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 4
1.2 Ein bisschen mehr Topologie 5
1.3 Stetigkeit 7
1.4 Lineare Abbildungen 8
1.5 Mehr über stetige Funktionen 11
1.6 Kompakte Menge 13

2 Differenzierbare Funktionen 17
2.1 Zusammenfassung 20
2.1.1 Das Differenzial 20
2.1.2 Richtungsableitung 20
2.1.3 Partielle Ableitung 20
2.2 Das Hauptkriterium der Differenzierbarkeit 20
2.3 Die geometrische Bedeutung des Gradients 22
2.4 Rechenregeln 22
2.5 Kettenregel 23
2.6 Mittelwertsatz und Schrankensatz 25
2.7 Höhere partielle Ableitungen 26

3 Das Taylorpolynom 29
3.1 Das Taylorpolynom zweiter Ordnung 30
3.2 Konvexität 34

4 Differentation parameterabhängiger Integrale 34

5 Differenzierbare Abbildungen 39
5.1 Differentiationregeln 40
5.2 Kettenregel 41
5.3 Schrankensatz 44
5.4 Satz der lokalen Umkehrbarkeit 47
5.4.1 Allgemeine Form des Fixpunktsatzes von Banach 47

1 Metrik und Topologie des euklidischen Raumes

$\mathbb{R}^n = \{(x_1, \cdots, x_n), x \in \mathbb{R}\}$. Wir führen verschiedene neue Begriffe in $\mathbb{R}^n$ ein:

- die Euklidische Norm
- der Euklidische Abstand
- die entsprechende Topologie.

Wir betrachten gleichzeitig die entsprechenden Verallgemeinerungen, d.h. die "Abstrakte Theorien" der

- Normierten Vektorräume
- Metrischen Räume
- Topologischen Räume.

Definition 1.1. Sei $x \in \mathbb{R}^n$ ($x = (x_1, \cdots, x_n)$, $x_i \in \mathbb{R}$). Die Euklidische Norm von x ist

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

(wir schreiben oft $\|x\|$ anstatt $\|x\|_e$).

Intuitiv: $\|x\|$ ="der Abstand zwischen x und 0". In der Tat, wenn $n = 2$, das Pytaghoras Theorem zeigt dass $\|x\|_e$ die Länge des Segments mit Extrema x and 0 ist.

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

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

Beweis. 1. ≥ 0 trivial

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

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

2.

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

3. Diese Aussage ist äquivalent zu

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

Wir rechnen

$$\sum_{i=1}^n (x_i + y_i)^2 = \sum_{i=1}^n \left(x_i^2 + y_i^2 + 2x_i y_i\right) = \|x\|^2 + \|y\|^2 \;\; \overbrace{2\sum_i x_i y_i}^{Skalarprodukt}$$

Wir definieren

$$\langle x, y \rangle := \sum_{i=1}^n x_i y_i$$

Wir brauchen dann die berühmte Cauchy-Schwartz Ungleichung, d.h.

$$\langle x, y \rangle \leq \|x\| \, \|y\| \; .$$

Diese Ungleichung ist der Inhalt des nächsten Satzes.

□

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)

$$t \to g(t) = \sum_{i=1}^{n} (x_i + t y_i)^2 = \left(\sum x_i^2\right) + 2t \sum x_i y_i + t^2 \sum y_i^2$$

$$= \|x\|^2 + 2t\langle x, y\rangle + \|y\|^2 t^2$$

Sei $t_0 = \frac{\langle x,y\rangle}{\|y\|^2}$, dann

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

$$\implies \langle x, y\rangle^2 \leq \|x\|^2 \|y\|^2 \implies |\langle x, y\rangle| \leq \|x\| \|y\|$$

□

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

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

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

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

$\| \cdot \|_2$ ist die Euklidische Norm.

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

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

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

3. $d(x, z) \leq d(x, y) + d(y, z)$ (Dreiecksungleichung)

Beweis. Die erste Zwei Aussagen sind trivial. Um die letzte zu beweisen:

$$\|x - z\| \leq \| \underbrace{x - y}_{=:v} \| + \| \underbrace{\|y - z\|}_{=:w} \| \,.$$

Aber $x - z = v + w$. Wir wenden die dritte Aussage von Lemma 1.2 an:

$$d(x, z) = \|v + w\| \leq \|v\| + \|w\| = d(x, y) + d(y, z)\,.$$

□

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

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

so dass

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

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

Beweis. Wir nutzen das gleiche Argument vom Lemma 1.7. □

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

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

(Wir werden auch oft $B_r(x)$ statta $K_r(x)$ nutzen.)

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, d.h.

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

Bemerkung 1.13. Die Dreiecksungleichung impliziert dass jede offene Kugel eine offene Menge ist. In der Tat, sei $y \in K_r(x)$. Dann $\rho := d(x,y) < r$. Sei $\tau := r - \rho > 0$. Falls $z \in K_\tau(y)$, dann $d(x,z) \leq d(x,y) + d(y,z) = \rho + d(y,z) < \rho + \tau = r$. D.h., $K_\tau(y) \subset K_r(x)$. Das beweist dass $K_r(x)$ eine Ungebung ihrer ganzen Elementen ist, d.h. $K_r(x)$ ist offen.

Satz 1.14. *1. $\varnothing$ und $\mathbb{R}^n$ sind offen*

2. Der Schnitt <u>endlich vieler</u> offener Mengen ist auch offen.

3. Die Vereinigung einer <u>beliebigen</u> Familie offener Mengen ist auch offen.

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

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

$$\forall i \in \{1, \ldots, N\} \qquad \exists r_i > 0 \text{ so dass } K_{r_i}(x) \subset U_i$$

Sei $r = \min\{r_i, \ldots, r_N\} > 0$;

$$\Longrightarrow K_r(x) \subset U_i \quad \forall i \Longrightarrow K_r(x) \subset U_1 \cap \cdots \cap U_N$$

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

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

$$\Longrightarrow \exists K_r(x) \subset U_\lambda \subset U.$$

□

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

1. $\varnothing, X \in O$
2. $U_1 \cap \cdots \cap_N \in O$ falls $U_i \in O$
3. $\bigcap_{\lambda \in \Lambda} U_\lambda \in O$ falls $U_i \in O$

O heisst die *Topologie*.

Satz 1.16. *Sei (X, d) ein metrischer Raum. Wir definieren die entsprechende offene Kugel mit Mittelpunkt $x \in X$ und Radius $r > 0$:*

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

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

1.1 Konvergenz

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

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

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

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

Dann schreiben wir

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

Satz 1.18.

$$x_k \to x_\infty \iff x_{ki} \to x_{\infty_i} \ \ \forall i \in \{1, \cdots, n\}$$

Beweis.

$$\|x_k - x_\infty\| = \sqrt{\sum_{i=1}^n (x_{ki} - x_{\infty_i})^2} \geq |x_{ki} - x_{k\infty}| \geq 0$$

$$\implies 0 \leq \lim_{k\to\infty} |x_{ki} - x_{k\infty}| \leq \lim \|x_k - x_\infty\| = 0$$

$$\|x_k - x_\infty\| = \underbrace{\sqrt{\sum_{i=1}^n \underbrace{(x_{ki} - x_{\infty_i})^2}_{\to 0}}}_{\to 0} \leq \sum_{i=1}^n |x_{ki} - x_{\infty_i}|$$

$$\implies \|x_k - x_\infty\| \to 0$$

Eine alternative Formulierung: $\lim\limits_{k\to\infty} x_k = \left(\lim\limits_{k\to\infty} x_{k1}, \cdots, \lim\limits_{k\to\infty} x_{kn}\right)$ □

Bemerkung 1.19. Die Folge $\{x_k\}$ konvergiert gegene x_∞ genau, dann wenn

$$\forall \varepsilon > 0 \quad \exists N : \|x_k - x_\infty\| < \varepsilon \ \text{ falls } \ k \geq N\,. \tag{1}$$

Eine äquivalente Formulierung von (1) ist

$$\text{für jede Umgebung } U \text{ von } x_\infty \text{ fast alle } x_k \in U. \tag{2}$$

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.*

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

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

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

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

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

□

Bemerkung 1.22. In einem metrischen Raum, Cauchy $\Leftarrow$ Konvergenz. Aber allgemein: die Cauchy Bedingung impliziert nicht die Konvergenz. Falls Cauchy $\Longrightarrow$ Konvergenz, dann ist der metrische Raum *vollständig.*

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

Satz 1.24. *1. Eine konvergente Folge ist beschränkt*

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

Beweis. Die erste Aussage ist eine Triviale Folgerung der Dreiecksungleichung. In der Tat, wenn x_k gegen x_∞ konvergiert, dann ist $\|x_k - x_\infty\|$ eine Nullfolge. Deswegen ist $\|x_k - x_\infty\|$ eine beschränkte Folge. Aber $0 \leq \|x_k\| \leq \|x_\infty\| + \|x_k - x_\infty\|$.

Wir beweisen nun die zweite Aussage.

$$\{x_k\} \text{ beschränkt} \Longrightarrow \{x_{k1}\}_{k\in\mathbb{N}} \text{ beschränkt}$$

$$\Longrightarrow \exists x_{k_j} : x_{k_j 1} \to x_1$$

Wir definieren $y_j := x_{k_j}$ $y_{j1} \to x_1$

$$y_j \text{ beschränkt} \Longrightarrow \exists j_l : y_{j_l 2} \to x_2$$

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

Nach $\ldots (n-2)$ Schritte finden wir eine Teilfolge w_r von x_k mit $w_{ri} \to x_i$. Deswegen

$$w_r \to (x_1, \cdots, x_n)$$

□

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. $\varnothing, \mathbb{R}^n$ sind abgeschlossen*

2. $G_1, \cdots, G_N$ abgeschlossen $\Longrightarrow G_1 \cup G_2 \cup \cdots \cup G_N$ abgeschlossen

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

Beweis. Diese Eigenschaften sind Folgerungen der entsprechenden Eigenschaften der offenen Mengen. □

Satz 1.28. *Sei $G \subset \mathbb{R}^n$. Dann G ist abgeschlossen genau dann, wenn*

$$\text{für jede konvergente } \{x_k\} \subset G \text{ gehört der Grenzwert zu } G. \tag{3}$$

Bemerkung 1.29. Es ist leicht zu sehen dass der folgende Beweis auch für metrische Räume gilt.

Beweis. $\Leftarrow$ Wir nehmen an dass (3) 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: $K_{\frac{1}{j}}(x) \not\subset G^c$, $j \in \mathbb{N} \setminus \{0\}$

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

$$\{x_j\} \subset G \;\; x_j \to x \;\; x \notin G$$

Aber diese letzte Aussage widerspricht (3). Wir schliessen deswegen dass G^c offen ist.

Wir beweisen nun die andere Aussage. Widerspruchsbeweis: G^c ist offen, aber $\exists \{x_k\} \subset G$ mit Grenzwert $x \notin G$, d.h. $x \in G^c$. Da G^c offen ist,

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

Aber die Konvergenz gegen x impliziert die Existenz von N s.d. $\|x_k - x\| < r$ für $k \geq N$. Deswegen

$$\|x_N - x\| < r \implies x_N \in K_r(x) \cap G \implies K_r(x) \cap G \neq \emptyset \implies \text{Widerspruch.}$$

□

Beispiel 1.30. 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 $y_k \to y$ und $\|y - x\| = r$.

Definition 1.31. Sei $\overline{K_r(x)} := \{y \in \mathbb{R}^n : \|y - x\| \leq r\}$ die *geschlossene* Kugel.

Übung 1.32. $\overline{K_r(x)}$ ist abgeschlossen.

Definition 1.33. $x \in \mathbb{R}^n$ ist ein Randpunkt von M falls

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

Definition 1.34. 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 1.35. *Sei $M \subset \mathbb{R}^n$. Dann $\partial M^c = \partial M$. Ausserdem,*

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. Die Aussage $\partial M = \partial M^c$ ist offen

Beweis von 1. Zuerst zeigen wir dass $M \setminus \partial M$ offen ist.

$$x \in M \setminus \partial M \implies x \in M \;\text{ und }\; \exists K_r(x) \;\text{ mit }\; K_r(x) \cap M^c = \varnothing$$
$$\implies K_r(x) \subset M$$

Sei $y \in K_r(x)$

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

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

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

Beweis von 2. Aus 1. folgt dass $M^c \setminus \partial M^c$ die grösste offene Teilmenge von M^c ist. Deswegen, $(M^c \setminus \partial M^c)^c$ die kleinste geschlossene Menge ist, die M entählt. Aber $(M^c \setminus \partial M^c)^c = (M^c)^c \setminus \partial M^c = M \setminus \partial M^c = M \setminus \partial M$.

□

1.3 Stetigkeit

Definition 1.36. Sei $f : \Omega_{\subset\mathbb{R}^n} \to \mathbb{R}^k$. f ist stetig an der Stelle $x \in \Omega$ falls $\forall \{x_k\} \subset \Omega$ mit $x_k \to x$.

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

Falls $f : \Omega \to \mathbb{R}^k$, $x \in \Omega$ und $y \in \mathbb{R}^k$ erfüllen die Bedingung

$$f(x_k) \to y \qquad \forall \text{ Folge } \{x_k\} \subset \Omega \setminus \{x\} \text{ mit } x_k \to x\,,$$

dann schreiben wir

$$\lim_{z\to x} f(z) = y\,.$$

Deswegen,

$$f \text{ stetig in } x \iff \lim_{z\to x} f(z) = f(x)\,.$$

Lemma 1.37. *Eine äquivalente Definition der Stetigkeit an der Stelle x:*

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

Bemerkung 1.38. Aus diesem Lemma folgt:

$$y = \lim_{z\to x} f(z) \iff (\forall\varepsilon > 0\, \exists\delta : 0 < \|z - x\| < \delta \implies \|f(z) - y\| < \varepsilon)\,.$$

Beweis. **ε-δ $\implies$ Folgendefinition.** Sei $x_k \to x$. Das Ziel ist $f(x_k) \to f(x)$ zu zeigen. D.h., $\forall\varepsilon > 0$ eine N zu finden s.d.

$$\underbrace{\underbrace{\|f(x_k) - f(x)\|}_{d(f(x_k),f(x))}}_{f(x_k)\in K_\varepsilon(f(x))} < \varepsilon \;\; \forall k \geq N$$

Sei $\varepsilon > 0$ gegegeben. Dann

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

Aber, da $x_k \to x$, $\exists N$ s.d.

$$\|x_k - x\| < \delta \;\; \forall k \geq N.$$

Für $k \geq N$ gilt

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

Folgendefinition $\implies$ (ε-δ)-Defintion. Widerspruchsannahme:

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

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

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

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

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

$x_j \to x$ aber $f(x_j) \not\to f(x)$ ∎

Definition 1.39. Die allgemeine Defintion der Stetigkeit für metrische Räume: Seien (X, d) und $(Y, \overline{d})$ zwei metrische Räume. Sei $f : X \to Y$. f ist stetig an der Stelle x falls:

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

$$\text{d.h.} \quad f(K\delta(x)) \subset K_\varepsilon(f(x))\,.$$

Definition 1.40. Eine $f : X \to Y$ heisst stetig falls f stetig an jeder Stelle $x \in X$ ist.

Satz 1.41. *Sei $f : X \to Y$ $((X, d), (Y\overline{d})$ 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.** Sei U eine Umgebung von $f(x) \implies \exists\delta > 0$ mit $K_\delta(f(x)) \subset U$

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

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

• **Umgebung $\implies$ Stetigkeit.** Sei $\delta > 0$ und $U := K_\delta(f(x))$. U ist eine 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

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

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

• **offen $\implies$ Stetigkeit.** Sei $x \in X$, und $\delta > 0$. $K_\delta(f(x))$ ist eine offene Menge.

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

Aber x gehḧort zu $f^{-1}(K_\delta(f(x)))$

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

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

□

1.4 Lineare Abbildungen

Definition 1.42. Eine Abbildung $L : V \to W$ (V, W Vektorräume) heisst linear, falls

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

Falls $L, L' : V \to W$ zwei lineare Abbildungen sind und $\lambda, \mu \in \mathbb{R}$, dann ist die Abbildung $v \mapsto \lambda L(v) + \mu L'(V)$ auch linear. Der Raum $\mathcal{L}(V, W) := \{L : V \to W \text{ linear}\}$ ist ein Vektorraum. Falls $V = \mathbb{R}^m$ und $W = \mathbb{R}^k$, dann $\exists$ eine Matrix (L_{ij}) mit

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

(L_{ij}) is die *Matrixdarstellung* der linearen Abbildung L.

Definition 1.43. Sei L_{ij} eine Matrix die die lineare Abbildung $L : \mathbb{R}^n \to \mathbb{R}^k$ darstellt. Die Hilbert-Schmidt Norm von L ist

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

Bemerkung 1.44. $\mathcal{L}(V,W) \sim \{L : (L_{ij})\ n \times k \text{ Matrizen}\} \sim \mathbb{R}^{nk}$. D.h., der Raum der $n \times k$ Matrizen ist ein Vektorraum. $\|.\|_{\mathrm{HS}}$ ist die Euklidische Norm.

Bemerkung 1.45. Sei $L : \mathbb{R}^n \to \mathbb{R}^k$ eine lineare Abbildung und $x \in \mathbb{R}^n$. Dann die Ungliechung

$$\|L(x)\|_e \leq \|x\|_e \|L\|_{\mathrm{HS}} \tag{4}$$

ist eine einfache Folgerung der Cauchy-Schwartz Ungliechung.

Beweis. Beweis von (4): $L(x) = y$

$$\|L(x)\|^2 = \sum_{i=1}^k y_i^2$$

$$= \sum_{i=1}^k \left(\sum_{j=1}^n L_{ij} x_j\right)^2 \overset{\text{Cauchy-Schwartz}}{\leq} \sum_{i=1}^k \left(\sum_{j=1}^n L_{ij}^2\right) \left(\sum_{j=1} x_j\right)^2$$

$$= \sum_{i=1}^k \sum_{j=1}^n L_{ij}^2 \|x\|^2 = \|x\|^2 \left(\sum_{i=1}^k \sum_{j=1}^n L_{ij}^2\right) = \|x\|^2 \|L\|_{HS}^2 .$$

□

Korollar 1.46. *Sei L wie oben, dann ist L stetig.*

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

$$\|L(x_k) - L(x)\| = \|L(x_k - x)\| \leq \|x_k - x\| \|L\|_{\mathrm{HS}} \to 0$$

$$\implies \|L(x_k) - L(x)\| \to 0 .$$

□

Definition 1.47. Sei $L : V \to W$ eine lineare Abbildung wobei $(V, \|.\|_V)$ und $(W, \|.\|_W)$ zwei endlich-dimensionierte normierte Vektorräume sind. Die Operatornorm von L ist:

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

Satz 1.48. $\|.\|_{L(V,W)}$ *ist eine Norm und*

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

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

Beweis. Der Kern ist die folgende Eigenschaft:

$$\|L\|_{L(V,W)} < +\infty \tag{5}$$

Das nehmen wir an ohne Beweis (für einen Beweis brauchen wir die Kompaktheit der geschlossenen Kugel , siehe Übungen). Wenn (5) gilt:

1.
$$\underbrace{\|L\|_{L(V,W)}}_{\text{Kern}} \quad \text{und} \quad \|L\|_{L(V,W)} = 0 \iff L = 0$$

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

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

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

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

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

2.
$$
\begin{aligned}
\|\lambda L\|_{L(V,W)} &= |\lambda| \, \|L\|_{L(V,W)} \\
\|\lambda L\|_{L(V,W)} &= \sup_{\|y\|_V \leq 1} \|\lambda L(v)\|_W \\
&= \sup_{\|y\|_V \leq 1} |\lambda| \, \|L(v)\|_W \\
&= |\lambda| \sup_{\|y\|_V \leq 1} \|L(v)\|_W \\
&= |\lambda| \, \|L\|_{L(V,W)}
\end{aligned}
$$

3.
$$
\begin{aligned}
&\|L + L'\|_{L(V,W)} \\
&= \sup_{\|y\|_V \leq 1} \|(L + L')(v)\|_{L(V,W)} \\
&= \sup_{\|y\|_V \leq 1} \|L(v) + L'(v)\|_{L(V,W)} \\
&\leq \sup_{\|y\|_V \leq 1} (\|L(v)\|_W + \|L'(v)\|_W) \\
&\leq \sup_{\|y\|_V \leq 1} \|L(v)\|_W + \sup_{\|y\|_V \leq 1} \|L'(v)\|_W \\
&= \|L\|_{L(V,W)} + \|L'\|_{L(V,W)}
\end{aligned}
$$

□

Bemerkung 1.49. Aus der Definition von $\|\cdot\|_{L(V,W)}$ folgt

$$
\|L(v)\|_W \leq \|L\|_{L(V,W)} \, \|v\|_V \qquad \forall v \in V \, . \tag{6}
$$

Falls $\|v\|_V = 1$, dann

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

Für $v = 0$ ist $L(v) = 0$ und deswegen ist die Ungliechung (6) trivial. Falls $\|v\|_V > 0$,

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

In der Tat, $\|L\|_{L(V,W)}$ ist die *optimale Konstante* in (6). D.h., für jede $C < \|L\|_{L(V,W)}$ $\exists v \in V$ mit $\|L(v)\|_W > C\|v\|_V$.

Korollar 1.50. *Seien V und W zwei endlichdimensionierte Vektorräume und $L : V \to W$ eine lineare Abbildung. Dann L ist stetig.*

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 (6)

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

$\implies$ Ungleichung erfüllt. □

Bemerkung 1.51. Seien $V = \mathbb{R}^n$ und $\|.\|_V$ die euklidische Norm, $W = \mathbb{R}^k$ und $\|\cdot\|_W$ die euklidische Norm. Dann (4) ist einfach die folgende Aussage:

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

In Matrixdarstellung:

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

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

In diesem Fall wir nutzen die Notation $\|\cdot\|_O$ für die Operatornorm.

1.5 Mehr über stetige Funktionen

Regeln für stetige Funktionen

Regel 1 Seien $f : X \to Y$, $g : X \to Z$ zwei stetige Funktionen (X, Y und Z topologische Räume). Dann

- falls $Y = Z$ ein normierter Vektorraum ist, $f + g$ ist auch stetig;
- falls Y ein normierter Vektorraum und $Z = \mathbb{R}$, gf ist auch stetig;
- falls $Y = Z = \mathbb{R}^n$ auch

  $$x \mapsto f(x) \cdot g(x) = \sum_{i=1}^n f_i(x) g_i(x)$$

  ist stetig.

Beweis. Wir geben den Beweis für den Fall $X \subset \mathbb{R}^m$. Der allgemeine Fall lassen wir als eine übung. In diesem Fall können wir die Folgendefinition der Stetigkeit anwenden.

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

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

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

$$g(x) = (g_1(x), \cdots, g_m(x))$$
$$f(x^k) = (f_1(x^k), \cdots, f_m(x^k))$$
$$f(x) = (f_1(x), \cdots, f_m(x))$$
$$(g+f)(x^k) = \big(g_1(x^k) + f_1(x^k), \ldots, g_m(x^k) + f_m(x^k)\big)$$
$$\to (g_1(x) + f_1(x), \ldots, g_m(x) + f_m(x)) = (g+f)(x)\,.$$

D.h.

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

DIe anderen Regeln folgen aus ähnlichen Argumente. $\square$

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

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

ist stetig.

Beweis. Sei U eine offene Menge in Z.

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

□

Definition 1.52. Sei $f : X \to \mathbb{R}$.

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

$f : X \to V$, $V, \|.\|_V$ normierter Vektorraum

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

Bemerkung 1.53. X Menge, $V, \|.\|$ ein normierter Vektorraum.

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

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

Definition 1.54. Eine Folge von Funktionen

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

konvergiert gleichmässig gegen f falls

$$\left\|f^k - f\right\| \to 0$$

Bemerkung 1.55. $x \in X$

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

Folgerung f^k konvergiert gleichmässig

$$\implies f^k(x) \to f(x) \;\; \forall x$$

Satz 1.56. *Sei X ein metrischer Raum und $f^k : X \to V$ eine Folge die gleichmässig gegen f konvergiert. Dann ist f stetig.*

Beweis. Seien $x \in X$ und $\varepsilon > 0$. Wir suchen $\delta > 0$ so dass

$$d(x, y) < \delta \implies \|f(x) - f(y)\| < \varepsilon\,. \tag{7}$$

Aus der gleichmässigen Konvergenz folgt die Existenz von N so dass

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

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

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

Siene nun x, y s.d. $d(x, y) < \delta$. Dann

$$\begin{aligned}
\|f(x) - f(y)\| &= \left\|(f(x) - f^N(x)) + (f^N(x) - f^N(x)) + (f^N(y) - f(y))\right\|_V \\
&\leq \left\|f(x) - f^N(x)\right\|_V + \left\|f^N(x) - f^N(y)\right\|_V + \left\|f^N(y) - f(y)\right\|_V \\
&< \left\|f^N - f\right\| + \frac{\varepsilon}{3} + \left\|f^N - f\right\| \\
&< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon
\end{aligned}$$

□

1.6 Kompakte Menge

Definition 1.57. 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 1.58. *Sei $K \subset \mathbb{R}^n$.*

$$K \text{ kompakt} \quad \iff \forall \left\{x^j\right\} \subset K \;\exists \text{ Teilfolge} x^{j_l} \text{ die gegen } x \in K \text{ konvergiert.} \tag{8}$$

Die Eingeschaft in der rechten Seite von (8) heisst *Folgenkompatkheit.* Der Satz 1.58 ist also die folgende Behauptung:

$$\text{falls } K \subset \mathbb{R}^n \text{ dann} \qquad K \text{ kompakt} \iff K \text{ folgenkompakt.}$$

Beweis. **Kompaktheit $\Longrightarrow$ Folgenkompaktheit.** Sei K kompakt und $\left\{x^j\right\} \subset K$ eine Folge.

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

Aus der Bolzano-Weiertsrass Eigenschaft $\exists x^{j_l} \to x \in \mathbb{R}^n$. Die abgeschlossenheit von $K \implies x \in K$.

Folgenkompaktheit $\Longrightarrow$ Abgeschlossenheit und Beschränktheit.

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

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

Widerspruch (weil $x = y$!).

Sei K nicht beschränkt.

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

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

Wenn $x^{j_l} \to x$. Aber das impliziert dass $\{\|x^{j_l}\|\}$ eine beschränkte Folge ist. (Wir wiederlegen das Argumebnt:

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

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

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

$$\implies \left\|x^{j_l}\right\| \to \|x\|)$$

Aber $\left\|x^{j_l}\right\| = j_l \to +\infty \implies$ Widerspruch. □

Wir geben noch eine zweite Characterisierung der kompaken Teilmenge von $\mathbb{R}^n$.

Definition 1.59. (Überdeckungseigenschaft) EIne Familie $\{U_\lambda\}_{\lambda \in \Lambda}$ von Teilmengen von $\mathbb{R}^n$ ist eine Überdeckung einer Menge E falls

$$\bigcup_{\lambda \in \Lambda} U_\lambda \supset E$$

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

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.

Beispiel 1.60. 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 $x \in U_x$!
2. Keine endliche Teilfamile von $\{U_x\}$ ist eine Überdeckung von $K_r(0)$. In der Tat, sei $\{U_{x_1}, \cdots, U_{x_N}\}$ eine beliebige endliche Teilfamilie. Sei
$$\rho := \max_{i\in\{1,\cdots,N\}} \|x_i\| < r$$
$\implies$ falls $\|y\| \geq \frac{\|x_i\|+r}{2}$ dann $y \notin U_{x_i}$. So, wenn $\|y\| \geq \frac{\rho+r}{2}$ dann
$$y \notin U_{x_1} \cup \cdots \cup U_{x_N}\,.$$
Aber $\frac{\rho+r}{2} < r$. So, wenn $\|y\| = \frac{\rho+r}{2}$, dann $y \in K_r(0)$.

Jede geschlossene Kugel hat die Überdeckungseigenschaft: das ist eine Konsequenz des nächsten Satzes.

Satz 1.61. *Sei $E \subset \mathbb{R}^n$*
$$E \text{ kompakt} \iff E \text{ hat die Überdeckungseigenschaft}$$

Bemerkung 1.62. Satz 1.61 kann auch so formuliert werden:
$$(E \text{ beschränkt und abgeschlossen}) \iff E \text{ hat die Überdeckungseigenschaft}\,.$$

Das Beispiel 1.60 erklärt wie so die Abgeschlossenheit nötig ist. Sei nun $E = \mathbb{R}^n$ und $U_n = K_{n+1}(0)$.
$$E \subset \bigcup_{n\in\mathbb{N}} U_n$$
Aber $\forall N \in \mathbb{N}$
$$\mathbb{R}^n = E \not\subset \bigcup_{n=0}^{N} U_n\,.$$
Dieses Beispiel zeigt wie so die Beschränkheit nötig ist.

Beweis. [Beweis des Satzes 1.61] E **ist nicht kompakt** $\implies$ **Überdeckungseigenschaft gilt nicht.** Da E nocht kompakt ist, $\exists\,\{x_i\} \subset E$ ohne konvergente Teilfolge in E. $\implies$ Zwei Möglichkeiten:

1. $\exists$ eine beschränkte Teilfolge von $\{x_i\}$. Bolzano-Weierstrass $\implies$ $\exists$ Teilfolge $\{y_i\} \subset \{x_i\} \subset E$ die gegen $y \in \mathbb{R}^n$ konvergiert. $y \notin E$.
2. $\{x_k\}$ besitzt beschränkte Teilfolge $\implies$ $\|x_i\| \to \infty$.

Beim ersten ist die folgende Menge offen:
$$U_0 := \mathbb{R}^n \setminus \underbrace{(\{y_i\} \cup \{y\})}_{E \text{ ist abgeschlossen}}$$
Beim zweiten gilt:
$$U_0 = \mathbb{R}^n \setminus \underbrace{\{x_i\}}_{F} \quad \text{ist offen}$$
$$U_n = U_0 \cup \{y_1, \cdots, y_{n-1}\} \quad n \geq 0$$
U_n ist auch offen.
$$\bigcup_{n=0}^{\infty} U_n = \begin{cases} \mathbb{R}^n \setminus \{y\} & \text{im Fall 1} \\ \mathbb{R}^n & \text{im Fall 2} \end{cases}$$
Aber jede endliche Familie
$$U_0 \cup U_1 \cup \cdots \cup U_n \not\supset E$$
in beiden Fällen lassen wir unendlich viele Punkte weg.

E **kompakt** $\implies$ E **besitzt die 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 Würfel. Jeder Würfel hat die Form

$$[k_1, k_1+1] \times [k_2, k_2+1] \times \cdots \times [k_n, k_n+1] \tag{9}$$

wobei $k_1, \ldots k_n \in \mathbb{Z}$. Nun, da E beschränkt ist, $\exists N \in \mathbb{N}$ so dass $[-N, N]^n \supset E$. Aber $[-N, N]^n$ können wir mit $M = (2N)^n$ Würfel der Form (9) überdecken:

$$E \subset W_1 \cup \cdots \cup W_M$$

Falls jedes $E \cap W_i$ mit einer endlichen Familie von $\{U_\lambda\}$ überdeckt wird, dann finde ich eine endliche Überdeckung von E wenn ich die Vereinung der entsprechenden endlichen Teilüberdeckungen von $E \cap W_i$ nehme. So, angenommen dass die Überdeckungseigenschaft nicht gilt, $\exists E_1 := E \cap W_i$ s.d.

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 $\frac{1}{2}$

$$\tilde{W}_1, \cdots, \tilde{W}_{2^n}\,.$$

Mit dem obigen Argument finden wir

$$E_2 := E \cap \tilde{W}_i : \text{ die Eigenschaften 1. und 2. mit } E_2 \text{ statt } E_1 \text{ noch gelten}$$

Induktiv

$$E \supset E_1 \supset E_2 \supset \cdots$$

jede $E_i \subset W^i$ Würfel mit Seite 2^{-i+1} und die beiden Eigenschaften 1. und 2. gelten mit E_i statt E_1.

Ausserdem, E_i ist nicht leer. Für jede i wählen wir $x_i \in E_i$. Dann $\{x_k\} \subset E$. Aber $\{x_k\}$ ist eine Cauchy-Folge: falls $j, k > i$, $x_k, x_j \subset E_i$ und E_i ist in einem Würfel mit Seite 2^{-i+1} enthalten. Deswegen $\|x_j - x_k\| \leq \sqrt{n}2^{-i+1}$. Die Vollestendigkeit von $\mathbb{R}^n$ garantiert die Existenz von $x \in \mathbb{R}^n$ s.d. $x_k \to x$. Da E abgeschlossen ist, $x \in E$. Deswegen $\exists U_\mu \in \{U_\lambda\}_\lambda$ s.d. $x \in U_\mu$. Da U_μ offen ist,

$$\exists K_r(x) \supset U$$

Aber, $x \in E_i$ für jedes i (weil $\{x_k\}_{k\geq i} \subset E_i$ und E_i ist abgeschlossen!). Sei nun $k \in \mathbb{N}$ s.d. $\sqrt{n}2^{-k+1} < r$. Falls $y \in E_k$, dann $\|y - x\| \leq \sqrt{n}2^{-k+1} < r$. Deswegen $E_k \subset K_r(x) \subset U_\mu$. So, die Familie $\{U_\mu\}$ ist endlich (entählt sogar einen einzigen Element!) und überdeckt E_k. Widerspruch! $\square$

Bemerkung 1.63. f stetig $\implies$ $f^{-1}(U)$ offen falls U offen: diese mächtige Characterisierung der Stetigkeit werden wir nun nutzen!

Korollar 1.64. *Sei $E \subset \mathbb{R}^n$ kompakt und $f\mathbb{R}^m \to \mathbb{R}^k$ stetig. Dann $f(E)$ ist kompakt.*

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

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

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

Korollar 1.65. *Wenn $f : \mathbb{R}^n \to \mathbb{R}$ stetig ist und $E \subset \mathbb{R}^n$ kompakt ist, besitzt f ein Maximum und ein Minimum auf E.*

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

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

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

$\implies$ s ist ein Maximum. □

Ohne Beweis:

Lemma 1.66. *[Lemma von Tietze] Sei $E \subset \mathbb{R}^m$ kompakt und $f : E \to \mathbb{R}$ stetig. Dann $\exists g : \mathbb{R}^n \to \mathbb{R}$ stetig s.d. $g|_E = f$.*

Wenn wir Lemma 1.66 und Korollar 1.65 kombinieren, erhalten wir den folgenden Satz:

Satz 1.67. *Wenn $E \subset \mathbb{R}^n$ kompakt ist und $f : E \to \mathbb{R}$ stetig ist, besitzt f ein Maximum und ein Minimum.*

Wir geben auch einen alternativen Beweis, unabhängig von Tietzes Lemma

Beweis. Sei $s = \sup\{f(x) : x \in R\}$ (es kann sein dass $s = \infty$). Dann $\exists\{x_k\} \subset E$ s.d. $f(x_k) \to s$. Die Kompaktheit von E impliziert die Existenz einer Teilfolge $\{x_{k_i}\}$ die gegen einen Element $x \in E$ konvergiert. Deswegen

$$s = \lim_{i\to\infty} f(x_{k_i}) = f(x)\,.$$

Ein ähnlichens Argument beweist die Existenz einer Minimumstelle. □

Zur Erinnerung: das Intervallschachtelungsprinzip in $\mathbb{R}$.
Sei I_j eine Intervallschachtelung d.h.:

1.
$$I_j = [a_j, b_j]$$

2.
$$I_0 \supset I_1 \supset \cdots \supset I_j \supset I_{j+1}$$

3.
$$b_j - a_j \to 0$$

Dann

$$\bigcap_{j=0}^{\infty} E_j \neq \varnothing$$

Ein Verallgemeinerung dieses Prinzips ist der Folgende

Satz 1.68. *Sei E_j eine Folge von kompakten Mengen mit $E_j \supset E_{j+1}$ $\forall j$ ($E_0 \subset \mathbb{R}^n$). Dann*

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

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

$$U_N \not\supset E_N \neq \varnothing \ \ E_N \subset E_0$$

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

Wir geben endlich eine Zusammenfusassung der Eigenschaften der stetigen Funktionen $f : \mathbb{R}^n \to \mathbb{R}^k$:

- $E \subset \mathbb{R}^k$ offen $\implies$ $f^{-1}(E)$ offen;
- $E \subset \mathbb{R}^k$ geschlossen $\implies$ $f^{-1}(E)$ geschlossen;
- $E \subset \mathbb{R}^n$ kompakt $\implies$ $f(E)$ kompakt.

Aber *Vorsicht!*

- $E \subset \mathbb{R}^n$ offen impliziert **nicht** $f(E)$ offen;
- $E \subset \mathbb{R}^n$ geschlossen impliziert **nicht** $f(E)$ geschlossen;
- $E \subset \mathbb{R}^k$ kompakt impliziert **nicht** $f^{-1}(E)$ kompakt.

2 Differenzierbare Funktionen

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

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

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

Definition 2.1. $f : U \to \mathbb{R}$, $U \subset \mathbb{R}^n$ offen, heisst differenzierbar in $a \in U$, falls es eine lineare Abbildung $L : \mathbb{R}^n \to \mathbb{R}$ gibt so dass

$$\lim_{h\to 0} \frac{f(a+h) - f(a) - L(h)}{\|h\|} = 0\,. \tag{10}$$

Bemerkung 2.2. $n = 1$: die Funktion ist differenzierbar genau dann, wenn die Ableitung existiert. In diesem fall gilt $L(h) = f'(a)h$.

Bemerkung 2.3. Die lineare Abbilung L in (10) ist eindeutig definiert. In der Tat seien L' und L zwei lineare Abbildungen die (10) erfüllen. Sei $v \in \mathbb{R}^n$ mit $\|v\| = 1$. Es gilt:

$$(L - L')(v) \overset{\text{linear und } \|v\|=1}{=} \lim_{t\downarrow 0} \frac{(L-L')(tv)}{\|tv\|} \overset{(10) \text{ mit } h=tv}{=} 0\,.$$

Deswegen $L = L'$.

Bemerkung 2.4. Wir können (10) auch anders beschreiben:

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

Dann gilt

$$(10) \iff \lim_{h\to 0} \frac{R(h)}{\|h\|} = 0 \tag{11}$$

Definition 2.5. L heisst Differential von f in a. Man schreibt $\mathrm{d}\, f|_a$. Sei nun $\{e_1, \cdots, e_n\}$ die Standardbasis $\mathbb{R}^n$, $h = (h_1, \cdots, h_n) \in \mathbb{R}^n$

$$\implies \mathrm{d}\, f|_a(h) = \mathrm{d}\, f|a \left(\sum_{i=1}^{k} h_i - e_i \right) = \sum_{i=1}^{n} h_i \,\mathrm{d}\, f|_a(e_i)$$

Definition 2.6.

$$\nabla f(a) = (\mathrm{d}\, f(a)e_1, \cdots, \mathrm{d}\, f(a)e_n)$$

heisst Gradient von f.

Die Affine Abbildung

$$Tf(x,a) = f(a) + \nabla df|_a(x-a)$$

ist die beste lineare Approximation der Funktion f an der Stelle a. Der Graph von Tf ist eine (hyper)Ebene von $\mathbb{R}^{n+1}$: die heisst die tangentiale Ebene.

Satz 2.7. *f differentierbar in $a \implies f$ ist stetig in a*

Beweis.

$$|f(a+b) - f(a)| = |\mathrm{d}\, f|_a(b) + R(b)| \leq \|\, \mathrm{d}\, f|_a\|_O \|h\| + \underbrace{|R(h)|}_{\to 0}$$

□

Beispiel 2.8. $f(x) = Ax+b$, $A \in M_a(1, n, \mathbb{R})$, $b \in \mathbb{R}$. Dann f ist differenzierbar und $df|_a(h) = a \cdot h$. In der Tat die Abbildung $L(h) := a \cdot h$ ist linear und

$$f(a+h) - f(a) - L(h) = 0 =: R(h)\,.$$

$\frac{R(h)}{\|h\|} \to 0$ trivialerweise!

Beispiel 2.9. $f(x) := x^T \cdot A \cdot x$, $A = (a_{ij}) \in \mathrm{Sym}(n, \mathbb{R})$

$$f(a+h) - f(a) = \underbrace{2a^T Ah}_{\mathrm{d}\,|_a(h)} + \underbrace{h^T Ah}_{R(h)}\,.$$

$L(h) := 2a^T Ah$ ist linear (in h), $R(h) = h^T \cdot A \cdot h$ $(= \sum h_i a_{ik} h_l)$. Wir haben $\|A \cdot h\| \leq \|A\|_O \|h\|$ und

$$|h^T \cdot A \cdot h| = |\underbrace{h^T \cdot (A \cdot h)}_{\text{Skalarprodukt von } h \text{ und } A \cdot h}| \overset{\text{Cauchy-Schwartz}}{\leq} \|h\| \|A \cdot h\| \leq \|A\|_O \|h\|^2\,.$$

Deswegen

$$\frac{|Rh|}{\|h\|} \leq \|A\|_O \|h\| \to 0\,.$$

Ziel Wir wollen $\mathrm{d}\, f_a(h)$ berechnen. Sei $t \in \mathbb{R}$. Dann

$$f(a+th) = f(a) + \mathrm{d}\, f(a)th + R(th)$$

$$\implies \mathrm{d}\, f|_a(h) = \lim_{t \to 0} \frac{f(a+th) - f(a)}{t} \tag{12}$$

Definition 2.10. $f : U \to \mathbb{R}$, $a \in U$. Die Richtungsableitung von f in Richtung $h \in \mathbb{R}^n$ ist der Grenzwert (falls er existiert)

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

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

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

Bemerkung 2.11. Wir haben <u>nicht</u> vorausgesetzt, dass f differenzierbar ist in a!

Satz 2.12. *Sei f in a differenzierbar. Dann existieren die Richtungsableitungen in jede Richtung. Insbesondere existieren die partiellen Ableitungen. Es gelten:*

$$\mathrm{d}\, f|_a(h) = \nabla f(a) \cdot h = \partial_n f(a) = \sum_{i=1}^n \partial_i f(a) h_i \tag{13}$$

und

$$\nabla f(a) = (\partial_1 f(a), \cdots, \partial_n f(a))$$

Beweis. Die xistenz der Richtungsableitung ist die Herleitung von 12. (13) ist eine triviale Konsequenz der Linearität von $df|_a$. □

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

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t\to 0} \frac{f(a+t_{ei}) - f(a)}{t}, \quad a = (a_1, \cdots, a_n)\,.$$

Wenn wir definieren

$$g_i(y) := f(a_1, \cdots, a_{i-1}, y, a_{i+1}, \cdots, a_n)$$

dann gilt

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t\to 0} \frac{g(a_i+t) - f(a_i)}{t} = g'(a_i)$$

Beispiel 2.13.

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

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

Beispiel 2.14.

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

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

$$\Longrightarrow \partial_h f(0,0) = \lim_{t\to 0} \frac{f(th_1, th_2) - f(0,0)}{k} = \lim_{t\to 0} \frac{t}{t} f(h_1,h_2) = f(h_1,h_2)$$

$$\begin{aligned} &\Longrightarrow \partial f(0,0) = f(h_1,h_2) \\ &\partial_{e_1} f(0,0) = f(1,0) = 0 \\ &\partial_{e_2} f(0,0) = f(0,1) = 0 \end{aligned}$$

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

$$\xRightarrow{\text{aus } 13} \underbrace{\partial_n f(0,0)}_{=\mathrm{d}\, f(a)h=0} = \underbrace{\partial_1 f(a)}_{0}(h_1) + \underbrace{\partial_2 f(a)}_{0}(h_2) = 0$$

$$\Longrightarrow \mathrm{d}\, f(a) = 0$$

Test $L = 0$

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

$\Longrightarrow$ f ist in $(0,0)$ NICHT DIFFERENZIERBAR.

D.h. **es kann sein dass die ganzen Richtungsableitungen existieren und die Funktion ist trotztdem nicht differenzierbar!**

2.1 Zusammenfassung

2.1.1 Das Differenzial

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

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

$$\lim_{h \downarrow 0} \frac{f(x+h) - f(x) - L(h)}{\|h\|} = 0 \tag{14}$$

(Zur Erinnunerung:

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

$$\iff (\forall \text{ Folgen } \{h_k\} \text{ die} \neq 0 \text{ aber} \to 0, \text{ es gilt } G(h_k) \to 0) \quad \Big)$$

Wenn f differenzierbar ist und (14) erfüllt, heisst L das Differential von f an der Stelle x:

$$L = \mathrm{d}\, f|_x \,.$$

2.1.2 Richtungsableitung

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

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

2.1.3 Partielle Ableitung

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

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

Falls $e_i = (0, \ldots, 0, \underbrace{1}_{i \text{ Stelle}}, 0, \ldots, 0)$

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

2.2 Das Hauptkriterium der Differenzierbarkeit

Die Existenz der Richtungsableitungen genugt nicht für die Differenzierbarkeit von f. Deswegen die Existenz der partiellen Ableitungen (d.h. von *manchen* Richtungsableitungen) impliziert **nicht** die Differenzierbarkeit.

Satz 2.15. *(Hauptkriterium der Differenzierbarkeit) Sei $f : U \to \mathbb{R}$ und U eine Umgebung von y. Falls $\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}$ $\underline{\text{in } U}$ existieren und stetig $\underline{\text{in } y}$ sind, dann ist f in y differenzierbar.*

Bemerkung 2.16. Aber *Vorsicht*: die Differenzierbarkeit von f impliziert **nicht** die Stetigkeit der partiellen Ableitungen!

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

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

Ziel L ist das Differential von f, d.h.

$$\lim_{h\to 0} \frac{f(x+h)-f(x)-L(h)}{\|h\|} = 0 \tag{15}$$

Wir schreiben

$$\begin{aligned}
f(x+h)-f(x) \quad = \quad & f(x+(h_1,\dots,h_n)) - f(y+(h_1,\dots,h_{n-1},0) \\
& +f(y+(h_1,\dots,h_{n-1},0) - f(y+(h_1,\dots,h_{n-2},0,0) \\
& +\dots \\
& +f(y+(h_1,\dots,h_i,0,\dots 0) - f(y+(h_1,\dots,h_{i-1},0,0,\dots 0) \qquad (i\text{te Zeile}) \\
& +\dots \\
& +f(y+(h_1,0,\dots,0)) - f(y) \qquad (16)
\end{aligned}$$

Sei nun $g_i(t)) = f(y+(h_1,\dots,h_{i-1},th_i,0,\dots,0)$. Die $i$te Zeile in (16) ist dann $g_i(1)-g_i(0)$. Aber

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

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

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

Deswegen die Existenz der Richtungsableitungen in einer Ungebung von y garantieren die Differenzierbarkeit der Funktion g_i falls $\|h\|$ klein genung ist. Ausserdem

$$\exists \xi_i \in [0,1] \ : \qquad i\text{te Zeile von (16)} = g_i'(\xi_i)$$

und so

$$i\text{te Zeile} = h_i \frac{\partial f}{\partial x_i}(y_1+h_1,\dots,y_{i-1}h_{i-1},y_i+\xi_i h_i,y_{i+1},\dots,y_n) = h_i \frac{\partial f}{\partial x_i}(y+\zeta_i)\ . \tag{17}$$

wobei $\zeta_i = (h_1,\dots,h_{i-1},\xi h_i,0,\dots,0)$ Wir setzten (17) in (16):

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

und deswegen

$$f(x+h)-f(x)-L(h) = \sum_{i=1}^n h_i \left(\frac{\partial f}{\partial x_i}(y+\zeta_i) - \frac{\partial f}{\partial x_i}(y) \right)\ . \tag{18}$$

Also,

$$\frac{|f(x+h)-f(x)-L(h)|}{\|h\|} \overset{(18)}{\leq} \sum_{i=1}^n \frac{|h_i| \left| \frac{\partial f}{\partial x_i}(y+\zeta_i) - \frac{\partial f}{\partial x_i}(y) \right|}{\|h\|} \tag{19}$$

Wenn $\|h\|\to 0$, $\|\zeta_i\|\to 0$. Die Stetigkeit von $\dfrac{\partial f}{\partial x_i}$ in y impliziert

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

Die rechte Seite von (19) $\to 0$ wenn $h\to 0 \implies$ (15). $\square$

2.3 Die geometrische Bedeutung des Gradients

Wir haben

$$df|_{x_0}(h) = \partial_h f(x_0) = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(x_0) \;= \nabla f(x_0) \cdot h$$

(manchmal wir schreiben auch $\langle \nabla f(x_0), h \rangle$). Deswegen,

$$|\partial_n f(x_0)| \overset{\text{Cauchy-Schwartz}}{\leq} \|\nabla f(x_0)\| \, \|h\|$$

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

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

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

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

Dann $\|K\| = 1$ 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.

2.4 Rechenregeln

Satz 2.17. *] Sei U eine Umgebung von $x \in \mathbb{R}^n$ und $f, g : U \to \mathbb{R}$ in x differenzierbar. Dann sind $f + g$ und fg auch differenzierbar in x und*

$$\mathrm{d}(f+g)|_x = \mathrm{d}\, f|_x + \mathrm{d}\, g|x \tag{20}$$

$$\mathrm{d}(fg) = f(x)\, \mathrm{d}\, g|x + g(x)\, \mathrm{d}\, f|_x . \tag{21}$$

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

$$\mathrm{d}\left(\frac{1}{f}\right)|_x = -\frac{1}{(f(x))^2}\, \mathrm{d}\, f|_x . \tag{22}$$

Korollar 2.18. *$g(x) \neq 0$, dann*

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

Beweis. [Beweis vom Satz 2.17] Die Regel (20) ist sehr einfach zu beweisen. Für die Regel (21) schreiben wir

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

und wir nutzen das gleiche Argument für den Fall einer reellen Variabel.

Wir beweisen nun (22). Da f stetig in x ist, $f(x+h) \neq 0$ falls $\|h\|$ klein genug ist. Deswegen ist $1/f$ wohldefiniert in einer Umgebung von x. Das Ziel ist eine lineare Abbildung L zu finden so dass

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

wobei

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

Wir schreiben

$$\lim_{h\to 0}\frac{\overbrace{\frac{1}{f(x+h)}-\frac{1}{f(x)}-\frac{1}{f(x)^2}(h)\,\mathrm{d}\,f|_x(h)}^{A}}{\|h\|}$$

und rechnen

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

Also,

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

Aber

$$\frac{B}{\|h\|}=-\underbrace{\frac{1}{f(x)f(x+h)}}_{\to f(x)^2\neq 0}\underbrace{\frac{f(x+h)-f(x)-\mathrm{d}\,f|_x(h)}{\|h\|}}_{\to 0\text{ weil } f \text{ diff. in } x}$$

und deswegen

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

Ausserdem

$$\frac{C}{\|h\|}=\underbrace{\frac{\mathrm{d}\,f|_x(h)}{\|h\|}}_{D}\frac{1}{f(x)}\underbrace{\left(\frac{1}{f(x)}-\frac{1}{f(x+h)}\right)}_{\to 0}$$

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

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

$$\implies D=\frac{|\mathrm{d}\,f|_x(h)|}{\|h\|}\le\|L\|\,.$$

Deswegen ist D beschränkt und

$$\lim_{h\to 0}\frac{C}{\|h\|}=0\,.$$

□

2.5 Kettenregel

Definition 2.19. Eine Kurve ist eine Abbildung $\gamma:[a,b]\to\mathbb{R}^n$.

Diese Definition bedeutet dass $\gamma(t)\in\mathbb{R}^n\ \forall t$. Seien nun $\gamma_i(t)$ die Koordinaten des Vektors $\gamma(t)$:

$$\gamma(t)=(\gamma_1(t),\cdots,\gamma_n(t))\,.$$

Jede $t\to\gamma_i(t)\in\mathbb{R}$ ist eine reelvertige Funktion einer Variabel.

Definition 2.20. Die Kurve γ heisst differenzierbar wenn jede γ_i differenzierbar ist. In diesem Fall definieren wir

$$\dot\gamma(t):=(\gamma'(t),\cdots,\gamma_n'(t))$$

Satz 2.21. *(Kettenregel 1. Version) Sei $f:U\to\mathbb{R}$ mit U Umgebung von x_0 und f differenzierbar in x_0. Sei $\gamma:[a,b]\to U$ eine differenzierbare Kurve mit $\gamma(t_0)=x_0$. Sei $g=f\circ\gamma$ (i.e. $g(t)=f(\gamma(t))$). Dann ist g in t_0 differenzierbar und*

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

Beweis. Das Ziel:

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

Wir definieren

$$R(h) = g(t_0+h)-g(t_0)-g(t_0)-h\left[\mathrm{d}\, f|_{\gamma(t_0)}(\dot\gamma(t_0))\,.\right] \tag{23}$$

Dann wollen wir die folgende Behauptung zeigen:

$$\lim_{h\to 0}\frac{R(h)}{|h|} = 0 \tag{24}$$

Wir führen eine neue Notation ein: wir sagen dass $R(h) = o(|h|)$ falls (23) gilt.

Aus der Differenzierbarkeit von f

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

d.h.

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

Die Differenzierbarkeit von γ impliziert

$$\lim_{k\to 0}\frac{\gamma(t_0+h)-\gamma(t_0)-\dot\gamma(t_0)}{h}\left(=: \frac{p(h)}{h}\right) = 0\,,$$

d.h.

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

Wir setzten

$$k = \gamma(t_0+h)-\gamma(t_0)$$

und schreiben

$$\begin{aligned}
g(t_0+h)-g(t_0) &= f(\gamma(t_0+h)) - g(\overbrace{\gamma(t_0)}^{x_0}) = f(\gamma(t_0)+k) - f(\gamma(t_0))\\
&= \mathrm{d}\, f|_{\gamma(t_0)}(k) + r(k)\\
&= \mathrm{d}\, f|_{\gamma(t_0)}(\gamma(t_0+h)-\gamma(t_0)) + r(k)\\
&= \mathrm{d}\, f|_{\gamma(t_0)}(h\dot\gamma(t_0)+p(h)) + r(k)\\
&\overset{\text{Linearität von } \mathrm{d}\, f}{=} h\,\mathrm{d}\, f|_{\gamma(t_0)}(\dot\gamma(t_0)) + \mathrm{d}\, f|_{\gamma(t_0)}(p(h)) + r(k)\,.
\end{aligned}$$

Deswegen

$$R(h) = g(t_0+h)-g(t_0)-h\,\mathrm{d}\, f|_{\gamma(t_0)}(\dot\gamma(t_0)) = df|_{\gamma(t_0)}(p(h)) + r(\gamma(t_0+h)-\gamma(t_0))\,.$$

$$\begin{aligned}
|R(h)| &\leq |\underbrace{df|_{\gamma(t_0)}}_{L}(p(h))| + |r(\gamma(t_0+h)-\gamma(t_0))|\\
&\leq \|L\|_O\,\|p(h)\| + \frac{r(\gamma(t_0+h)-\gamma(t_0)}{\|h\|}
\end{aligned}$$

Aber $p(h) = o(|h|) \implies \|L\|_O\,\|p(h)\| = o(|h|)$. Nun beweisen wir auch

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

Wir unterscheiden zwei Fälle. Falls

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

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

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

dan schreiben wir

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

Nun

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

(weil $k = \gamma(t_0+h)-\gamma(t_0) \to 0$ wenn $h \to 0$). Ausserdem

$$\frac{\gamma(t_0+h)-\gamma(t_0)}{h} == \underbrace{\dot\gamma(t_0)}_{\text{konstant}} + \underbrace{\frac{p(h)}{h}}_{\to 0}$$

Deswegen

$$\lim_{h\to 0}\frac{\|\gamma(t_0+h)-\gamma(t_0)\|}{|h|} = \|\dot\gamma(t_0)\|$$

$$\implies \frac{|R(h)|}{\|h\|} \to 0$$

$\implies$ Differenzierbarkeit und Kettenregel! □

Bemerkung 2.22. Als Korollar der Kettenregel erhalten wir das folgene geometrische Korollar: der Gradient ist orthogonal zur Niveaumenge der Funktion (Höhenlinien, wenn der Definitionsbreich der Funktion 2-dimensioniert ist). In der Tat, sei $\gamma : [a,b] \to U$ eine differenzierbare Kurve, U offen. Sei $f : U \to \mathbb{R}$ 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.

$$\langle \nabla f(\gamma(t)), \dot\gamma(t)\rangle = 0,$$

weil

$$0 = g'(t) = (f(\gamma(t)))' \overset{\text{Kettenregel}}{=} \langle \nabla f(\gamma(t)), \dot\gamma(t)\rangle$$

2.6 Mittelwertsatz und Schrankensatz

Sei $f : [a,b] \to \mathbb{R}$ eine differenzierbare Funktion. Dann $\exists \xi \in]a,b[$ s.d.

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

Sei nun:

$$f : U \mapsto \mathbb{R} \;\; \text{differenzierbar auf } U$$
$$x, y \in U \;\; \text{so dass das Segment } \; [x,y] \subset U$$

Das Segment $[x,y]$ ist die Menge

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

Wir definieren

$$\gamma(t) := x + t(y-x) \qquad \text{und} \qquad g = f \circ \gamma \quad (\text{d.h. } g(t) = f(\gamma(t)))\,.$$

Dann

$$f(y) - f(x) = g(1) - g(0)\,.$$

Ausserdem, γ ist differenzierbar und

$$\dot\gamma(\tau) = (\gamma_1'(\tau), \cdots, \gamma_n'(\tau)) = (y_1 - x_1, \cdots, y_n - x_n) = y - x$$

Aus dem Mittelwertsatz für reelwertige Funktionen einer Variabel $\exists \tau \in]0,1[$ s.d.

$$f(y) - f(x) = g(1) - g(0) = g'(\tau) \overset{\text{Kettenregel}}{=} \mathrm{d}\, f|_{\gamma(\tau)}(\dot\gamma(\tau)) = \mathrm{d}\, f|_{\gamma(\tau)}(y-x)$$

D.h. $\exists \xi \in [x,y]$ s.d.

$$f(y) - f(x) = \mathrm{d}\, f|_\xi (y-x) = \partial_{y-x} f(\xi) \tag{25}$$

Satz 2.23. *(Mittelwertsatz) U offen, $[x,y] \subset U$ und $f : U \to \mathbb{R}$ differenzierbar. Dann $\exists \xi \in]x,y[$ so das (25) gilt.*

Definition 2.24. Sei $U \subset \mathbb{R}^n$ eine Menge. U heisst sternförmig mit Zentrum $x_0 \in U$: wenn $[x_0, x] \subset U \forall x \in U$

Satz 2.25. *(Schrankensatz) Sei U eine offene Menge, die sternförmig ist und $f : U \to \mathbb{R}$ eine differenzierbare Funktion mit*

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

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

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

Sei (X,d) ein metrischer Raum. $f : (X,d) \to \mathbb{R}$ heisst Lipschitz falls $\exists L < \infty$ so dass

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

Korollar 2.27. *Sei U offen und konvex und $f : U \to \mathbb{R}$ eine differenzierbare Funktion mit beschränkten partiellen Ableitungen. Dann ist f Lipschitz.*

2.7 Höhere partielle Ableitungen

Sei

$$f : \mathbb{R}^n \supset \Omega \to \mathbb{R}$$

Die partiellen Ableitungen von f:

$$\frac{\partial f}{\partial x_i}(x) = \lim_{\varepsilon\to 0} \frac{f(x+\varepsilon e_i) - f(x)}{\varepsilon} \qquad \text{wobei } e_i = (0, \cdots, 1, \cdots, 0)$$

Falls die partielle Ableitung $\dfrac{\partial f}{\partial x_i}$ überall existiert dann bekommen wir eine neue Funktion

$$\Omega \ni x \mapsto \frac{\partial f}{\partial x_i} \in \mathbb{R}\,.$$

Wir können diese neune Funktion noch ableiten. Wir definieren

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(x) := \frac{\partial\left(\dfrac{\partial f}{\partial x_i}\right)}{\partial x_j}(x) = \lim_{\varepsilon\downarrow 0} \frac{\dfrac{\partial f}{\partial x_i}(x+\varepsilon e_j) - \dfrac{\partial f}{\partial x_i}(x)}{\varepsilon}\,.$$

Wenn auch diese überall existiert, können wir noch ableiten:

$$\frac{\partial^3 f}{\partial x_k \partial x_j \partial x_i}(x) := \lim_{\varepsilon\downarrow 0} \frac{\frac{\partial^2 f}{\partial x_i}(x+\varepsilon e_j) - \frac{\partial^2 f}{\partial x_j \partial x_i}(x)}{\varepsilon}\,.$$

Und so weiter. Die Anzahl Ableitungen dir wir nehmen ist die *Ordnung* der höheren Partiellen Ableitung. D.h.

$$\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}$$

ist eine partielle Ableitung mit Ordnung k.

Ausserdem wir nutzen die Notation

$$\frac{\partial^2 f}{\partial x_i^2} = \frac{\partial^2 f}{\partial x_i \partial x_i} \quad \frac{\partial^3 f}{\partial x_i^3} = \frac{\partial^3 f}{\partial x_i \partial x_i \partial x_i}$$

und so weiter.

Satz 2.28 (Lemma von Schwarz)**.** *Sei $f : \Omega \to \mathbb{R}$ eine Funktion die in einer Umgebung von $p \in \Omega$ die partielle Ableitungen $\dfrac{\partial f}{\partial x_i}$, $\dfrac{\partial f}{\partial x_j}$ 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 2.29. Wir kontrollieren die Plausibilität dieses Satzes mit einer ziemlichen grossen Famile von Funktionen: Die Polynome. Sei

$$f(x_1, x_2) = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} a_{ij} x_1^i x_2^j$$

Dann wir können explizit die folgenden partiellen Ableitungen rechnen:

$$\frac{\partial f}{\partial x_1} = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} i a_{ij} x_1^{i-1} x_2^j$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} ij a_{ij} x_1^{i-1} x_2^{j-1}$$

$$\frac{\partial f}{\partial x_2} = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} j a_{ij} x_1^{i} x_2^{j-1}$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} ij a_{ij} x_1^{i-1} x_2^{j-1}\,.$$

Beispiel 2.30. Aber, ohne gewisse Annahmen, ist der Satz falsch. Sei zum Beispiel $V : \mathbb{R} \to \mathbb{R}$ eine Funktion die nicht differenzierbar ist und definieren wir

$$v : \mathbb{R}^2 \to \mathbb{R} \qquad v(x_1, x_2) = V(x_2)$$

Dann,

$$\frac{\partial f}{\partial x_1} = 0 \qquad \text{und} \qquad \frac{\partial^2 f}{\partial x_2 \partial x_1} = 0\,.$$

Aber $\dfrac{\partial f}{\partial x_2}$ existiert nicht und deswegen auch $\frac{\partial^2 f}{\partial x_1 \partial x_2}$ nocht existiert.

Beweis. [Beweis des Lemmas von Schwarz] Die Idee ist ein Ärt von Mittelwertsatzßu benutzen.

Schritt 1 Von Dimension $n \to 2$

$$f(x_1, \cdots, x_i, \cdots, x_j, \cdots, x_n)$$

$$p = (p_1, \cdots, p_i, \cdots, p_j, \cdots, p_n)$$

Wir definieren $g : \mathbb{R}^2 \supset U \to \mathbb{R}$ als

$$g(y, z) = g(p_1, \cdots, p_{i-1}, y, p_{i+1}, \cdots, p_{j-1}, z, p_{j+1}, \cdots, p_n)\,.$$

Dann,

$$\frac{\partial f}{\partial x_i}(p) = \frac{\partial g}{\partial y}(p_i, p_j) \qquad \frac{\partial f}{\partial x_j}(p) = \frac{\partial g}{\partial z}(p_i, p_j)$$

$$\frac{\partial f}{\partial x_j \partial x_i} = \frac{\partial^2 g}{\partial z \partial y}(p_i, p_j) \qquad \frac{\partial f}{\partial x_i \partial x_j}(p) = \frac{\partial g}{\partial y \partial z}(p_i, p_j)$$

(Wir rechnen zum Beispiel

$$\begin{aligned}\frac{\partial f}{\partial x_i}(p) &= \lim_{\varepsilon to 0} \frac{f(p_1,\ldots,p_i+\varepsilon,\ldots,p_j,\ldots p_n)-f(p)}{\varepsilon}\\ &= \lim_{\varepsilon to 0} \frac{g(p_1+\varepsilon,p_2)-g(p_1,p_2)}{\varepsilon} = \frac{\partial g}{\partial y}(p_i,p_j). \quad \Big).\end{aligned}$$

Deswegen, oBdA beweisen wir nun den Fall $n = 2$ des Satzes.

Schritt 2 Sei $f : \mathbb{R}^2 \supset \Omega \to \mathbb{R}$ und $(a,b) \in \Omega$. Wir wissen dass $\frac{\partial f}{\partial x_1}$, $\frac{\partial f}{\partial x_2}$ und $\frac{\partial^2 f}{\partial x_2 \partial x_1}$ in einer Umgebunv von $p = (a,b)$ existieren und $\frac{\partial^2 f}{\partial x_2 \partial x_1}$ stetig auf p ist. Zu beweisen: $\frac{\partial^2 f}{\partial x_2 \partial x_1}(p)$ existiert und

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

Für jede $h,k \in \mathbb{R}\setminus\{0\}$ wir definieren den Rechteck Q mit Ecken (a,b), $(a+h,b)$, $(a,b+k)$, $(a+h,b+k$. D.h. $Q = [a,a_h] \times [b,b+k]$. Wir definieren

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

und bemerken dass

$$\begin{aligned}\lim_{k\to 0}\lim_{h\to 0}\frac{D_Q f}{hk} &= \lim_{k\to 0}\lim_{h\to 0}\frac{f(a+h,b+k)-f(a,b+k)}{hk} - \frac{f(a+h,b)-f(a,b)}{hk}\\ &= \lim_{k\to 0}\frac{\frac{\partial f}{\partial x_1}(a,b+k)-\frac{\partial f}{\partial x_1}(a,b)}{k} = \frac{\partial^2 f}{\partial x_2\partial x_1}(a,b) \qquad (26)\end{aligned}$$

und

$$\lim_{h\to 0}\left(\lim_{k\to 0}\frac{D_Q f}{hk}\right) = \lim_{h\to 0}\frac{\frac{\partial f}{\partial x_2}(a+h,b)-\frac{\partial f}{\partial x_2}(a,b)}{h}\,. \qquad (27)$$

Die Existenz des Grenzwerts in (27) impliziert die Existenz der partiellen Ableitung $\frac{\partial^2 f}{\partial x_1 \partial_2}(a,b)$. In diesem Fall ist es auch

$$\lim_{h\to 0}\lim_{k\to 0}\frac{D_Q f}{hk} = \frac{\partial^2 f}{\partial x_1 \partial x_2}(a,b)$$

Wir werden nun die existenz dieses zweiten Grenzwerts beweisen. Gleichzeitig erhalten wir dass die Grenzwerte in (26) und (27) gleich sind (i.e. wir können "h und k im Grenzwert vertauschen").

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

$$\frac{D_Q f}{hk} = \frac{\partial^2 f}{\partial x_2 \partial x_1}(\xi,\zeta) \qquad (28)$$

Das folgt wenn wir zwei Mal den Mittelwertsatz anwenden. OBdA nehmem wir $h,k > 0$ an. Dann

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

wobei

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

und ξ eine Stelle in $]X,X+h[$ ist. g ist in der Tat differenzierbar und

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

Deswegen, wenn wir einen zweiten Mal den Mittelwertsatz anwenden,

$$\begin{aligned}\frac{D_Q f}{hk} &= \frac{1}{k}\left(\frac{\partial f}{\partial x_1}(\xi, b+k) - \frac{\partial f}{\partial x_1}(\xi, b)\right) \\ &= \frac{\partial f}{\partial x_2}\left(\frac{\partial f}{\partial x_1}\right)(\xi,\zeta) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(\xi,\zeta)\,.\end{aligned}$$

Nun nutzen wir die Stetigkeit der Funktion $\frac{\partial^2 f}{\partial x_2 \partial x_1}$:

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

□

3 Das Taylorpolynom

Definition 3.1. Sei nun $a \in \Omega \subset \mathbb{R}^n$, $f : \Omega \to \mathbb{R}$ und $w \in \mathbb{R}^n$. Falls die ganzen Ableitungen mit Ordnung k in a existieren, dann definieren wir

$$:= \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} \frac{\partial^k f(a)}{\partial x_{i_1}\cdots \partial x_{i_k}} w_{i_1}\cdots w_{i_k}$$

und das Taylor Polynom

$$T^k_x f(z) = f(x) + \mathrm{d}\, f|_x (z-x) + \cdots + \frac{1}{k!}\,\mathrm{d}\, f^{(k)}|_x (z-x)^k$$

Definition 3.2. Eine Funktion $f : \Omega \to \mathbb{R}$ heisst C^k falls die ganzen partiellen Ableitungen mit Ordnung $\leq k$ überall existieren und stetig sind.

Satz 3.3 (Verallgemeinerte Lagrange Fehlerabschätzung)**.** *Sei $f \in C^{k+1}$ und $K_r(a) \in \Omega$. Dann, $\forall x \in K_r(a)$ $\exists \xi \in [x,k]$ s.d.*

$$R^k_a f(x) := f(x) - T^k_x f(x) = \frac{1}{(k+1)!}\,\mathrm{d}\, f^{(k+1)}|_\xi (x-a)^{k+1}\,. \tag{29}$$

Falls $f \in C^k$, dann $f(x) - T^k_x f(x) = o(\|x\|^k)$.

Beweis. **Teil 1: Beweis von** (29) Sei $g(t) := f(tx + (1-t)a)$. Wir wenden die Kettenregel $k+1$ Mal und rechnen:

$$\begin{aligned} g'(t) &= df|_{tx+(1-t)a}(x-a) \\ g''(t) &= d^2 f|_{tx+(1-t)a}(x-a)^2 \\ &\cdots \\ g^{(k+1)}(t) &= d^{(k+1)} f|_{tx+(1-t)a}(x-a)^{k+1}\end{aligned} \tag{30}$$

Die Lagrange Fehlerabschätzung für Funktionen einer Variable gibt die existenz einer Stelle $\tau \in]0,1[$ s.d.

$$g(1) = \sum_{i=0}^{k} \frac{1}{i!} g^{(i)}(0) + \frac{1}{(k+1)!} g^{(k+1)}(\tau)\,. \tag{31}$$

(Zur Erinnerung: wir nutzen die Konvention Konvention $g^{(0)}(0) = g(0)$ und deswegen

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

Die Stelle $\xi := \tau x + (1-\tau)a$ liegt auf dem Segment $[a,x]$. Mit den Formeln (30) schreiben wir (31) als

$$\begin{aligned} f(x) &= g(1) = \sum_{i=0}^{i} \frac{1}{i!} df^{(i)}|_a (x-a)^i + \frac{1}{(k+1)!} df^{(k+1)}|_\xi (a-x)^{k+1} \\ &= T_a^k f(x) + \frac{1}{(k+1)!} df^{(k+1)}|_\xi (a-x)^{k+1} \end{aligned}$$

Teil 2 Sei nun $f \in C^k$. Wir nuzten (29) und (für $x \in K_r(a)$) schreiben

$$f(x) = T_a^{k_1} f(x) + \frac{1}{k!} df^{(k)}|_\xi (a-x)^k \,. \tag{32}$$

Deswegen

$$\begin{aligned} & \left|f(x) - T_a^k f(x)\right| = \left| \frac{1}{k!} df^{(k)}|_a (a-x)^k - \frac{1}{k!} df^{(k)}|_\xi (a-x)^k \right| \\ = & \frac{1}{k!} \left| \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} \left(\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(a) - \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(\xi) \right) (x_{i_1} - a_{i_1}) \dots (x_{i_k} - a_{i_k}) \right| \\ \leq & \frac{\|x-a\|^k}{k!} \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} \left| \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(a) - \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(\xi) \right| \,. \end{aligned} \tag{33}$$

Die Stetigkeit der partiellen Ableitungen impliziert

$$\lim_{\xi \to 0} \left(\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(a) - \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(\xi) \right) = 0 \,.$$

$x \to a$ impliziert $\xi \to 0$ und aus (33) schliessen wir

$$\lim_{x \to a} \frac{\left|f(x) - T_a^k f(x)\right|}{\|x-a\|^k} = 0 \,.$$

$\square$

Falls f beliebig mal differenzierbar ist (in diesem Fall schreiben wir $f \in C^\infty(\Omega)$; d.h. die ganzen partiellen Ableitungen existieren und sind stetig), können wir die Taylorreihe schreiben:

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

Definition 3.4. Eine Funktion $f \in C^\infty(\Omega)$ heisst analytisch wenn $\forall x \in \Omega$ $\exists B_r(x) \subset \Omega$ mit der Eigenschaft dass:

$$T_x(z) = f(z) \;\; \forall z \in B_r(x) \,.$$

In diesem Fall schreiben wir $f \in C^\omega(\Omega)$.

3.1 Das Taylorpolynom zweiter Ordnung

Wir schreiben noch ein Mal die Approximation mit dem Taylonrpolynom zweiter Ordnung für eine C^2 Funktion:

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

$$+ \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x)(z_i - x_i)(z_j - x_j) + R(z) \,. \tag{34}$$

Aus dem Satz 3.3 wissen wir dass $R(z) = o(\|z-x\|^2)$. Falls $f \in C^3$ dann wissen wir noch mehr: $R(z) = O(\|z-x\|^3)$ (wir führen hier eine neue Notation ein: wenn g eine nichtnegative Funktion ist, die Schreibung $R(z) = O(g(z))$ bedeutet die Existenz einer Umgebung U von x und einer Konstant C s.d. $|R(z)| \leq Cg(z)$ $\forall z \in U$).

Wir definieren die Hessche 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 3.5. Schwarz $\implies$ $Hf(x)$ ist symmetrisch wenn alle Ableitungen zweiter Ordnung stetig sind.

Wir rechnen

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

und deswegen

$$\sum_{j=1}^{n}(z_j - x_j)\sum_{i_1}^{n}\frac{\partial^2 f}{\partial x_j \partial x_i}(x)(z_j - x_j)$$
$$= \langle z-x, Hf(x)(z-x)\rangle = (z-x)^T Hf(x)(z-x)$$

Wenn A eine $n \times n$ Matrix, die Abbildung

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

ist eine "quadratische Form" auf $\mathbb{R}^n$. $w^T A w$ ist das Matrix Produkt der: $1 \times n$ Matrix w^T ("eine Zeile"), $n \times n$ Matrix A und $n \times 1$ Matrix w ("eine Spalte").

Das Taylorpolynom zweiter Ordnung ist dann

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

Korollar 3.6. *Falls $f \in C^3(\Omega)$ und $B_r(x) \subset \Omega$*

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

d.h.

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

Korollar 3.7. *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.

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

Beweis. Die Taylorapproximation mit Ordnung 1:

$$f(z) = T_x^1 f(z) + \frac{1}{2}(z-x)^T Hf(\zeta)(z-x)$$

Dann,

$$f(z) - T_x^2 f(z) = \frac{1}{2}(z-x)^T Hf(\zeta)(z-x) - \frac{1}{2}(z-x)^T Hf(x)(z-x)$$

$$= \frac{1}{2}(z-x)^T(Hf(\zeta)-Hf(x))(z-x)$$
$$\leq \frac{1}{2}\|z-x\|\,\|Hf(\zeta)-Hf(x)(z-x)\|$$
$$\leq \frac{1}{2}\|z-x\|\,\|Hf(\zeta)-Hf(x)\|_O\,\|z-x\|$$
$$= \frac{1}{2}\|z-x\|^2\,\|Hf(\zeta)-Hf(x)\|_O$$
$$\frac{\left|f(z-)-T_x^2 f(z)\right|}{\|z-x\|^2} \leq \frac{1}{2}\|Hf(\zeta)-Hf(x)\|_O$$
$$\|\zeta-x\| \leq \|z-x\|^x$$

Stetigkeit der Ableitungen 2. Ordnung

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

□

Definition 3.8. $X \subset \mathbb{R}^n$, $f : X \to \mathbb{R}$. f hat in $a \in X$ ein lokales Minimum/Maximum

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

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

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

Satz 3.9. *(Notwendiges Kriterium für lokale Extrema). Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine partiell differentierbare Funktion die ein Extremum in $a \in U$ hat. Dann gilt*

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

D.h. wenn f differenzierbar ist, dann gilt $\mathrm{d}f|_a = 0$

Beweis. $F(t) = f(a+te_i)$ (für t sehr klein, so dass $a+te_i \in U$). F hat ein lokales Extremum in 0, d.h. $F'(0) = \partial_i f(a) = 0$. □

Definition 3.10. Sei f differenzierbar. Eine Stelle a mit $\mathrm{d}f|_a = 0$ heisst <u>kritischer Punkt</u>. Man sagt auch f ist <u>stationär</u> in a.

Bemerkung 3.11. Lokale Extremum $\implies \not\Leftarrow$ kritischer Punkt.

Satz 3.12. *(Hinreichendes Kriterium für lokale Extrema) $U \subset \mathbb{R}^n$ offen, $f \in C^2(U,\mathbb{R})$ mit $\mathrm{d}f|_a = 0$. Dann*

$$Hf(a) > 0 \implies a \text{ lokales Minimum}$$
$$Hf(a) < 0 \implies a \text{ lokales Maximum}$$
$$Hf(a) \text{ indefinit} \implies a \text{ kein Extremum}$$

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 3.13.
- $Hf(a) > 0$ bedeutet $H_f(a)$ <u>positiv definit</u>, d.h.

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

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

$$v^t Hf(a)v > 0$$
$$w^t Hf(a)w < 0$$

Beweis.

$H_f(a) > 0$

$$\mathrm{d}\, f|_a = 0 \xRightarrow{\text{Taylor}} f(a+h) = f(a) + \frac{1}{2} h^T H_f(a) h + R(h)$$

mit

$$\frac{R(h)}{\|h\|^2} \to 0 \quad \text{wenn } \|h\| \to 0.$$

$f \in C^2$

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

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

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

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

d.h. f hat in a ein lokales Minimum

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

$Hf(a) <> 0$

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

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

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

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

$\implies$ Beh

$\square$

Bemerkung 3.14. 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 3.15. $f(x,y) = y^2(x-1) + x^2(x+1)$

$$\mathrm{d}\, f|_{(x,y)} = (y^2 + 3x^2 + 2x, 2(x-1)y)$$

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

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

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

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

3.2 Konvexität

Definition 3.17. $U \subset \mathbb{R}^n$ heisst konvex

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

Definition 3.18. $f: U \to \mathbb{R}$ heisst konvex

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

- 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 3.19. f ist konvex

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

Satz 3.20. *(Konvexitätskriterium) Sei $f: U \to \mathbb{R}, C^2$ $U \subset \mathbb{R}^n$ offen, konvex. Es gilt:*

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

Bemerkung 3.21. Umkehrung von 1 gilt nicht, z.B. $f(x,y) = x^4 + y^4$

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

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

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

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

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

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

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

2. Analog wie die zweite Richtung im Ersten.

□

4 Differentation parameterabhängiger Integrale

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

$$F(x) = \int_a^b f(x,t) \,\mathrm{d}\,t$$

Satz 4.1. *(Differentationssatz) Falls*

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

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

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

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

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

Der Satz bedeutet also dass, mit den obigen Annahmen, dürfen wir die Ableitung und das Integral vertauschen:

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

Beweis. Sei $x \in U$ und $e_i = (0, \cdots, \underbrace{1}_{ite\ \text{Stelle}}, \cdots, 0)$. Wir rechnen

$$\begin{aligned}
\frac{\partial F}{\partial x_i}(x) &= \lim_{\varepsilon\to 0}\frac{F(x+\varepsilon e_i)-F(x)}{\varepsilon}\\
&= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left\{\int_a^b f(x+\varepsilon e_i,t)\,\mathrm{d}\,t - \int_a^b f(x,t)\,\mathrm{d}\,t\right\}\\
&= \lim_{\varepsilon\to 0}\int_a^b \frac{f(x+\varepsilon e_i,t)-f(x,t)}{\varepsilon}\,\mathrm{d}\,t
\end{aligned}$$

Deswegen

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

$$\iff \lim_{\varepsilon\to 0}\left\{\int_a^b \frac{f(x+\varepsilon e_i,t)-f(x,t)}{\varepsilon}\,\mathrm{d}\,t - \int_a^b \frac{\partial f}{\partial x_i}(x,t)\,\mathrm{d}\,t\right\} = 0$$

$$\iff \lim_{\varepsilon\to 0}\left\{\int_a^b \left[\frac{f(x+\varepsilon e_i,t)-f(x,t)}{\varepsilon} - \frac{\partial f}{\partial x_i}(x,t)\,\mathrm{d}\,t\right]\right\} = 0$$

Wir behaupten mehr, d.h.

$$A(\varepsilon) := \int_a^b \left| \underbrace{\frac{f(x+\varepsilon e_i,t)-f(x,t)}{\varepsilon}}_{(\text{Mittelwetsatz})=\frac{\partial f}{\partial x_i}(\xi_\varepsilon(t),t)} - \frac{\partial f}{\partial x_i}(x,t)\right|\mathrm{d}\,t \overset{\varepsilon\to 0}{\to} 0$$

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

$$A(\varepsilon) = \int_a^b \left|\frac{\partial f}{\partial x_i}\,(\xi(t),t) - \frac{\partial f}{\partial x_i}(x,t)\right|\mathrm{d}\,t\,.$$

Wir bemerken dass

$$\lim_{\varepsilon\to 0}\xi_\varepsilon(t) = x$$

und, wegen der Stetigkeit von $\dfrac{\partial f}{\partial x_i}$,

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

In der Tat ist diese Konvergenz gleichmässig, d.h.

Behauptung 4.2. $\forall\delta > 0\ \exists\varepsilon_0 > 0$ *so dass*

$$|\varepsilon| \le \varepsilon_0 \implies \sup_{t\in[a,b]}\left|\frac{\partial f}{\partial x_i}(\xi_\varepsilon(t),t) - \frac{\partial f}{\partial x_i}(x,t)\right| < \delta$$

$\implies$

$$\limsup_{\varepsilon\to 0} A(\varepsilon) \le \sup_{|\varepsilon|<\varepsilon_0} A(\varepsilon) \le \int_a^b \delta\,\mathrm{d}\,t = \delta(b-a)$$

δ ist beliebig

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

Die Behauptung 4.2 folgt aus dem nächsten Lemma. □

Lemma 4.3. *Sei $g : U \times [a,b] \to \mathbb{R}$ 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 $g_y \to g_x$ gleichmässig für $y \to x$.

Bemerkung 4.4. 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

$$\Bigg\| \underbrace{(\xi,\tau)}_{\in \mathbb{R}^{n+1}} - \underbrace{(x,t)}_{\in \mathbb{R}^{n+1}} \Bigg\| < \delta(x,t)$$

$$\|(\xi,\tau) - (x,t)\| = \sqrt{\|\xi - x\|^2 + (t-\tau)^2}$$

$\forall (x,t)$ Sei

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

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

$$\|(y,t) - (x,\tau)\| < \sqrt{\frac{1}{2}\delta^2(x,t) + \frac{1}{2}\delta^2(x,t)} = \delta(x,t)$$

$$\implies (y,t) \in B_{\delta(x,t)}(x,t)$$

Deswegen $U_{x,t} \subset B_{\delta(x,t)}(x,t)$. Wir bemerken dass K kompakt istm, weil

$$\mathbb{R} \ni t \mapsto (x,t)$$

eine stetige Funktion ist und K das Bild von $[a,b]$ durch diese Abbildung ist. $\{U_{x,t} : t \in [a,b]\}$ ist eine offene Überdeckung von K. Kompaktheit $\implies$ $\exists \{U_{x_i,t_i} : i \in \{1,\cdots,N\}\}$ Überdeckung von K. Sei

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

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

$$(x,t),(y,t) \in U_{x_i,t_i} \subset B_{\delta(x_i,t_i)}(x_i,t_i)$$

$$\implies |g(y,t) - g(x_i,t_i)| < \frac{\varepsilon}{10}$$

und

$$\implies |g(x,t) - g(x_i,t_i)| < \frac{\varepsilon}{10}$$

$$\implies |g(x,t) - g(y,t)| < \frac{\varepsilon}{5}$$

$$\implies \sup_{y \in B_\delta(x)} |g(x,t) - g(y-t)| \le \frac{\varepsilon}{5} < \varepsilon \quad \forall t \in [a,b]$$

□

Korollar 4.5. *Sei $g : U \times [a,b] \to \mathbb{R}$ stetig. Dann*

$$F(x) = \int_a^b g(x,t)\,\mathrm{d}\,t$$

ist eine stetige Funktion

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

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

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

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

□

Bemerkung 4.6. Im Differentiationssatz ist $\dfrac{\partial f}{\partial x_i}$ eine stetige Funktion. Da

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

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

Der folgende Satz ist eine sehr wichtige Konsequenz der Differentiationssatz.

Satz 4.7. *Sei $f : \mathbb{R}^2 \supset U \to \mathbb{R}$ eine stetige Funktion. Sei $R = [a,b] \times [c,d] \subset U$. Dann:*

$$\int_a^b \int_c^d f(s,t)\,\mathrm{d}\,t\,\mathrm{d}\,s = \int_c^d \int_a^b f(s,t)\,\mathrm{d}\,s\,\mathrm{d}\,t$$

Beweis. Wir definieren

$$\Phi(x,y) = \int_a^x \int_c^y f(s,t)\,\mathrm{d}\,t\,\mathrm{d}\,s$$

$$\Psi(x,y) = \int_c^y \int_a^x f(s,t)\,\mathrm{d}\,s\,\mathrm{d}\,t$$

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

Schritt 1 Φ und Ψ sind stetig differenzierbar und $\nabla\Phi = \nabla\Psi$
(Kein Problem mit Definition. Die FUnktion sind wohldefiniert fur $(x,y) \in]a-\varepsilon, b+\varepsilon[\times]c-\varepsilon, d+\varepsilon[$ wobei $\varepsilon > 0$ klein genug ist). Sei y fixiert. wir setzen

$$\phi(x) = \int_c^y f(x,t)\,\mathrm{d}\,t$$

ϕ ist stetig wwil f stetig ist. Der Fundamentalsatz der Integralrechnung gibt

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

$\dfrac{\partial\Phi}{\partial x}$ ist eine stetige Funktion in der Variabel x. Wir beweisen nun dass $\dfrac{\partial\Phi}{\partial x}$ eine stetige Funktion von zwei Variablen ist. Sei (x_0, y_0), $\varepsilon > 0$. Dann $\exists\delta$

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

Sei x fixiert:

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

wobei M das Maximum von f ist.

Deswegen, für $\bar\delta \le \frac{\varepsilon}{2M}$

$$|y-y_0|<\bar\delta$$

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

Wenn

$$\|(x,y)-(x_0,y_0)\|] < \min\left\{\delta,\bar\delta\right\}$$

$\implies$ $|x-x_0|<\delta$ und $|y-y_0|<\bar\delta$

$$\left|\frac{\partial\Phi}{\partial x}(x,y)-\frac{\partial\Phi}{\partial x}(x_0,y_0)\right|$$

$$\le \left|\frac{\partial\Phi}{\partial x}(x,y)-\frac{\partial\Phi}{\partial x}(x,y_0)\right|+\left|\frac{\partial\Phi}{\partial x}(x,y_0)-\frac{\partial\Phi}{\partial x}(x_0,y_0)\right|<\frac{\varepsilon}{2}$$

Das gleiche Argument: $\dfrac{\partial\Psi}{\partial y}$ exisiert und ist stetig. Sei nun

$$\psi(x,y):=\int_a^x f(s,y)\,\mathrm{d}\,s$$

$$\frac{\partial\Psi}{\partial x}=\frac{\partial}{\partial x}\int_c^y \psi(x,t)\,\mathrm{d}\,t \overset{?}{=} \int_c^y \frac{\partial\psi}{\partial x}(x,t)\,\mathrm{d}\,t$$

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

$$\frac{\partial\psi}{\partial x}(x,t)=\frac{\partial}{\partial x}\int_a^x f(s,t)\,\mathrm{d}\,s \overset{\text{Fundamentalsatz}}{=} f(x,t) \tag{35}$$

$$\frac{\partial\Psi}{\partial x}=\int_c^y f(x,t)\,\mathrm{d}\,t \overset{!}{=} \frac{\partial\Phi}{\partial x}$$

Das gleiche Argument impliziert auch $\dfrac{\partial\Psi}{\partial x}=\dfrac{\partial\Phi}{\partial x}$.

Schritt 2. $\Phi=\Psi$ Sei $\alpha:=\Phi-\Psi$. Aus dem Schritt 1 wissen wir dass α differenzierbar ist und $\mathrm{d}\,\alpha=0$. Seien

$$(x_0,y_0),(x_1,y_1)\in[a-\varepsilon,b+\varepsilon[\times]c-\varepsilon,d+\varepsilon[$$

Da $[(x_0,y_0)(x_1,y_1)]$ im Definitionsbereich ist, wir können den Schrankensatz anwenden:

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

Deswegen

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

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

$y=d,x=b \implies$ den Satz. □

5 Differenzierbare Abbildungen

$f : \mathbb{R}^n \supset \Omega \to \mathbb{R}^m$

Definition 5.1. f ist in x_0 differenzierbar falls $\exists L : \mathbb{R}^n \to \mathbb{R}^m$ lineare Abbildung mit

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

d.h. wenn

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

dann

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

(In $\varepsilon - \delta$ Form:

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

D.h. "$\|R(h)\| \to 0$ schneller als $\|(\| h)$: $R(h) = o(\|h\|)$ in kleines-o-Notation. Deswegen

$$f \text{ diff in } \ x_0 \iff \exists L \lim \ \text{ mit } \ f(x_0+h) - f(x_0) + L(h) + o(\|h\|)\big) \,. \tag{36}$$

Bemerkung 5.2. f differenzierbar in $x_0 \implies$ stetig in x_0
f differenzierbar in $x_0 \implies \exists!$ lineare Abbildung die (36) erfüllt. Wir nennen L das Differential von f und bezeichnen es mit $\mathrm{d}\, f|_{x_0}$

$f : U \to \mathbb{R}^m$

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

$\forall i$ gibt es n verschieden partielle Ableitungen: $\dfrac{\partial f_i}{\partial x_j}$ n. DIe Matrixdarstellung einer linearen Abbildung $L : \mathbb{R}^n \to \mathbb{R}^m$ besteht as $m \times n$ Koeffizienten:

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

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

Wir definieren m lineare Abbildungen $\mathbb{L}_i : \mathbb{R}^n \to \mathbb{R}$ als $\mathbb{L}_i(x) = L_i \cdot x = \langle L_i, x\rangle$. Dann

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

Sei $f : U \to \mathbb{R}^m$ differenzierbar in x_0 und sei $L = \mathrm{d}\, f|_{x_0}$. Dann:

$$\frac{\overbrace{f(x_0+h) - f(x_0) - L(h)}^{A}}{\|h\|} \to 0 \tag{37}$$

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

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

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

Deswegen

$$(37) \iff \lim_{h \to 0} \frac{f_i(x_0+h) - f_i(x_0) - \mathbb{L}_i(h)}{\|h\|} = 0 \ \forall i \in \{1, \cdots, m\}$$

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

Im naechsten Satz fassen wir zusammen die Konsequenzen dieses Arguments.

Satz 5.3. *Sei $f : \underbrace{U}_{\subset \mathbb{R}^n} \to \mathbb{R}^m$ mit U offen und $f = (f_1, \cdots, f_m)$*

1. *f ist differenzierbar in x_0 $\iff$ f_i differenzierbar in x_0 $\forall i \in \{1, \cdots, m\}$*
2. $$\mathrm{d}\, f|_{x_0}(h) = \begin{pmatrix} \mathrm{d}\, f_1|_{x_0}(h) \\ \vdots \\ \mathrm{d}\, f_m|_{x_0} \end{pmatrix}$$
3. $$\mathrm{d}\, f|_{x_0}(h) = \begin{pmatrix} \nabla f_1(x_0) \cdot h \\ \vdots \\ \nabla f_n(x_0) \cdot h \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \frac{\partial f_1}{\partial x_2}(x_0) & \cdots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \frac{\partial f_i}{\partial x_j}(x_0) & \vdots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \cdots & \cdots & \frac{\partial f_m}{\partial x_n}(x_0) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}$$

   Das ist die Jacobi Matrix.

5.1 Differentiationregeln

Die erste Differentiationregel ist eine einfache Folgerung vom Satz 5.3.

Satz 5.4. *Seien $f, g; U \to \mathbb{R}^m$ 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 $\mathrm{d}\, f|_{x_0} + \mathrm{d}\, g|_{x_0}$

Seien nun $f : U \to \mathbb{R}^m$ und $g : U \to \mathbb{R}^m$ differenzierbar in x_0 und

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

Jede gf_i ist differenzierbar: die Funktion fg ist dann auch differenzierbar. Wir wollen $\mathrm{d}(gf)$ rechnen:

$$\mathrm{d}(gf) = \mathrm{d} \left. \begin{pmatrix} gf_1 \\ \vdots \\ gf_m \end{pmatrix} \right|_{x_0} (h) = \begin{pmatrix} \mathrm{d}(gf_1)|_{x_0}(h) \\ \vdots \\ \mathrm{d}(gf_m)|_{x_0}(h) \end{pmatrix}$$

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

Jacobi-Matrix

$$\begin{pmatrix} \frac{\partial g}{\partial x_1}(x_0) f_1(x_0) + g(x_0) \frac{\partial f_1}{\partial x_1}(x_0) & \cdots & \frac{\partial g}{\partial x_n}(x_0) f_1(x_0) + g(x_0) \frac{\partial f_1}{\partial x_n}(x_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial g}{\partial x_1}(x_0) f_m(x_0) + g(x_0) \frac{\partial f_m}{\partial x_1}(x_0) & \cdots & \frac{\partial g}{\partial x_n}(x_0) f_m(x_0) + g(x_0) \frac{\partial f_m}{\partial x_n}(x_0) \end{pmatrix}$$

$$= \begin{pmatrix} \frac{\partial g}{\partial x_1}(x_0) f_1(x_0) & \cdots & \frac{\partial g}{\partial x_n}(x_0) f_1(x_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial g}{\partial x_1}(x_0) f_m(x_0) & \cdots & \frac{\partial g}{\partial x_n}(x_0) f_m(x_0) \end{pmatrix}$$

$$+ \begin{pmatrix} g(x_0) \frac{\partial f_1}{\partial x_1}(x_0) & \cdots & g(x_0) \frac{\partial f_1}{\partial x_n}(x_0) \\ \vdots & \ddots & \vdots \\ g(x_0) \frac{\partial f_m}{\partial x_1}(x_0) & \cdots & g(x_0) \frac{\partial f_m}{\partial x_n}(x_0) \end{pmatrix}$$

$$\underbrace{\left(\frac{\partial g}{\partial x_j}(x_0) f_i(x_0) \right)}_{= f(x_0) \otimes \nabla g(x_0)} + \overbrace{\underbrace{g(x_0) \left(\frac{\partial f_i}{\partial x_j}(x_0) \right)}_{\text{Jacobi}}}^{A}$$

$$\mathrm{d}(gf)|_{x_0} = \underbrace{g(x_0)\, \mathrm{d}\, f|_{x_0}}_{A} + \overbrace{f(x_0) \otimes \mathrm{d}\, g|_{x_0}}^{\text{lineare Abbildung mit Rang 1}}$$

$$\mathrm{d}(gf)|_{x_0}(h) = g(x_0) \left[\mathrm{d}\, f|_{x_0}(h) \right] + \left[f(x_0) \right] \mathrm{d}\, g|_{x_0}(h)$$

5.2 Kettenregel

Satz 5.5. *Seien*

$$f : \underbrace{U}_{\subset \mathbb{R}^n} \to \underbrace{V}_{\subset \mathbb{R}^n} \quad \textit{und} \quad g : V \to \mathbb{R}^k .$$

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

$$\mathrm{d}(g \circ f)|_a = \mathrm{d}\, g|_b \circ \mathrm{d}\, f|_a \tag{38}$$

Beweis. Die Differenzierbarkeit von f in a bedeutet

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

Die Differential von g in b bedeutet

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

$$g(f(a+h)) = g(\underbrace{f(a)}_{b} + k) = g(b) + \mathrm{d}\, g|_b(k) + \bar{R}(k)$$

$$= g(b) + \mathrm{d}\, g|_b \left(\mathrm{d}\, f|_a(h) + R(h) \right) + \bar{R}(k)$$

Linearität von $\mathrm{d}\, g|_b$

$$= \underbrace{g(b)}_{g \circ f(a)} + \underbrace{\mathrm{d}\, g|_b(\mathrm{d}\, f|_a(h))}_{\text{ist linear in } h} + \underbrace{\mathrm{d}\, g|_b(R(h)) + \bar{R}(k)}_{:= \rho(h)}$$

Wir werden zeigen dass
$$\rho(h) = o(\|h\|)\,.$$

Schritt 1. Linearität von $h \mapsto \mathrm{d}\,g|_b(\mathrm{d}\,f|_a(h))$.

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

Schritt 2: $\rho(h) = o(\|h\|)$.

$$\begin{aligned}
\frac{\rho(h)}{\|h\|} &\leq \frac{|\mathrm{d}\,g|_b(R(h))|}{\|h\|} + \frac{|\bar{R}(k)|}{\|h\|} \\
&\leq \frac{\|\mathrm{d}\,g|_b\|_0\,\|R(h)\|}{\|h\|} + \frac{\left\|\bar{R}(k)\right\|}{\|h\|} \qquad (39)
\end{aligned}$$

Wir wissen dass
$$\frac{\|R(h)\|}{\|h\|} \to 0$$
und deswegen konvergiert der erste Teil von (39) zu null. Ausserdem
$$\frac{\left\|\bar{R}(k)\right\|}{\|h\|} = \begin{cases} 0 & \text{falls } \ k = 0 \\ \frac{\left\|\bar{R}(k)\right\|}{\|k\|}\,\frac{\|k\|}{\|h\|} & \end{cases}$$
$$\begin{aligned}
\|k\| = \|\mathrm{d}\,f|_a(h) + R(h)\| &\leq \|\mathrm{d}\,f|_a(h)\| + \|R(h)\| \\
&\leq \|\mathrm{d}\,f|_a\|_0\,\|h\| + \|R(h)\| \qquad (40)
\end{aligned}$$

Da
$$\frac{\|R(h)\|}{\|h\|} \to 0$$
$\exists \delta > 0$ so dass
$$\|h\| < \delta \implies \frac{\|R(h)\|}{\|h\|}$$
Falls $\|h\| < \delta$
$$\|k\| \leq (\|\mathrm{d}\,f|_a\|_0 + 1)\,\|h\|$$
Deswegen: wenn $\|h\| \to 0$, dann $\|k\| \to 0$ und für $\|h\| < \delta$
$$\frac{\left\|\bar{R}(k)\right\|}{\|h\|} \leq \underbrace{\frac{\left\|\bar{R}(k)\right\|}{\|k\|}}_{\to 0 \ \text{ für } \ \|h\| \to 0} (\|\mathrm{d}\,f|_a\|_0 + 1)$$

Deswegen:
$$\begin{aligned}
0 &\leq \limsup_{\|h\| \to 0} \frac{\|\rho(h)\|}{\|h\|} \\
&\leq \lim_{\|h\| \to 0} \frac{\left\|\bar{R}(k)\right\|}{\|h\|} + \lim_{\|h\| \to 0} \frac{\|\mathrm{d}\,g|_b(R(h))\|}{\|h\|} = 0 + 0 = 0 \\
&\implies \lim_{\|h\| \to 0} \frac{\|\rho(h)\|}{\|h\|} = 0
\end{aligned}$$

□

Bemerkung 5.6. Sei $n = m = k = 1$. $b = f(a)$

$$\mathrm{d}\, f|_a(h) = f'(a)h$$

$$\mathrm{d}\, g|_b(k) = g'(b)k$$

$$\begin{aligned}\mathrm{d}\, g|_b \circ \mathrm{d}\, f|_a(h) &= \mathrm{d}\, g|_b(\mathrm{d}\, f|_a(h)) = \mathrm{d}\, g|_b(f'(a)h)\\ &= g'(b)f'(a)h = g'(f(a))f'(a)h\end{aligned} \tag{41}$$

$\phi = g \circ f$

$$\mathrm{d}\, \phi|_a(h) = \phi'(a)h = (g \circ f)'(a)h$$

(38) (d.h. die allgemeine Kettenregel) impliziert

$$\begin{aligned}\mathrm{d}\, \phi|_a(h) \quad &= \quad \mathrm{d}(g \circ f)|_a(h) = \mathrm{d}\, g|_b \circ \mathrm{d}\, f|_a(h)\\ &\overset{(41)}{=} \quad g'(f(a))f'(a)h\end{aligned}$$

$$\begin{aligned}\implies &(g \circ f)'(a) \cancel{h} = g'(f(a))f'(a) \cancel{h}\\ \implies &\underbrace{(g \circ f)'(a) = g'(f(a))f'(a)}_{\text{alte Kettenregel}}\end{aligned}$$

Bemerkung 5.7. Kettenregel für die Jacobi-Matrizen. Sei M die Jacobi-Matrix für $\mathrm{d}\, g|_{b(=f(a)}$ und N düe für $\mathrm{d}\, f_a$. Die Jacobi für $\mathrm{d}(g \circ f)|_a$ ist MN
$\implies g = (g_1, \cdots, g_k)$ $f = (f_1, \cdots, f_m)$ Es gibt eine Formel für $\dfrac{\partial (g \circ f)_i}{\partial x_j}$

$$\mathrm{d}\, g_b \circ \mathrm{d}\, g|_a(w) = \mathrm{d}\, g|_b(\underbrace{\mathrm{d}\, f|_a(w)}_{v})$$

$$\begin{aligned}\mathrm{d}\, g|_b \circ \mathrm{d}\, f|_a(w) &= \mathrm{d}\, g|_b(v)\\ = \left(\sum_{i=1}^m M_{1i}v_i, \sum_{i=1}^m M_{2i}v_i, \cdots, \sum_{i=1}^m M_{ki}v_i\right)&\\ = \left(\sum_{i=1}^m M_{1i}\sum_{j=1}^n N_{ij}w_j, \cdots, \sum_{i=1}^m M_{ki}\sum_{j=1}^n N_{ij}w_j\right)&\end{aligned}$$

$$v = \mathrm{d}\, f|_a(w) = \left(\sum_{j=1}^n N_{1j}w_j, \cdots, \sum_{j=1}^n N_{mj}v_j\right)$$

$$\iff v_i = \sum_{j=1}^n N_{ij}w_j$$

$$\mathrm{d}\, g|_b \circ \mathrm{d}\, f|_a(v) = \left(\sum_{i=1}^m \sum_{j=1}^n M_{1i}N_{ij}v_j, \cdots, \sum_{i=1}^m \sum_{j=1}^n M_{ki}N_{ij}v_j\right)$$

(Sei A die Matrix

$$A_{lj} = \sum_{i=1}^m M_{li}N_{ij} \iff A = M \cdot N$$

$$= \left(\sum_{j=1}^n A_{1j}v_j, \cdots, \sum_{j=1}^n A_{kj}v_j\right)$$

Deswegen ist A die Matrixdarstellung von

$$\mathrm{d}\, g|_b \circ \mathrm{d}\, f|_a = \mathrm{d}(g \circ f)|_a$$

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

Bemerkung 5.8. $f : U \to V \subset \mathbb{R}^m$ $f = (f_1, \cdots, f_m)$, $f_i(x) = f(x_1, \cdots, x_n)$
$g : V \to \mathbb{R}^k$ $g = (g_1, \cdots, g_k)$, $g_j(x) = g(y_1, \cdots, y_m)$

$$g \circ f(x) = (g_1(f(x)), \cdots, g_k(f(x)))$$
$$g_j(x) = g_j(f_1(x), \cdots, f_m(x))$$
$$g_j(y) = g_j(f(x_1, \cdots, x_n), \cdots, f_m(x_1, \cdots, x_n))$$
$$A_{lj} = \frac{\partial}{\partial x_j}(g_l \circ f)(a)$$
$$M_{li} = \frac{\partial g_l}{\partial y_i}(b) = \frac{\partial g_l}{\partial y_i}(f(a))$$
$$N_{ij} = \frac{\partial f_i}{\partial x_j}(a)$$

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

Korollar 5.9. *Sei $f : U \to V(\subset \mathbb{R}^m)$ und $\phi : V \to \mathbb{R}$ 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.

5.3 Schrankensatz

Definition 5.10. sei $f : U \to \mathbb{R}^m$ eine Abbildung. Wir schreiben $f \in C^k(U, \mathbb{R}^m)$ falls die partielle Ableitungen jeder f_i mit Ordnung $\leq k$ existieren und stetig sind ($f = (f_1, \cdots, f_m)$).

Satz 5.11. *Sei $\Omega \subset \mathbb{R}^n$ eine offene Menge, $f \in C^1(\Omega, \mathbb{R}^k)$ und $\gamma[a, b] \to \Omega$ eine $\mathbb{C}^1$ Kurve. Dann:*

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

Zur Erinnerung: $\gamma : [a, b] \to \Omega \subset \mathbb{R}^n$, $\gamma = (\gamma_1, \cdots, \gamma_n)$, $\dot{\gamma} = (\gamma_1', \cdots, \gamma_n')$.

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

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

Kettenregel

$$\mathrm{d}\, \phi|_t = \mathrm{d}\, f|_{\gamma(t)} \,\mathrm{d}\, \gamma|_t \tag{42}$$
$$\phi : [a, b] \to \mathbb{R}^k$$
$$\mathrm{d}\, \phi|_t : \mathbb{R} \to \mathbb{R}^k \text{ lineare Abbildung}$$

$\phi = (\phi_1, \cdots, \phi_k)$
$$\begin{pmatrix} \dfrac{\partial \phi_1}{\partial t} \\ \vdots \\ \dfrac{\partial \phi_k}{\partial t} \end{pmatrix} = \begin{pmatrix} \phi_1' \\ \vdots \\ \phi_k' \end{pmatrix} = \dot{\phi}$$

Sei $A(x)$ die Jacobi-Matridx für $\mathrm{d}\, f|_x$ (d.h. $A_{ij}(x) = \dfrac{\partial f}{\partial x_i}(x)$). Kettenregel:

$$\underbrace{\dot{\phi}(t) = A(\gamma(t)) \cdot \dot{\gamma}(t)}_{\text{Matrix-Darstellung von (42)}}$$

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

ϕ_i' ist eine stetige Funktion:

$$\phi_i'(t) = \sum_{j=1}^n A_{ij}(\gamma(t))\gamma_j'(t) = \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(\gamma(t))\gamma_j'(t)$$

Nun
$$\phi(b) - \phi(a) = \begin{pmatrix} \int_a^b \phi_1'(t)\,\mathrm{d}\,t \\ \vdots \\ \int_a^b \phi_k'(t)\,\mathrm{d}\,t \end{pmatrix}$$

und

$$\|f(\gamma(b)) - f(\gamma(a))\| = \|\phi(b) - \phi(a)\|^2 = \sqrt{\sum_{i=1}^k \left(\int_a^b \phi_i'(t)\,\mathrm{d}\,t \right)^2}$$

Wir brauchen nun die folgende“Dreiecksungleichun”:

$$\sqrt{\sum_{i=1}^k \left(\int_a^b \phi_i'(t)\,\mathrm{d}\,t \right)^2} \leq \int_a^b \|\dot{\phi}(t)\|\,\mathrm{d}\,t\,. \tag{43}$$

Diese Ungleichung folgt aus dem Lemma 5.13 unten. Mit der schreiben wir

$$\begin{aligned} \|f(\gamma(b)) - f(\gamma(a))\| &\leq \int_a^b \|\dot{\phi}(t)\|\,\mathrm{d}\,t = \int_a^b \|A(\gamma(t)) \cdot \dot{\gamma}(t)\|\,\mathrm{d}\,t \\ &\leq \int_a^b \|A(\gamma(t)\|_O \|\dot{\gamma}(t)\,\mathrm{d}\,t = \int_a^b \|df|_{\gamma(t)}\|_O \|\dot{\gamma}(t)\,\mathrm{d}\,t \\ &\leq \sup_{t \in [a,b]} \left\| \mathrm{d}\, f|_{\gamma(t)} \right\|_O \end{aligned}$$

□

Bemerkung 5.12. In der Tat $\sup_{t\in[a,b]} \left\|\mathrm{d}\, f|_{\gamma(t)}\right\|_O$ ist ein Maxim wegen der Stetigkeit der Abbildung $t \mapsto \left\|\mathrm{d}\, f|_{\gamma(t)}\right\|_O$.

Lemma 5.13. *Sei $g : [a, b] \to \mathbb{R}^k$ eine stetige Funktion. Dann*

$$\sqrt{\sum_{i=1}^k \left(\int_a^b g_i \right)^2} \leq \int_a^b \|g\|\;.$$

<u>*Dreiecksungleichung*</u>

Beweis. Sei $\varepsilon > 0$ und Treppenfunktion α_i so dass $g_i - \varepsilon \le \alpha_i \le g_i + \varepsilon$, $\alpha_i - \varepsilon \le g_i \le \alpha_i + \varepsilon$. Dann

$$\int_a^b \alpha_i - (b-a)\varepsilon \le \int_a^b g_i \le \int_a^b \alpha_i + (b-a)\varepsilon$$

d.h.

$$\left| \int_a^b g_i - \int_a^b \alpha_i \right| \le (b-a)\varepsilon$$

Deswegen

$$\left| \sqrt{\sum_{i=1}^k (\int g_i)^2} - \sqrt{\sum_{i=1}^k \int \alpha_i^2} \right| \le \sqrt{\sum_{i=1}^k \left(\int g_i - \int \alpha_i \right)^2} \le \sqrt{k}(b-a)\varepsilon \tag{44}$$

Sei nun $\alpha = (\alpha_1, \dots, \alpha_n)$. Dann

$$\left| \int_a^b \|g\| - \int_a^b \|\alpha\| \right| \le \int_a^b |\|g\| - \|\alpha\|| \le \int_a^b \|g - \alpha\| \le \int_a^b \sqrt{k}\varepsilon = \sqrt{k}(b-a)\varepsilon\,. \tag{45}$$

Wir werden bewesein dass

$$\sqrt{\sum \left(\int_a^b \alpha_i \right)^2} \le \int_a^b \|\alpha\| \tag{46}$$

(44), (45) und (46) implizieren

$$\begin{aligned} \sqrt{\sum \left(\int_a^b g_i \right)^2} &\le \sqrt{\sum \left(\int_a^b \alpha_i \right)^2} + (b-a)\sqrt{k}\varepsilon \\ &\le \int_a^b \|\alpha\| + (b-a)\sqrt{k}\varepsilon \le \int_a^b \|g\| + 2(b-a)\sqrt{k}\varepsilon \end{aligned}$$

Wenn $\varepsilon \downarrow 0$:

$$\sqrt{\sum \left(\int_a^b g_i \right)^2} \le \int_a^b \|g\|$$

Beweis von (46)**.** Ohne Beschränkung der Allgemeinheit: $\exists$ eine Zerteilung von $[a, b]$

$$a = c_0 < c_1 < \dots < c_N = b$$

so dass 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

$$a_j = \begin{pmatrix} a_{1,j} \\ \vdots \\ a_{k,j} \end{pmatrix}$$

$$\sqrt{\sum_{i=1}^k \left(\int_a^b \alpha_i \right)^2} = \sqrt{\sum_{i=1}^k \left(\sum_{j=1}^N |I_j| \alpha_{i,j} \right)^2} = \|a\|$$

wobei

$$a := \sum_{j=1}^N |I_j| a_j = \begin{pmatrix} \sum_{j=1}^N |I_j| \alpha_{1,j} \\ \vdots \\ \sum_{j=1}^N |I_j| \alpha_{1,j} \end{pmatrix}.$$

Deswegen

$$\begin{aligned} \|a\| &= \left\| \sum_{j=1}^{N} |I_j| a_j \right\| \overset{\text{Dreiecksungleichung}}{\leq} \sum_{j=1}^{N} \| |I_j| a_j \| \\ &= \sum_{j=1}^{N} |I_j| \, \|a_j\| = \int_a^b \|\alpha\| \end{aligned}$$

□

Korollar 5.14. *$f \in C^1(\Omega, \mathbb{R}^k)$ und $[p,q] \subset \Omega$. Dann:*

$$\|f(p) - f(q)\| \leq \max_{z \in [p,q]} \|\mathrm{d}\, f|_z\|_O \, \|p - q\|$$

Beweis. Wenden den Schrankensatz an f und $\gamma : [0,1] \to \Omega$ ist $\gamma(a) = (1 - s)p + sq$. Da $\dot{\gamma} = q - p$,

$$\|f(p) - f(q)\| \leq \max_{s \in [0,1]} \left\| \mathrm{d}\, f|_{\gamma(s)} \right\|_O \underbrace{\int_0^1 \|\dot{\gamma}(s)\| \,\mathrm{d}\, s}_{\|p-q\|} = \max_{z \in [p,q]} \|\mathrm{d}\, f|_z\|_O \, \|p - q\| \ .$$

□

5.4 Satz der lokalen Umkehrbarkeit

Satz 5.15. *Sei $\Phi : \underbrace{U}_{\subset \mathbb{R}^n} \to \mathbb{R}^n$ (U offene Menge) eine $\mathbb{C}^1$-Abbildung und sei $a \in U$ so dass $\mathrm{d}\, \Phi|_a$ umkehrbar ist. Dann $\exists U_0$ offene Umgebung von a so dass $V := \Phi(U_0)$ eine offene Umgebung von $\Phi(a)$ und die Einschränkung*

$$\Phi : U_0 \to V$$

ein Diffeomorphismus ist.

Lemma 5.16. *(Banachscher Fixpunktsatz) Sei $C \subset \mathbb{R}^n$ eine abgeschlossene Menge und sei $\phi : C \mapsto C$ eine Abbildung mit folgender Eigenschaft:*

$$\|\phi(x) - \phi(y)\| \leq \lambda \, \|x - y\| \qquad\qquad \forall x, y \in C$$

wobei $0 \geq \lambda < 1$ (unabhängig von x, y).
Dann $\exists x \in C$ so dass $\phi(x) = x$ (d.h. x ein Fixpunkt von ϕ ist.

Definition 5.17. Eine Abbildung

$$\phi : X \mapsto X \ \ \text{(mit } X \text{ metrischer Raum)}$$

heisst Kontraktion falls $\exists \lambda < 1$ so dass

$$\mathrm{d}(\phi(x), \phi(y)) \leq \lambda \, \mathrm{d}(x, y) \qquad\qquad \forall x, y \in X$$

5.4.1 Allgemeine Form des Fixpunktsatzes von Banach

Satz 5.18. *Jede Kontraktion auf einem <u>vollständigen</u> metrischen Raum besitzt einen Fixpunkt.*

Beweis. Sei $x_0 \in X$ (bzw. in $C \subset \mathbb{R}^n$)

$$\begin{pmatrix} x_1 = \phi(x_0) \\ x_2 = \phi(x_1) \\ \vdots \\ x_k = \phi(x_{k-1}) \end{pmatrix}$$

Behauptungen:

1. $\{x_k\}$ ist eine Cauchyfolge

$$\xRightarrow{\text{Vollständigkeit von } X} \exists x \lim_{k\to\infty} x_k$$

2. $\phi(x) = x$

$1 \implies 2$ weil

$$\phi(x) = \lim_{k\to\infty} \phi(x_k) = \lim_{x_{k+1}\to\infty} = x$$

□

Beweis.

$$\mathrm{d}(x_0, x_1) = M \geq 0$$

$$\begin{aligned} \mathrm{d}(x_{k+1}, x_k) &= \mathrm{d}\left(\phi(x_k), \phi(x_{k-1})\right) \\ &\leq \lambda\, \mathrm{d}(x_k, x_{k-1}) \leq \cdots \leq \lambda^2\, \mathrm{d}(x_{k-1}, x_{k-2}) \\ &\cdots \leq \lambda^k\, \mathrm{d}(x_1, x_0) = \lambda^k M \end{aligned}$$

$$\begin{aligned} &\mathrm{d}(x_{k+j} x_k) \\ &\leq \mathrm{d}(x_{k+j}, x_{k+j-1}) + \mathrm{d}(x_{k+j-1}, x_{k+j-2}) + \cdots + \mathrm{d}(x_{k+1}, x_k) \\ &\leq \lambda^{k+j-1} M + \lambda^{k+j-2} M + \cdots + M\lambda^k \end{aligned}$$

$$\begin{aligned} \mathrm{d}(x_{k+j}, x_k) &\leq M\lambda^k (1 + \lambda + \cdots + \lambda^{j-1}) \\ &< M\lambda^k \sum_{i=0}^{\infty} \lambda^i \\ &= \frac{M\lambda^k}{1-\lambda} \end{aligned}$$

Deswegen $\forall m > n \geq N$ ($\lambda^N \to$ für $N \to +\infty$

$$\mathrm{d}(x_m, x_n) \leq \frac{M}{1-\lambda}\lambda^N$$

$\forall \varepsilon > 0\ \exists N$ so dass

$$\frac{M\lambda^N}{1-\lambda} < \varepsilon$$

$$\implies \mathrm{d}(x_m, x_n) < \varepsilon \qquad\qquad \forall n > m \geq N$$

Das ist die Cauchyeigenschaft $\implies$ $\{x_k\}$ ist eine Cauchyfolge □

Beweis des Satzes

Schritt 1 Wir suchen eine Umgebung von W von $\Phi(a)$, wo wir immer ein Urbild von $\in W$ finden. D.h.

$$\Phi(x) = y \tag{47}$$

besitzt eine Lösung x.
OBdA nehmen wir an $a = 0$ und $\mathrm{d}\,\Phi|_a = \mathrm{Id}$ (In der Tat, nehmen wir an dass

$$L = \mathrm{d}\,\Phi|_a \neq \mathrm{Id}$$

Sei

$$\Phi' = L^{-1} \circ \Phi$$

und

$$\mathrm{d}\,\Phi'|_x = L^{-1} \circ \mathrm{d}\,\Phi|_x$$

$\implies$ Φ' ist eine $\mathbb{C}^1$-Funktion.

$$\mathrm{d}\,\Phi|_0 = L^{-1} \circ \mathrm{d}\,\Phi|_0 = L^{-1} \circ L = \mathrm{Id}$$

$\implies$ Satz an Φ' anwenden

$$\Psi'(\Phi'(x)) = x \implies \Psi'(L^{-1}(\Phi(x))) = x$$
$$\implies \Psi := \Phi' \circ L^{-1}$$

die gesuchte Umkehrung von Φ ist $V := (V'))$ Wir wollen zeigen dass, wenn $\|y - \Phi(0)\| < \delta$, dann die Gleichung 47 lösbar ist.

$$47 \iff \underbrace{y + x - \Phi(x)}_{x \mapsto \phi_y(x)} = x$$

$\phi_y : U \to \mathbb{R}^n$ $\exists \eta > 0$ so dass

$$\phi_y : \overline{B_\eta}(0) \mapsto \overline{B_\eta}(0)$$

eine Kontraktion ist.

1. ϕ_y bildet $\overline{B_y}(0)$ in $\overline{B_\eta}(0)$
2. $\|\phi_y(z) - \phi_y(w)\| \leq \frac{1}{2}\|z - w\|$

Das zweite:

$$\begin{aligned} &\|\phi_y(z) - \phi_y(w)\| \\ &= \|y + z - \Phi(z) - y - w + \Phi(w)\| \\ &= \|(\Phi(w) - \Phi(z)) - (w - z)\| \\ &= \left\| \underbrace{\Phi(w) - w}_{\Lambda(w)} - \underbrace{\Phi(z) - z}_{\Lambda(z)} \right\| \end{aligned}$$

Λ ist $\mathbb{C}^1$

$$\mathrm{d}\,\Lambda|_0 = \mathrm{d}\,\Phi|_0 - \mathrm{Id} = 0$$
$$\|\mathrm{d}\,|_0\|_{HS} = 0$$

$\implies$ $\exists \eta > 0$ so dass

$$B_{\leq \eta}(0) \ni x \implies \|\mathrm{d}\,\Lambda|_x\|_{HS} \leq \frac{1}{2}$$

$z, w \in \overline{B_\eta}(0)$ und $\in B_\eta(0)$

$$\begin{aligned} \|\phi_y(z) - \phi_y(w)\| &= \|\Lambda(z) - \Lambda(w)\| \\ &\overset{\text{Schrankensatz}}{\leq} \left(\max_{B_\eta(0)} \|\mathrm{d}\,\Lambda\|_O \right) \|z - w\| \\ &\frac{1}{2}\|z - w\| \end{aligned}$$

$$\phi_y(0) = y - \Phi(0) + 0 = y - \Phi(0)$$

$\delta = \frac{\eta}{2}$, $\|\phi_y(0)\| \leq \frac{1}{2}$. Sei $z \in \overline{B_\eta}(0)$

$$\begin{aligned} \|\phi_y(z)\| &\|\phi_y(z) - \phi_y(0)\| + \|\phi_y(0)\| \\ &< \|\phi_y(z) - \phi_y(0)\| + \frac{\eta}{2} \\ &\leq \frac{1}{2}\|z - 0\| + \frac{\eta}{2} \\ &\leq \frac{1}{2}\eta + \frac{1}{2}\eta \\ &= \eta \\ &\implies \|\phi_y(z)\| < \eta \end{aligned}$$

So
$$\phi_y : \overline{B_\eta}(0) \mapsto B_\eta(0)$$
Banach: $\forall y \in B_{\frac{\eta}{2}}(\Phi(0))$, $\exists x \in B_\eta(0)$ und $\in \overline{B_\eta}(0)$ mit
$$\phi_y(x) = x \iff \Phi(x) = y$$
Sei $V := B_\delta(\Phi(0))$ (offen und Umgebung von $\Phi(0)$)
$$\underbrace{B_\eta(0) \cap \Phi^{-1}(V)}_{\text{ist eine offene Menge}} = U_0 \qquad\qquad \text{(offen und Umgebung von 0)}$$
$$Phi : U_0 \to V$$

1. Φ ist surjektiv: $\forall y \in V$, $\exists x \in B_\eta(0)$ mit $\Phi(x) = y$
$$\implies x \in \Phi^{-1}(V) \cap B_\eta(0) = U_0$$

2. Φ ist injektiv
$$\begin{aligned}\|\Phi(x) - \Phi(z)\| &= \|(x + \Lambda(x)) - (z + \Lambda(z))\| \\ &\implies \|\Phi(x) - \Phi(z)\| \\ &\leq \|x - z\| - \|\Lambda(x) - \Lambda(z)\| \\ &\leq \|x - z\| - \frac{1}{2}\|x - z\| \\ &\leq \frac{1}{2}\|x - z\|\end{aligned}$$

$\implies$ Φ ist injektiv. (Alternativerweise wenn ϕ eine Kotraktion ist, der Fixpunkt ovn ϕ ist eindeutig: $\phi(p) = p$, $\phi(q) = q$
$$\begin{aligned}\mathrm{d}(p, q) &= \mathrm{d}(\phi(p), \phi(q)) \\ &\leq \lambda\, \mathrm{d}(p, q) \\ (1 - \lambda)\, \mathrm{d}(p - q) &\leq 0 \\ &\overset{\lambda<1}{\implies} \mathrm{d}(p, q) = 0 \\ &\implies p = q\end{aligned}$$

Schritt 2 Sei $\Phi : V \mapsto U_0$ die Umkehrfunktion von Φ. Ψ ist stetig. Seien $\xi, \zeta \in V$, $x = \Phi(\xi), z = \Phi(\zeta) \implies \Phi(x) = \xi$, $\Phi(z) = \zeta$. Aber:
$$\begin{aligned}&\|\Phi(x) - \Phi(z)\| \leq \frac{1}{2}\|x - z\| \\ \implies &\underbrace{2\|\xi - \zeta\| \geq \|\Psi(\xi) - \Phi(\zeta)\|}_{\text{Lipschitz-Bedingung für }\Phi\text{: stetig}}\end{aligned}$$

Schritt 3

Bemerkung 5.19. $\Phi : U_0 \to V$ ist differenzierbar und $\mathrm{d}\,\Phi|_x$ ist umkehbar $\forall x \in U_0$.
$$\Phi(x) = x - \Lambda(x)$$
$$\mathrm{d}\,\Phi|_x = \mathrm{Id} - \mathrm{d}\,\Lambda|_x$$
Wir wissen, dass
$$\|\mathrm{d}\,\Lambda|_x\|_{HS} \leq \frac{1}{2} \qquad\qquad \forall x \in U_0 \subset B_\eta(0)$$
$$\mathrm{d}\,\Phi|_x(v) = v - \mathrm{d}\,\Lambda|_x(v)$$

$$\begin{aligned}
&\|\mathrm{d}\,\Phi|_x(v)\| \\
&\geq \|v\| - \|\mathrm{d}\,\Lambda|_x(v)\| \\
&\geq \|v\| - \frac{1}{2}\|v\| \\
&\geq \frac{1}{2}\|v\|
\end{aligned}$$

$\implies \operatorname{Ker}(\mathrm{d}\,\Phi|_x) = \{0\} \implies \mathrm{d}\,\Phi|_x$ ist injektiv $\implies$ Surjektivität $\implies \mathrm{d}\,\Phi|_x$ ist umkehrbar

Lemma 5.20. *Falls $\Phi : U_0 \to V$ eine $\mathbb{C}^1$ umkehrbare Abbildung so dass*

- $\mathrm{d}\,\Phi|_x$ *umkehrbar* $\forall x \in U_0$ *ist*
- *die Umkehrfunktion $\Psi : V \to U_0$ stetig ist, dann ist auch Ψ eine $\mathbb{C}^1$ Abbildung.*