_A + \overbrace{f(x_0) \otimes dg|_{x_0}}^{\text{lineare Abbildung mit Rang 1}} \\
 d(gf)|_{x_0}(h) &= g(x_0) [df|_{x_0}(h)] + [f(x_0)] dg|_{x_0}(h)
 \end{aligned}$$

Satz 2.8.

$$\begin{aligned}
 f &: \underbrace{U}_{\subset \mathbb{R}^n} \rightarrow \underbrace{V}_{\subset \mathbb{R}^n} \\
 g &: V \rightarrow \mathbb{R}^k
 \end{aligned}$$

Falls f in a differenzierbar ist und g in $b = f(a)$ differenzierbar ist, dann ist $g \circ f$ in a differenzierbar und

$$d(g \circ f)|_a \stackrel{?}{=} dg|_b \circ df|_a \quad (24)$$

Beweis. Differential von f in a mit $\frac{R(h)}{\|h\|} \rightarrow 0$

$$f(a+h) = f(a) + df|_a(h) + \overbrace{R(h)}^{o(\|h\|)}$$

Differential von g in b mit $\frac{\bar{R}(k)}{\|k\|} \rightarrow 0$

$$g(b+k) = g(b) + dg|_b(k) + \underbrace{\bar{R}(k)}_{o(\|k\|)}$$

$$g(f(a+h)) = g(\underbrace{f(a)}_b + k) = g(b) + dg|_b(k) + \bar{R}(k)$$

Linearität von $dg|_b$

$$\begin{aligned}
 &= g(b) + dg|_b(df|_a(h) + R(h)) + \bar{R}(k) \\
 &= \underbrace{g(b)}_{g \circ f(a)} + \underbrace{dg|_b(df|_a(h))}_{\text{ist linear in } h} + \underbrace{dg|_b(R(h)) + \bar{R}(k)}_{:=\rho(h)} \\
 &\quad \rho(h) = o(\|h\|)
 \end{aligned}$$

□

Lemma 2.9. *Linearität:*

$$\begin{aligned}
 dg|_b \circ df|_a(\lambda_1 h_1 + \lambda_2 h_2) & \quad \lambda_1, \lambda_2 \in \mathbb{R}, h_1, h_2 \in \mathbb{R}^n \\
 &= dg|_b (df|_a(\lambda_1 h_1 + \lambda_2 h_2)) \\
 &= dg|_b \left(\lambda_1 \overbrace{df|_a(h_1)}^{\in \mathbb{R}^m} + \lambda_2 \overbrace{df|_a(h_2)}^{\in \mathbb{R}^m} \right) \\
 &= \lambda_1 dg|_b(df|_a(h_1)) + \lambda_2 dg|_b(df|_a(h_2)) \\
 &= \lambda_1 dg|_b \circ df|_a(h_1) + \lambda_2 dg|_b \circ df|_a(h_2)
 \end{aligned}$$

Lemma 2.10.

$$\begin{aligned}
 \frac{\rho(h)}{\|h\|} &\leq \frac{|dg|_b(R(h))|}{\|h\|} + \frac{|\bar{R}(j)|}{\|h\|} \\
 &\leq \frac{\|dg|_b\|_0 \|R(h)\|}{\|h\|} + \frac{\|\bar{R}(h)\|}{\|h\|} \\
 &\quad \frac{\|R(h)\|}{\|h\|} \rightarrow 0 \\
 \frac{\|\bar{R}(h)\|}{\|h\|} &= \begin{cases} 0 & \text{falls } k = 0 \\ \frac{\|\bar{R}(k)\|}{\|k\|} \frac{\|k\|}{\|h\|} & \end{cases} \\
 \|k\| &= \|df|_a(h) + R(h)\| \leq \|df|_a(h)\| + \|R(h)\| \\
 &\leq \|df|_a\|_0 \|h\| + \|R(h)\| \\
 &\quad \frac{\|R(h)\|}{\|h\|} \rightarrow 0
 \end{aligned} \tag{25}$$

Sei $\varepsilon = 1$, $\exists \delta > 0$ so dass

$$\|h\| < \delta \implies \frac{\|R(h)\|}{\|h\|}$$

Falls $\|h\| < \delta$

$$25 \leq (\|df|_a\|_0 + 1) \|h\|$$

Deswegen: wenn $\|h\| \rightarrow 0$, dann $\|k\| \rightarrow 0$ und für $\|h\| < \delta$

$$\frac{\|\bar{R}(k)\|}{\|h\|} \leq \underbrace{\frac{\|\bar{R}(k)\|}{\|k\|}}_{\rightarrow 0 \text{ für } \|h\| \rightarrow 0} (\|df|_a\|_0 + 1)$$

Deswegen:

$$\begin{aligned}
 0 &\leq \limsup_{\|h\| \rightarrow 0} \frac{\|\rho(h)\|}{\|h\|} \\
 &\leq \lim_{\|h\| \rightarrow 0} \frac{\|\bar{R}(k)\|}{\|h\|} + \lim_{\|h\| \rightarrow 0} \frac{\|dg|_b(R(h))\|}{\|h\|} = 0 + 0 = 0 \\
 &\implies \lim_{\|h\| \rightarrow 0} \frac{\|\rho(h)\|}{\|h\|} = 0
 \end{aligned}$$

Bemerkung 2.11. $n = m = k = 1$

$$f(a+h) = f(a) + \underbrace{df|_a}_{\text{Differential}}(h) + o(\|h\|)$$

$b = f(a)$

$$df|_a(h) = f'(a)h$$

$$dg|_b(k) = g'(b)k$$

$$dg|_b = df|_a(k) = dg|_b(df|_a(h)) = dg|_b(f'(a)h)$$

$$= g'(b)f'(a)h = g'(f(a))f'(a)h \quad (26)$$

$$\phi = g \circ f$$

$$d\phi|_a(h) = \phi'(a)h = (g \circ f)'(a)h$$

24 d.h. die allgemeine Kettenregel

$$\begin{aligned} d\phi|_a(h) &= d(g \circ f)|_a(h) \\ &= dg|_b \circ df|_a(h) \stackrel{26}{=} g'(f(a))f'(a)h \\ \implies (g \circ f)'(a)h &= g'(f(a))f'(a)h \\ \implies \underbrace{(g \circ f)'(a)}_{\text{alte Kettenregel}} &= g'(f(a))f'(a) \end{aligned}$$

Bemerkung 2.12. Kettenregel für die Jacobi-Matrizen: Formel 24 \iff Sei M die Jacobi-Matrix für $dg|_{b(=f(a))}$ und N die für $df|_a$. Dann ist die Jacobi für $d(g \circ f)|_a$ ist MN

$\implies g = (g_1, \dots, g_k) \quad f = (f_1, \dots, f_m)$ Es gibt eine Formel für $\frac{\partial(g \circ f)_i}{\partial x_j}$

$$dg_b \circ df|_a(w) = dg_b(\underbrace{df|_a(w)}_v)$$

$$\begin{aligned} dg_b \circ df|_a(w) &= dg_b(v) \\ &= \left(\sum_{i=1}^m M_{1i}v_i, \sum_{i=1}^m M_{2i}v_i, \dots, \sum_{i=1}^m M_{ki}v_i \right) \\ &= \left(\sum_{i=1}^m M_{1i} \sum_{j=1}^n N_{ij}w_j, \dots, \sum_{i=1}^m M_{ki} \sum_{j=1}^n N_{ij}w_j \right) \end{aligned}$$

$$\begin{aligned} v = df|_a(w) &= \left(\sum_{j=1}^n N_{1j}w_j, \dots, \sum_{j=1}^n N_{mj}w_j \right) \\ \iff v_i &= \sum_{j=1}^n N_{ij}w_j \end{aligned}$$

$$dg_b \circ df|_a(w) = \left(\sum_{i=1}^m \sum_{j=1}^n M_{1i}N_{ij}w_j, \dots, \sum_{i=1}^m \sum_{j=1}^n M_{ki}N_{ij}w_j \right)$$

(Sei A die Matrix

$$\begin{aligned} A_{lj} &= \sum_{i=1}^m M_{li}N_{ij} \iff A = M \cdot N \\ &= \left(\sum_{j=1}^n A_{1j}w_j, \dots, \sum_{j=1}^n A_{kj}w_j \right) \end{aligned}$$

Deswegen ist A die Matrixdarstellung von

$$dg_b \circ df|_a = d(g \circ f)|_a$$

$\iff A$ ist die Jacobi-Matrix für $d(g \circ f)|_a$

Bemerkung 2.13. $f : U \rightarrow V \subset \mathbb{R}^m$ $f = (f_1, \dots, f_m)$, $f_i(x) = f(x_1, \dots, x_n)$
 $g : V \rightarrow \mathbb{R}^k$ $g = (g_1, \dots, g_k)$, $g_j(x) = g(y_1, \dots, y_m)$

$$\begin{aligned} g \circ f(x) &= (g_1(f(x)), \dots, g_k(f(x))) \\ g_j(x) &= g_j(f_1(x), \dots, f_m(x)) \\ g_j(y) &= g_j(f(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)) \end{aligned}$$

$$\begin{aligned} A_{lj} &= \frac{\partial}{\partial x_j}(g_l \circ f)(a) \\ M_{li} &= \frac{\partial g_l}{\partial y_i}(b) = \frac{\partial g_l}{\partial y_i}(f(a)) \\ N_{ij} &= \frac{\partial f_i}{\partial x_j}(a) \end{aligned}$$

$$\begin{aligned} \frac{\partial}{\partial x_j}(g_l \circ f)(a) &= A_{lj} = \sum_{i=1}^m M_{li} N_{ij} \\ &= \sum_{i=1}^m \frac{\partial g_l}{\partial y_i}(f(a)) \frac{\partial f_i}{\partial x_j}(a) \end{aligned}$$

Korollar 2.14. Sei $f : U \rightarrow V \subset \mathbb{R}^m$ und $\phi : V \rightarrow \mathbb{R}$ mit:

- $a \in U$ und U offen
- $b \in V$, V offen und $b = f(a)$
- f differenzierbar in a und ϕ differenzierbar in b

Dann ist $\phi \circ f$ differenzierbar in a und

$$\frac{\partial \phi \circ f}{\partial x_j}(a) = \sum_{i=1}^m \frac{\partial \phi}{\partial y_i}(f(a)) \frac{\partial f_i}{\partial x_j}(a)$$

Das ist die "konkrete" allgemeine Kettenregel.

Satz 2.15. Sei $\Omega \subset \mathbb{R}^n$ eine offene Menge, $f : \Omega \rightarrow \mathbb{R}^k$ (eine differenzierbare Abbildung mit stetigen partiellen Ableitungen) eine \mathbb{C}^1 Funktion und $\gamma[a, b] \rightarrow \Omega$ eine \mathbb{C}^1 Kurve.

Definition 2.16. $f \in \mathbb{C}^k(U, \mathbb{R}^m)$ falls die partielle Ableitungen von f_i mit Ordnung $\leq k$ existieren und stetig sind ($f = (f_1, \dots, f_m)$). Dann:

$$\|f(\gamma(b)) - f(\gamma(a))\| \leq \left[\sup_{t \in [a, b]} \|df|_{\gamma(t)}\|_0 \right] \underbrace{\int_a^b \|\dot{\gamma}(t)\| dt}_{\text{Länge der Kurve}}$$

$$\gamma : [a, b] \rightarrow \Omega \subset \mathbb{R}^n, \gamma = (\gamma_1, \dots, \gamma_n), \dot{\gamma} = (\gamma'_1, \dots, \gamma'_n)$$

Beweis. Sei $\phi : [a, b] \rightarrow \mathbb{R}^k$ die Funktion

$$\phi(t) := f(\gamma(t)) = f \circ \gamma$$

Kettenregel

$$d\phi|_t = df|_{\gamma(t)} d\gamma|_t \tag{27}$$

$$\phi : [a, b] \rightarrow \mathbb{R}^k$$

$d\phi|_t : \mathbb{R} \rightarrow \mathbb{R}^k$ lineare Abbildung

$$\phi = (\phi_1, \dots, \phi_k)$$

$$\begin{pmatrix} \frac{\partial \phi_1}{\partial t} \\ \vdots \\ \frac{\partial \phi_k}{\partial t} \end{pmatrix} = \begin{pmatrix} \phi'_1 \\ \vdots \\ \phi'_k \end{pmatrix} = \dot{\phi}$$

Sei A die Jacobi-Matrix für $d f$. Kettenregel:

$$\dot{\phi}(t) = \underbrace{A(\gamma(t)) \cdot \dot{\gamma}(t)}_{\text{Matrix-Darstellung von 27}}$$

$$f(\gamma(b)) - f(\gamma(a)) = \phi(b) - \phi(a) = \begin{pmatrix} \phi_1(b) - \phi_1(a) \\ \vdots \\ \phi_k(b) - \phi_k(a) \end{pmatrix}$$

ϕ'_i ist eine stetige Funktion:

$$\begin{aligned} \phi'_i(t) &= \sum_{j=1}^n A_{ij}(\gamma(t)) \gamma'_j(t) \\ &= \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(\gamma(t)) \gamma'_j(t) \\ \phi(b) - \phi(a) &= \begin{pmatrix} \int_a^b \phi'_1(t) dt \\ \vdots \\ \int_a^b \phi'_k(t) dt \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \|f(\gamma(b)) - f(\gamma(a))\|^2 &= \|\phi(b) - \phi(a)\|^2 \\ &= \sum_{i=1}^k \left(\int_a^b \phi'_i(t) dt \right)^2 \\ &\stackrel{\text{Dreiecksungleichung}}{\leq} \int_a^b \end{aligned}$$

Lemma 2.17. Dreiecksungleichung Seien $f, g : [a, b] \rightarrow \mathbb{R}$ zwei stetige Funktionen. Dann:

$$\left| \int_a^b f g dt \right|^2 \leq \int_a^b f^2 \int_a^b g^2$$

Beweis.

$$\begin{aligned} \int_a^b \left(t f - \frac{1}{t} g \right)^2 &\geq 0 \quad \forall t \in \mathbb{R} \setminus \{0\} \\ \int_A t^2 f^2 + \int_A \frac{1}{t^2} g^2 - \int_a^b 2 f g &\geq 0 \\ \int_A f g &\leq \underbrace{\frac{t^2}{2} \int_a^b f^2 + \frac{1}{2t^2} \int_A g^2}_{\Phi(t) = \frac{t^2}{2} \alpha + \frac{1}{2t^2} \beta} \quad \alpha, \beta > 0, \Phi'' > 0 \end{aligned}$$

$$\Phi'(t) = t\alpha - \frac{\beta}{t^3}$$

$$\Phi'(t) = 0 \iff t^4 = \frac{\beta}{\alpha} \iff t^2 = \sqrt{\frac{\beta}{\alpha}}$$

Wähle

$$t^2 = \sqrt{\frac{\beta}{\alpha}} = \frac{1}{2} \sqrt{\frac{\beta}{\alpha}} \alpha + \frac{\beta}{\sqrt{\frac{\beta}{\alpha}}} = \frac{1}{2} \sqrt{\beta \alpha} + \frac{1}{2} \sqrt{\beta \alpha}$$

□

$$\implies \int_a^b fg \leq \int_a^b f^2 \int_a^b g^2 \quad (= \sqrt{\beta\alpha}) \quad (28)$$

Zur Erinnerung: $\alpha, \beta > 0 \implies 28$

$$\begin{aligned} \alpha = 0 &\implies \int_A f^2 = 0 \implies f^2 \equiv 0 \\ &\implies f \equiv 0 \implies ?? \text{ ist trivial} \\ &\implies 28 \text{ gilt ohne Einschränkung} \end{aligned}$$

□

Korollar 2.18. Wenn wir das gleiche Argument mit $-f$ und g anwenden, dann:

$$\int_a^b (-f)g \leq \sqrt{\int_a^b (f)^2 \int_a^b g^2} \quad (29)$$

$$\begin{aligned} \xRightarrow{28+29} \left| \int_a^b fg \right| &\leq \sqrt{\int_a^b f^2} \sqrt{\int_a^b g^2} \\ &\implies \left(\int_a^b fg \right)^2 \leq \int_a^b f^2 \int_a^b g^2 \end{aligned}$$

Lemma 2.19. Dreiecksungleichung (2te Version) $f, g : [a, b] \rightarrow \mathbb{R}^k$ zwei stetige Funktionen. Dann:

$$\left| \sum_{i=1}^k \int_a^b f_i g_i \right| \leq \int_a^b \|f\|^2 \int_a^b \|g\|^2$$

Beweis. Das Argument:

$$\begin{aligned} &\int_a^b \left\| tf - \frac{1}{t}g \right\|^2 \geq 0 \\ &\quad \int_a^b \sum_{i=1}^k \left(tf_i - \frac{1}{t}g_i \right)^2 \\ &= \int_a^b \left(t^2 \sum f_i^2 + \frac{1}{t^2} \sum g_i^2 - 2 \sum f_i g_i \right) \\ &\implies \frac{t^2}{2} \int_a^b \|f\|^2 + \frac{1}{2t^2} \int_a^b \|g\|^2 \\ &\quad \geq \int_a^b \sum f_i g_i \end{aligned}$$

Optimierung von $t \implies$ die Ungleichung

$$\int_a^b \sum f_i g_i \leq \sqrt{\int_a^b \|f\|^2 \int_a^b \|g\|^2}$$

□

$$\begin{aligned} &\|f(\gamma(b)) - f(\gamma(a))\| \\ &= \|\phi(b) - \phi(a)\| \\ &= \sqrt{\sum \left(\int_a^b \phi'(s) \, ds \right)^2} \\ &\stackrel{\text{Dreiecksungleichung}}{\leq} \int_a^b \overbrace{\sqrt{\sum (\phi'(s))^2}}^{\|\dot{\phi}\|} \, ds \end{aligned}$$

Lemma 2.20. Sei $g : [a, b] \rightarrow \mathbb{R}^k$ eine stetige Funktion.

$$\sqrt{\sum \left(\int_a^b g_i \right)^2} \leq \int_a^b \|g\|$$

Dreiecksungleichung

Beweis. Sei $\varepsilon > 0$ und Treppenfunktion α_i so dass $g_i - \varepsilon \leq \alpha_i \leq g_i + \varepsilon$,
 $\alpha_i - \varepsilon \leq g_i \leq \alpha_i + \varepsilon$.

$$\begin{aligned} & \int_a^b \alpha_i - (b-a)\varepsilon \\ & \leq \int_a^b g_i \\ & \leq \int_a^b \alpha_i + (b-a)\varepsilon \\ & \left| \int_a^b g_i - \int_a^b \alpha_i \right| \leq (b-a)\varepsilon \\ & \left| \sqrt{\sum \left(\int_a^b g_i \right)^2} - \sqrt{\sum \left(\int_a^b \alpha_i \right)^2} \right| \\ & \leq \sqrt{\sum \left(\int_a^b g_i - \int_a^b \alpha_i \right)^2} \\ & \leq \sqrt{k(b-a)^2 \varepsilon^2} \\ & \leq \sqrt{k}(b-a)\varepsilon \\ & \left| \int_a^b \|g\| - \int_a^b \|\alpha\| \right| \\ & \leq \int_a^b \left| \|g\| - \|\alpha\| \right| \\ & \leq \int_a^b \|g - \alpha\| \\ & \leq \int_a^b \sqrt{k} \varepsilon \\ & = (b-a)\varepsilon \end{aligned}$$

Behauptung 2.21.

$$\sqrt{\sum \left(\int_a^b \alpha_i \right)^2} \leq \int_a^b \|\alpha\| \tag{30}$$

$$\begin{aligned} 30 \implies \sqrt{\sum \left(\int_a^b g_i \right)^2} & \leq \sqrt{\sum \left(\int_a^b \alpha_i \right)^2} + (b-a)\sqrt{k}\varepsilon \\ & \stackrel{30}{\leq} \int_a^b \|\alpha\| + (b-a)\sqrt{k}\varepsilon \\ & \leq \int_a^b \|g\| + 2(b-a)\sqrt{k}\varepsilon \end{aligned}$$

Wenn $\varepsilon \downarrow 0$:

$$\sqrt{\sum \left(\int_a^b g_i \right)^2} \leq \int_a^b \|g\|$$

Ohne Beschränkung der Allgemeinheit: \exists eine Zerteilung von $[a, b]$

$$a = c_0 < c_1 < \dots < c_N = b$$

jedes α_i ist konstant auf $[c_{j-1}, c_j] = I_j$. Die Konstante ist $a_{i,j}$

$$\alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix}$$

ist konstant auf I_j mit Wert

$$a_j = \begin{pmatrix} a_{1,j} \\ \vdots \\ a_{k,j} \end{pmatrix}$$

$$\sqrt{\sum_{i=1}^k \left(\int_a^b \alpha_i \right)^2} = \sqrt{\sum_{i=1}^k \left(\sum_{j=1}^N |I_j| \alpha_{i,j} \right)^2} = \|a\|$$

$$a := \sum_{j=1}^N |I_j| a_j$$

$$= \begin{pmatrix} \sum_{j=1}^N |I_j| \alpha_{1,j} \\ \vdots \\ \sum_{j=1}^N |I_j| \alpha_{k,j} \end{pmatrix}$$

$$= \left\| \sum_{j=1}^N |I_j| a_j \right\|$$

$$\stackrel{\text{Dreiecksungleichung}}{\leq} \sum_{j=1}^N \| |I_j| a_j \|$$

$$= \sum_{j=1}^N |I_j| \|a_j\|$$

$$= \int_a^b \|\alpha\|$$

$\|\alpha\|$ hat den Wert $\sqrt{\sum_{i=1}^k \alpha_i^2}$ und deswegen auf I_j ist dieser Wert

$$\sqrt{\sum_{i=1}^k a_{i,j}^2} = \|a_j\| \equiv \|\alpha\|$$

□

Beweis. vom Schrankensatz. $\phi = f \circ \gamma$, $\phi(t) = f(\gamma(t))$

$$\begin{aligned}
 & f(\gamma(b)) - f(\gamma(a)) \\
 &= \phi(b) - \phi(a) \\
 &= \begin{pmatrix} \int_a^b \phi'_1(s) \, ds \\ \vdots \\ \int_a^b \phi'_k(s) \, ds \end{pmatrix} \\
 &\xRightarrow{\text{Dreiecksungleichung}} \|f(\gamma(b)) - f(\gamma(a))\| \\
 &\leq \int_a^b \|\dot{\phi}(s)\| \, ds \\
 &\stackrel{\text{Kettenregel}}{=} \int_a^b \|\mathrm{d}f|_{\gamma(s)}(\dot{\gamma}(s))\| \, ds \\
 &\leq \int_a^b \|\mathrm{d}f|_{\gamma(s)}\|_O \|\dot{\gamma}(s)\| \, ds \\
 &\leq \max_{s \in [a,b]} \|\mathrm{d}f|_{\gamma(s)}\|_O \int_a^b \|\dot{\gamma}(s)\| \, ds
 \end{aligned}$$

Vorsicht: $\|\cdot\|_O$ ist eine Norm, deswegen stetig!

□

Korollar 2.22. $f : \Omega \rightarrow \mathbb{R}^k \subset \mathbb{C}^1$ Funktion, $[p, q] \subset \Omega$. Dann:

$$\|f(p) - f(q)\| \leq \max_{z \in [p,q]} \|\mathrm{d}f|_z\| \|p - q\|$$

Beweis. Wenden den Satz an f und $\gamma : [0, 1] \rightarrow \Omega$ ist $\gamma(s) = (1-s)p + sq$, $\dot{\gamma} = q - p$

$$\begin{aligned}
 & \|f(p) - f(q)\| \leq \\
 & \max_{s \in [p,q]} \|\mathrm{d}f|_{\gamma(s)}\| \underbrace{\int_0^1 \|\dot{\gamma}(s)\| \, ds}_{\|p-q\|}
 \end{aligned}$$

□