

MITTSCHRIEB

Analysis II

Sommersemester 2025

Emma Bach

Vorlesung gehalten von
Prof. Dr. Michael RŮŽIČKA

May 26, 2025

Inhalt

1	Wiederholung	2
2	Der Euklidische Raum	3
2.1	Abbildungen und Koordinatenfunktionen auf \mathbb{R}^n	5
2.2	Mehrdimensionale Ableitungen	7
2.3	Differenzierbarkeit	8

Chapter 1

Wiederholung

Chapter 2

Der Euklidische Raum

Lemma 2.1. Sei $(V, \langle -, - \rangle)$ ein euklidischer Vektorraum. Dann wird durch

$$\|u\| = \sqrt{\langle u, u \rangle}$$

auf V eine Norm erklärt. Diese bezeichnet man als die durch das Skalarprodukt induzierte Norm.

Definition 2.2. Sei $(V, \langle -, - \rangle)$ ein euklidischer Vektorraum. Die Vektoren $u, v \in V$ heißen **orthogonal**, wenn

$$\langle u, v \rangle = 0$$

ist. Für $u, v \in V \setminus \{0\}$ wird die reelle Zahl

$$\phi = \arccos \frac{\langle u, v \rangle}{\|u\| \|v\|}$$

als der Winkel zwischen u und v bezeichnet.

Anmerkung 2.3. Es gilt

$$\frac{|\langle u, v \rangle|}{\|u\| \|v\|} \leq 1$$

Lemma 2.4. Für $X = (x_1, \dots, x_n) \in \mathbb{R}^n$ sei

$$\|X\|_{\max} := \max\{|x_1|, \dots, |x_n|\}$$

Dann ist $\|\cdot\|_{\max}$ eine Norm auf \mathbb{R}^n und es gilt

$$\|X\|_{\max} \leq \|X\| \leq \sqrt{n} \|X\|_{\max}$$

Satz 2.5. Die Menge \mathbb{Q}^n der Punkte mit rational Koordinaten ist dicht in \mathbb{R}^n .

Beweis. Sei $X \in \mathbb{R}^n$ und $\varepsilon \in \mathbb{R}^+$. Da \mathbb{Q} dicht in \mathbb{R} ist gilt

$$\forall i \in \{1, \dots, n\} : \exists y_i \in \mathbb{Q} : |x_i - y_i| \leq \frac{\varepsilon}{\sqrt{n}}$$

Durch Lemma 2.4 folgt:

$$\|x - y\| \leq \sqrt{n} \|X - Y\| < \varepsilon$$

□

Satz 2.6. Sei $(X_k)_{k \in \mathbb{N}}$ eine Folge aus \mathbb{R}^n . Sei $X_k = (x_1^{(k)}, \dots, x_n^{(k)})$. Dann gilt:

$$\lim_{k \rightarrow \infty} X_k = X \Leftrightarrow \forall i : \lim_{k \rightarrow \infty} x_i^{(k)} = x_i$$

Insbesondere ist X_k eine Cauchyfolge, wenn die Komponenten Cauchyfolgen sind.

Beweis. $X_k \rightarrow X$, $i \in \{1, \dots, n\}$, $\varepsilon \in \mathbb{R}^+$. Dann gilt

$$\exists k_0 \in \mathbb{N} : \forall k \geq k_0 : \|X_k - X\| \leq \varepsilon \implies \forall i : \left| x_i^{(k)} - x_i \right| < \varepsilon \implies \lim_{k \rightarrow \infty} x_i^{(k)} = x_i$$

Und umgekehrt:

$$\forall i : x_i^{(k)} \rightarrow x_i, \varepsilon \in \mathbb{R}^+ \implies \exists k_0^i \in \mathbb{N} : \forall k \geq k_0^i \left| x_i^{(k)} - x_i \right| \leq \frac{\varepsilon}{\sqrt{n}}$$

$$k_0 := \max\{k_0^1, \dots, k_0^n\} \implies \forall k \geq k_0 : \left| x_i^{(k)} - x_i \right| < \frac{\varepsilon}{\sqrt{n}} \implies \|X_k - X\| \leq \sqrt{n} \|X_k - X\| < \varepsilon$$

□

Satz 2.7. Für konvergente Folgen $(X_k), (Y_k) \in \mathbb{R}^n$, $(\lambda_k) \in \mathbb{R}$ gilt:

$$\lim_{k \rightarrow \infty} (X_k + Y_k) = \lim_{k \rightarrow \infty} X_k + \lim_{k \rightarrow \infty} Y_k \quad (2.1)$$

$$\lim_{k \rightarrow \infty} \lambda_k X_k = \left(\lim_{k \rightarrow \infty} \lambda_k \right) \left(\lim_{k \rightarrow \infty} X_k \right) \quad (2.2)$$

$$\lim_{k \rightarrow \infty} \langle X_k, Y_k \rangle = \left\langle \lim_{k \rightarrow \infty} X_k, \lim_{k \rightarrow \infty} Y_k \right\rangle \quad (2.3)$$

Satz 2.8. \mathbb{R}^n ist vollständig.

Beweis. Ist X_k eine Cauchyfolge in \mathbb{R}^n , so sind nach Satz 2.6 alle Teilfolgen Cauchy in \mathbb{R} . Also:

$$\exists x_i \in \mathbb{R} : x_i^{(k)} \rightarrow x_i \implies \exists X \in \mathbb{R}^n : X_k \rightarrow X$$

□

Satz 2.9. (Bolzano-Weierstrass:) Jede beschränkte Folge in \mathbb{R}^n besitzt eine konvergente Teilfolge.

Beweis. Sei (X_k) eine beschränkte Folge in \mathbb{R}^n . Nach 2.4 müssen die Komponentenfolgen ebenfalls beschränkt sein. Nach dem eindimensionalen Fall des Satzes von Bolzano-Weierstrass existieren also konvergente Teilfolgen der Koordinatenfolgen. Angenommen, die konvergente Teilfolge der ersten Komponente ist gegeben durch $x_1^{(k_n)} \rightarrow x_1$. So ist $x_2^{(k_n)}$ ebenfalls eine beschränkte Teilfolge, also existiert eine Teilfolge $x_2^{(k_{n_m})}$ welche in den ersten beiden Komponenten konvergiert. Führt man dieses Verfahren induktiv fort, erhält man eine konvergente Teilfolge von (X_k) . □

Satz 2.10. Sei $(A_i)_{i \in \mathbb{N}}$ eine Folge abgeschlossener beschränkter nichtleerer Teilmengen des \mathbb{R}^n , sodass $A_1 \supseteq A_2 \supseteq \dots$. Dann ist $\bigcap_{i \in \mathbb{N}} A_i \neq \emptyset$

Beweis. $A_i \neq \emptyset \implies \exists X_i \in A_i$ s.d. $(X_i)_{i \in \mathbb{N}}$ eine Folge ist. Da A_i beschränkt ist ist $(X_i)_{i \in \mathbb{N}}$ beschränkt, also hat X_i eine konvergente Teilfolge X_{i_k} mit Limes X . Es gilt $X_{i_k} \in A_{i_k} \subseteq A_i$, also ist X ein Berührungspunkt von A_i , also $X \in A_i$. □

Satz 2.11. Jede abgeschlossene beschränkte Teilmenge des \mathbb{R}^n ist kompakt.

Beweis. Analog zur eindimensionalen Version, wobei statt Intervallen $[a_i, b_i]$ Hyperwürfel $[a_i^{(1)}, b_i^{(1)}] \times \dots \times [a_i^{(n)}, b_i^{(n)}]$ genutzt werden müssen. □

Satz 2.12. Seien $\|\cdot\|_1$ und $\|\cdot\|_2$ Normen auf \mathbb{R}^n . So existieren $k, K \in \mathbb{R}^+$ mit

$$\forall X \in \mathbb{R}^n : k \|X\|_1 \leq \|X\|_2 \leq K \|X\|_1$$

Beweis. Diese Normenäquivalenz bildet eine Äquivalenzrelation. Es reicht also, zu zeigen, dass jede Norm $\|\cdot\|_2$ äquivalent zu einer spezifischen Norm $\|\cdot\|_1$ ist. Wir wählen $\|\cdot\|_{\max}$. Sei (E_i) die Standardbasis des \mathbb{R}^n . Wir definieren:

$$K := \|E_1\|_2 + \dots + \|E_n\|_2$$

Dann gilt:

$$\begin{aligned} \|X\|_2 &= \|x_1 E_1 + \dots + x_n E_n\| \\ &\leq |x_1| \|E_1\|_2 + \dots + |x_n| \|E_n\|_2 \\ &\leq \|X\|_{\max} K \quad [\text{citation needed}] \end{aligned}$$

Es bleibt die Rückrichtung zu zeigen.

Lemma 2.13. $f(X) := \|X\|_2$ ist stetig.

Beweis.

$$|\|X\|_2 - \|Y\|_2| \leq \|X - Y\|_2 \leq K\|X - Y\|_{\max} \leq K\|X - Y\|$$

Also ist $\|\cdot\|_2$ stetig bezüglich der euklidischen Norm $\|\cdot\|$. □

Wir definieren nun:

$$A := \{X \in \mathbb{R}^n \mid \|X\|_{\max} = 1\}$$

Diese Menge ist beschränkt. Wir wollen zeigen, dass sie außerdem abgeschlossen ist. Sei $X_i \rightarrow X$, $X_i \in A$. Es gilt:

$$|\|X_i\|_{\max} - \|X\|_{\max}| \leq \|X_i - X\|_{\max} \leq \|X_i - X\|$$

Also konvergiert jede Menge, also ist A kompakt, also auch abgeschlossen. Dementsprechend muss f auf A ein Minimum k annehmen. Wir wissen $f \geq 0$, also ist $k \geq 0$. Es gilt sogar $k > 0$, da keiner der Vektoren in A der Nullvektor ist. Nun gilt also $\forall X \in A : \|X\|_2 \geq k$. Wir definieren:

$$\lambda := \frac{1}{\|X\|_{\max}}$$

$$\|\lambda X\|_{\max} = |\lambda| \|X\|_{\max} = 1$$

$$|\lambda| \|X\|_2 = \|\lambda X\|_2 \geq k \implies \|X\|_2 \geq k \|X\|_{\max}$$

□

Anmerkung 2.14. Im unendlichdimensionalen Fall gilt Satz 2.12 nicht.

2.1 Abbildungen und Koordinatenfunktionen auf \mathbb{R}^n

In diesem Abschnitt betrachten wir Funktionen $F : \mathbb{R}^n \rightarrow \mathbb{R}^k$. Betrachten wir zuerst den Spezialfall Linearer Funktionen, also $\forall X, Y \in \mathbb{R}^n : \forall \lambda, \mu \in \mathbb{R} : F(\lambda X + \mu Y) = \lambda F(X) + \mu F(Y)$.

Sei (E_i) die Standardbasis des \mathbb{R}^n und sei (E'_i) die Standardbasis des \mathbb{R}^k . Nun gilt:

$$F(E_j) = \sum_{i=1}^k a_{ij} E'_i$$

Daraus erhalten wir Koeffizienten a_{ij} , welche eine Matrix bilden. Umgekehrt können wir aus den Koeffizienten die Abbildung F rekonstruieren, indem wir definieren:

$$\begin{aligned} F(X) &= F\left(\sum_{j=1}^n x_j E_j\right) \\ &= \sum_{j=1}^n x_j F(E_j) \\ &= \sum_{j=1}^n x_j \sum_{i=1}^k a_{ij} E'_i \\ &= \sum_{i=1}^k \left(\sum_{j=1}^n a_{ij} x_j\right) E'_i \end{aligned}$$

[missing stuff here]

Definition 2.15. Wir bezeichnen als $p_i : M \rightarrow k$ die Projektion eines Vektors auf die i -te Komponente.

Satz 2.16. Sei M ein metrischer Raum, $F : M \rightarrow \mathbb{R}^n$ eine Abbildung und $x \in M$. Dann ist F stetig in x genau dann, wenn $p_i \circ F$ stetig für alle i ist.

Beweis. 1. p_i ist stetig. Ist also F stetig folgt direkt, dass auch $p_i \circ F$ stetig ist.

2. Angenommen, $p_i \circ F$ ist stetig $\forall i, \varepsilon \in \mathbb{R}^+$. Da $p_i \circ F$ stetig ist existiert eine Umgebung U_i von x , sodass $|f_i(x) - f_i(y)| < \frac{\varepsilon}{\sqrt{n}} \forall y \in U_i$. Ebenso für die anderen Komponenten. Nun gilt:

$$\|F(y) - F(x)\| \leq \sqrt{n} \|F(x) - F(y)\|_{\max} \leq \varepsilon$$

□

Analog gilt das Selbe für Stetigkeit auf M , gleichmäßige Stetigkeit, etc.

Definition 2.17. Sei $M \subseteq \mathbb{R}^n$, $F : M \rightarrow \mathbb{R}^k$ eine Abbildung, x_0 ein Häufungspunkt, $y \in \mathbb{R}^k$. Dann definieren wir:

$$\lim_{x \rightarrow x_0} F(x) = y \Leftrightarrow \forall \varepsilon \in \mathbb{R}^+ : \exists \delta \in \mathbb{R}^+ : \forall x \in M \setminus \{x_0\} : \|x - x_0\| \leq \delta \implies \|F(x) - y\| < \varepsilon$$

F ist stetig in x_0 genau dann, wenn $\lim_{x \rightarrow x_0} F(x) = F(x_0)$.

Satz 2.18. Sei $M \subseteq \mathbb{R}^n$, $F : M \rightarrow \mathbb{R}^k$ eine Abbildung, $X_0 \in M$ ein Häufungspunkt, $Y \in \mathbb{R}^k$ und $f_i = p_i \circ F$. Dann gilt:

$$\lim_{X \rightarrow X_0} F(X) = Y \Leftrightarrow \forall i : \lim_{X \rightarrow X_0} f_i(X) = y_i$$

Beweis. Analog zu Beweis 2.16.

□

Korollar 2.19.

$$F(X) \rightarrow Y, G(X) \rightarrow Z \implies F(X) + G(X) \rightarrow Y + Z$$

2.2 Mehrdimensionale Ableitungen

Beispiel 2.20. Sei $f : M \rightarrow \mathbb{R}$ definiert auf einer offenen Menge $M \subseteq \mathbb{R}^n$.

$$f(X) = f(x_1, \dots, x_n) \text{ bzgl. der Standardbasis}$$

Wir können aber auch $X = \sum x'_i E'_i$ bezüglich einer beliebigen anderen Basis darstellen. Also:

$$f(X) = f(x_1, \dots, x_n) = g(x'_1, \dots, x'_n)$$

Da f in der Regel nicht linear ist, ist ein solcher Basiswechsel sehr viel komplizierter als in der Linearen Algebra! Wo möglich ist es also besser, über $f(X)$ zu reden.

Definition 2.21. Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $\bar{X} \in M$. Betrachte die Abbildung

$$t \rightarrow f(\bar{x}_1, \dots, \bar{x}_{i-1}, t, \bar{x}_{i+1}, \dots, \bar{x}_n),$$

welche eine Mehrdimensionale Funktion $f(x_1, \dots, x_n)$ auf eine eindimensionale Funktion $f(t)$ abbildet.
Achtung: Wir nehmen hier implizit eine Darstellung bezüglich der Standardbasis an!

Beispiel 2.22. Betrachte folgende Funktion:

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

f ist an $(0, 0)$ partiell differenzierbar, die Partiellen Ableitungen sind 0. Allerdings gilt

$$\forall x : f(x, x) = \frac{1}{2}$$

Also ist f an 0 nicht stetig! Es existieren also Funktionen, die an einem Punkt partiell Differenzierbar sind, an dem sie nicht stetig sind.

Idee: Fordere partielle Differenzierbarkeit bezüglich jeder möglichen Basis, also partielle Differenzierbarkeit in jedem Vektor.

Beispiel 2.23.

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^4 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

Wir betrachten die “Linearisierung” $t \rightarrow f(t, \alpha t)$. Einsetzen liefert:

$$f(t, \alpha t) = \frac{\alpha t}{t^2 + \alpha^2}$$

Diese Funktion ist differenzierbar, also ist f differenzierbar bezüglich beliebiger Basen. Das reicht jedoch immer noch nicht:

$$f(a, a^2) = \frac{a^2 a^2}{a^4 + a^4} = \frac{1}{2}$$

Also ist f immer noch nicht stetig - es ist stetig für Folgen, welche den Nullpunkt durch Geraden erreichen, aber nicht, wenn wir durch kompliziertere Pfade gegen den Nullpunkt gehen.

Wir wollen die Begriffe aus der Analysis I über Stetigkeit und Ableitbarkeit retten, also brauchen wir einen komplizierteren Ableitungsbegriff.

2.3 Differenzierbarkeit

Sei f eine beliebige Funktion $\mathbb{R} \rightarrow \mathbb{R}$. Die Ableitung gibt uns die Tangente der Funktion an einem beliebigen Punkt, also die beste affine Approximation der Funktion an diesem Punkt.

Definition 2.24. Eine Funktion $F : \mathbb{R}^n \rightarrow \mathbb{R}^k$ heißt **affin**, wenn es eine Lineare Funktion $L : \mathbb{R}^n \rightarrow \mathbb{R}^k$ und eine Konstante $Z \in \mathbb{R}^k$ gibt, sodass:

$$F(X) = L(X) + Z$$

Sei $g : \mathbb{R} \rightarrow \mathbb{R}$ affin, also $g(x) = cx + t$ für $c, t \in \mathbb{R}$. Sei $f : \mathbb{R} \rightarrow \mathbb{R}$. Wir wollen eine beliebige Funktion f an der Stelle x_0 approximieren. Für eine gute Approximation wollen wir $f(x_0) = g(x_0)$, also erhalten wir:

$$g(x) = c(x - x_0) + f(x_0).$$

Schreibe $x = x_0 + h$ und lasse h gegen 0 gehen.

$$h \rightarrow f(x_0 + h) - g(x_0 + h) = f(x_0 + h) - f(x_0) - ch$$

Wir sagen, die Approximation ist gut, wenn $f(x_0 + h) - f(x_0) - ch$ schneller gegen 0 geht als h selbst, also:

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0) - ch}{h} = 0 \quad (2.4)$$

Was äquivalent ist zu:

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

Wir sagen also, f ist in x_0 differenzierbar, genau dann, wenn eine lineare Abbildung L existiert, sodass:

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

Diese geometrische Intuition, nach der die Ableitung die beste affine Approximation der Funktion an einem gegebenen Punkt ist, können wir auf den \mathbb{R}^n übertragen. Analog zu der Interpretation affiner Funktionen als Geraden in \mathbb{R} , also der Ableitung als das Finden einer Tangentengeraden auf dem Funktionengraph, sucht man beim Ableiten einer Mehrdimensionalen Funktion eine Tangenten(hyper-)ebene auf dem Funktionengraph.

Definition 2.25. Sei $M \subset \mathbb{R}^n$ offen, $F : M \rightarrow \mathbb{R}^k$ eine Abbildung, sei $X_0 \in M$. Die Abbildung F heißt **differenzierbar** am Punkt X_0 , wenn es eine Lineare Abbildung $L : \mathbb{R}^n \rightarrow \mathbb{R}^k$ gibt, sodass:

$$\lim_{H \rightarrow 0} \frac{F(X_0 + H) - F(X_0) - L(H)}{\|H\|} = 0.$$

Wir nennen sie das **Differenzial von F im Punkt X_0** und notieren sie als DF_{X_0} . F heißt differenzierbar, wenn sie differenzierbar an jedem Punkt $X \in M$ ist.

Anmerkung 2.26. Differenzierbarkeit kann analog über die Eigenschaften des Restglieds $R(X, X_0)$ definiert werden: Sei

$$f(X) = f(X_0) + Df_{X_0}(X - X_0) + R(X, X_0).$$

Dann ist f genau dann differenzierbar, wenn:

$$\lim_{X \rightarrow X_0} \frac{R(X, X_0)}{\|X - X_0\|} = 0$$

Satz 2.27. *Gibt es ein Differential, ist es eindeutig bestimmt.*

Beweis. Seien L_1, L_2 Differentiale. Es folgt:

$$\lim_{H \rightarrow 0} \frac{L_1(H) - L_2(H)}{\|H\|} = 0$$

Sei $X \in \mathbb{R}^n \setminus \{0\}$. Dann gilt:

$$\lim_{t \rightarrow 0} \frac{L_1(tX) - L_2(tX)}{\|tX\|} = 0$$

$$\implies \frac{L_1(X) - L_2(X)}{\|X\|} = 0$$

$$\implies L_1(X) - L_2(X) = 0$$

also sind die beiden Differentiale identisch. □

Anmerkung 2.28. Unserer Differenzierbarkeitsbegriff wird insbesondere in der älteren Literatur oft als **totale Differenzierbarkeit** bezeichnet.

Satz 2.29. *Ist $F : M \rightarrow \mathbb{R}^k$ an einem Punkt X_0 differenzierbar, so ist F an diesem Punkt stetig.*

Beweis. Sei F differenzierbar. Da die Differenzierbarkeit über den Limes des Differentialquotienten definiert ist folgt direkt:

$$\begin{aligned} \forall \varepsilon \in \mathbb{R}^+ : \exists \delta_1 \in \mathbb{R}^+ : \forall H \in M : (X_0 + H \in M) \wedge (0 \leq \|H\| \leq \delta_1) \\ \implies \frac{\|F(X_0 + H) - F(X_0) - DF_{X_0}(H)\|}{\|H\|} \leq \frac{\varepsilon}{2} \\ \implies \|F(X_0 + H) - F(X_0) - DF_{X_0}(H)\| \leq \frac{\varepsilon}{2} \|H\| \end{aligned}$$

Da DF_{X_0} eine lineare Abbildung ist ist DF_{X_0} gleichmäßig stetig, also gilt:

$$\exists \delta_2 \in \mathbb{R}^+ : \|H\| < \delta_2 \implies \|DF_{X_0}(H)\| \leq \frac{\varepsilon}{2}$$

Also gilt für $\|H\| \leq \delta := \min\{\delta_1, \delta_2, 1\}$

$$\begin{aligned} & \|F(X_0 + H) - F(X_0)\| \\ &= \|F(X_0 + H) - F(X_0) - DF_{X_0}(H) + DF_{X_0}(H)\| \\ &\leq \|F(X_0 + H) - F(X_0) - DF_{X_0}(H)\| + \|DF_{X_0}(H)\| \\ &\leq \frac{\varepsilon}{2} \|H\| + \frac{\varepsilon}{2} \\ &\leq \varepsilon \end{aligned}$$

□

Satz 2.30. Sind F und G differenzierbar, so auch $F + G$, und es gilt

$$D(F + G)_{X_0} = DF_{X_0} + DG_{X_0}$$

Beweis.

$$\begin{aligned} & \lim_{H \rightarrow 0} \frac{(F + G)(X_0 + H) - (F + G)(X_0) - (DF_{X_0} + DG_{X_0})H}{\|H\|} \\ &= \lim_{H \rightarrow 0} \frac{F(X_0 + H) - F(X_0) - DF_{X_0}H}{\|H\|} + \lim_{H \rightarrow 0} \frac{G(X_0 + H) - G(X_0) - DG_{X_0}H}{\|H\|} \\ &= 0 \end{aligned}$$

□

Satz 2.31. Kettenregel: Seien $M \subseteq \mathbb{R}^n$, $N \subseteq \mathbb{R}^m$ offen, seien $F : M \rightarrow N$, $G : N \rightarrow \mathbb{R}^m$ Abbildungen, sei X_0

Beweis. Sei $F(X_0) = Y_0$, $F(X_0 + H) - F(X_0) = Z$, $H \in \mathbb{R}^n \setminus \{0\}$, $X_0 + H \in M$.

$$\begin{aligned} & \frac{1}{\|H\|} ((G \circ F)(X_0 + H) - (G \circ F)(X_0) - DG_{F(X_0)} \circ DF_{X_0}(H)) \\ &= \frac{1}{\|H\|} (G(Y_0 + Z_H) - G(Y_0) - DG_{Y_0}(Z_H)) \\ &= \frac{1}{\|H\|} (DG_{Y_0}(F(X_0 + H) - F(X_0)) - DG_{Y_0}(DF_{X_0}(H))) \\ &= \frac{1}{\|H\|} DG_{Y_0}((F(X_0 + H) - F(X_0)) - DF_{X_0}(H)) \end{aligned}$$

$$\lim_{H \rightarrow 0} \frac{1}{\|H\|} DG_{Y_0}((F(X_0 + H) - F(X_0)) - DF_{X_0}(H)) = DG_{Y_0}(0) = 0$$

$$\frac{1}{\|H\|} (G(Y_0 + Z_H) - G(Y_0) - DG_{Y_0}(Z_H)) = \begin{cases} 0 & Z_H = 0 \\ \frac{1}{\|H\|} (G(Y_0 + Z_H) - G(Y_0) - DG_{Y_0}(Z_H)) & Z_H \neq 0 \end{cases}$$

Der Term zweite Term in $Z_H \neq 0$ geht gegen 0 für $H \rightarrow 0 \implies Z_H = F(X_0 + H) - F(X_0) \rightarrow 0$

$$\begin{aligned} \frac{\|Z_H\|}{\|H\|} &= \frac{\|F(X_0 + H) - F(X_0)\|}{\|H\|} \\ &= \frac{\|DF_{X_0}(H) - R(X_0, H)\|}{\|H\|} \\ &\leq \frac{\|DF_{X_0}(H)\|}{\|H\|} + \frac{\|R(X_0, H)\|}{\|H\|} \\ &\leq \frac{\|DF_{X_0}(H)\|}{\|H\|} + \frac{\|R(X_0, H)\|}{\|H\|} \\ &\stackrel{???}{\leq} \frac{\|DF_{X_0}\| \|H\|}{\|H\|} + \frac{\|R(X_0, H)\|}{\|H\|} \\ &= \|DF_{X_0}\| + \frac{\|R(X_0, H)\|}{\|H\|} \\ &\leq c \end{aligned}$$

□

Satz 2.32. Seien $I \subseteq \mathbb{R}$, $N \subseteq \mathbb{R}^k$ offen, seien $F : I \rightarrow N$, $G : N \rightarrow \mathbb{R}^n$ Abbildungen.

Ist F differenzierbar in $t_0 \in I$ und G differenzierbar in $F(t_0)$, so gilt:

$$(G \circ F)'(t_0) = DG_{F(t_0)}(F'(t_0))$$

Beweis. Gemäß Kettenregel gilt $D(G \circ F) = DG_{F(t_0)} \circ DF_{t_0}$. Nun gilt:

$$\begin{aligned} h(G \circ F)'(t_0) &= hD(G \circ F)_{t_0}(1) \\ &= D(G \circ F)_{t_0}(h) \\ &= DG_{F(t_0)}(DF_{t_0}(h)) \\ &= hDG_{F(t_0)}(F'(t_0)) \end{aligned}$$

□

Mittelwertsatz: $f : [x, y] \rightarrow \mathbb{R}$, dann $\exists y : f(y) - f(x) = f'(z)(y - x)$. Im Allgemeinen ist dieser im Mehrdimensionalen Fall leider falsch.

Betrachte allerdings die folgende Ungleichung, welche die Wichtigste Konsequenz des Mittelwertsatzes ist: $|f(y) - f(x)| \leq |f'(z)||y - x| \leq c|y - x|$. Diese kann im Allgemeinen erhalten werden.

$F : M \subset \mathbb{R}^n \rightarrow \mathbb{R}^k$, $X, Y \in M$. Sei $[X, Y] = \{(1 - \lambda)X + \lambda Y\}$ die Verbindungslinie zwischen den beiden Vektoren.

Satz 2.33. Sei $M \subseteq \mathbb{R}^n$ offen, $X, Y \in M$ mit $[X, Y] \subseteq M$. Die Abbildung $F : M \rightarrow \mathbb{R}^k$ sei stetig in M und differenzierbar in den Punkten $(1 - \lambda)X + \lambda Y$ mit $\lambda \in (0, 1)$. Gilt

$$\forall \lambda \in (0, 1) : \forall (1 - \lambda)X + \lambda Y : \|DF_Z\| \leq c$$

so gilt auch

$$\|F(Y) - F(X)\| \leq c\|Y - X\|$$

Beweis. Angenommen $G : [0, 1] \rightarrow \mathbb{R}^k$ ist stetig auf $[0, 1]$ und differenzierbar auf $(0, 1)$. So gilt

$$\forall t \in (0, 1) : \|G'(t)\| \leq c$$

Sei $\varepsilon \in \mathbb{R}^+$ und

$$A := \{t \in [0, 1] \mid \|G(t) - G(0)\| \leq (c + \varepsilon)t + \varepsilon\}$$

Da G stetig in 0 ist gilt $[0, \tau] \subseteq A$.

Sei $s = \sup A$. Es gilt $0 < s \leq 1$, also ist G stetig in s .

Da $t \in A \implies t \leq s$

$$\|G(t) - G(0)\| \leq (c + \varepsilon)t + \varepsilon \rightarrow s \implies \|G(s) - G(0)\| \leq (c + \varepsilon)s + \varepsilon$$

also $s \in A$. Angenommen, $s < 1$. Dann gilt $\exists h > 0 : s + h < 1$.

$$\begin{aligned} \left\| \frac{G(s+h) - G(s)}{h} - G'(s) \right\| &\leq \varepsilon \\ \implies \left\| \frac{G(s+h) - G(s)}{h} \right\| &\leq \varepsilon + G'(s) \leq c + \varepsilon \end{aligned}$$

$$\begin{aligned}
\|G(s+h) - G(0)\| &\leq \|G(s+h) - G(s)\| + \|G(s) - G(0)\| \\
&\leq (c+\varepsilon)h + (c+\varepsilon)s + \varepsilon \\
&\leq (c+\varepsilon)(s+h) + \varepsilon
\end{aligned}$$

Daraus folgt $s+h \in A$. Da s das Supremum ist, ist dies ein Widerspruch. Also gilt $h=1$.

$$\begin{aligned}
\forall \varepsilon \in \mathbb{R}^+ : \|G(1) - G(0)\| &\leq c + \varepsilon + \varepsilon = c + 2\varepsilon \\
&\implies \|G(1) - G(0)\| \leq c
\end{aligned}$$

Sei F wie im Satz. Sei $K : [0, 1] \rightarrow \mathbb{R}^n : t \rightarrow (1-t)X + tY$. Diese Abbildung ist affin, also differenzierbar. Es gilt $K'(t) = Y - X$. $F \circ K$ ist diffbar in $(0, 1)$

$$D(F \circ K)_t = DF_{K(t)} \circ DK_t$$

$$(F \circ K)'(t) = DF_{K(t)}(K'(t)) = DF_{K(t)}(Y - X)$$

$$\|(F \circ K)'(t)\| = \|DF_{K(t)}(Y - X)\| \leq \|DF_{K(t)}\| \|Y - X\| \leq c \|Y - X\|$$

Mit $G := F \circ K$ und $c := c \|Y - X\|$

$$\|F(Y) - F(X)\| \leq c \|Y - X\|$$

□