



TECHNISCHE UNIVERSITÄT BERLIN

LERNSKRIPT ZUR VORLESUNG UND ÜBUNG

Analysis II & III für Mathematiker

gelesen von Prof. Dr. König im Sommersemester 2018 und
Wintersemester 2018/2019.

Übungsleiter: Dr. Alexander Schmeding

Inhaltsverzeichnis

I Analysis II: Mehrdimensionale Differentiation	3
1 Grundlagen der Topologie	3
1.1 Metrische Räume	4
1.2 Topologie in metrischen Räumen	18
1.3 Konvergenz	34
1.4 Kompaktheit und Vollständigkeit	45
1.5 Zusammenhang	56
1.6 Stetige Abbildungen	61
1.7 Fünf wichtige Sätze über stetige Abbildungen	69
1.8 Normierte Vektorräume	73
2 Grundlagen der mehrdimensionalen Differentiation	83
2.1 Die Ableitung	84
2.2 Rechenregeln für differenzierbare Abbildungen	90
2.3 Richtungsableitungen, partielle Ableitungen	99
2.4 Höhere Ableitungen und der Satz von SCHWARZ	103
2.5 Schrankensatz und der Satz von TAYLOR	111
2.6 Lokale Extrema mit und ohne Nebenbedingungen	116
2.7 Differentialoperatoren der klassischen Vektoranalysis	120
3 Die großen Sätze der mehrdimensionalen Differentialrechnung	127
3.1 Der Umkehrsatz	127
3.2 Implizite Funktionen	133
3.3 Der Rangsatz	137
4 Mannigfaltigkeiten	142
5 Differentialgleichungen	149
5.1 Lineare Differentialgleichungen	149
5.2 Der Hauptsatz über lineare Differentialgleichungen	150
5.3 Lineare Differentialgleichungen mit konstanten Koeffizienten	153
5.4 Skalare lineare Differentialgleichungen höherer Ordnung	159
II Analysis III: Maß- und Integrationstheorie	165
6 Das RIEMANN-Integral für Funktionen $\mathbb{R}^d \rightarrow \mathbb{R}$	165
6.1 Das HAARSche Maß	166
6.2 Die Transformationsformel	174
6.3 Partielle Integration	182
6.4 Der LAPLACE-Operator in krummlinigen Koordinaten	188
6.5 Integrale halbstetiger Funktionen	191
6.6 Berechnung einige Volumina	198
7 Einführung in die Maßtheorie	202

7.1 Motivation und Einführung: Das LEBESGUE-Maß auf \mathbb{R}^2	203
Alle wichtigen Definitionen, Lemmas und Sätze	i
Abbildungsverzeichnis	iii

TEIL I

Analysis II: Mehrdimensionale Differentiation

1 Grundlagen der Topologie

- Bevor wir mit der Analysis von Funktionen mehrerer Variablen beginnen können, müssen wir deren Definitionsbereiche, also höher-dimensionale Räume genauer kennenlernen.
- Wir legen die Grundlagen für die Definition von Konvergenz in solchen Räumen und damit für die Definitionen von Stetigkeit und Differenzierbarkeit von Funktionen auf solchen Räumen.



Motivation

Der \mathbb{R}^n hat mehrere fallend ineinander ($4 \subset 3 \subset 2 \subset 1$) enthaltenen Strukturen. Er ist ein

- ① *topologischer Raum*. Erklärbare Begriffe: offen, abgeschlossen, Umgebung, Konvergenz und Stetigkeit von Funktionen, (Weg)Zusammenhang, Hausdorff'sch.
- ② *metrischer Raum*. Erklärbare Begriffe: Abstand, Kugel, ε -Umgebung, Beschränktheit, Vollständigkeit, CAUCHY-Folge, glm. Konvergenz.
- ③ *normierter Vektorraum*. EB: Linearität, Länge.
- ④ *Vektorraum mit Skalarprodukt / (Prä)HILBERT-Raum*. EB: Winkel zwischen zwei Vektoren

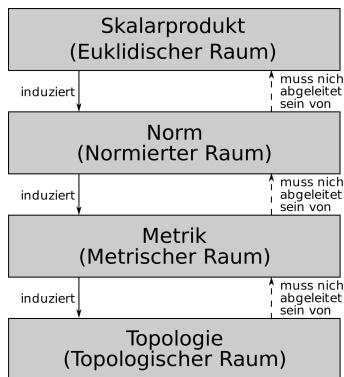


Abb. 1: [Quelle]

1.1 Metrische Räume

- Wir lernen, was eine Metrik zur Abstandsmessung von Punkten in einem Raum (d.h. in einer Menge) ist

DEFINITION 1.1.1 (METRISCHER RAUM, METRIK)

Ein **metrischer Raum** ist ein Paar (X, d) bestehend aus einer Menge X und einer Abbildung (der **Metrik**)

$$d : X \times X \rightarrow [0, \infty)$$

mit folgenden Eigenschaften für alle $x, y, z \in X$:

$$d(x, y) \geq 0 \text{ und } d(x, y) = 0 \iff x = y \quad (\text{positive Definitheit}) \quad (\text{D1})$$

$$d(x, y) = d(y, x) \quad (\text{Symmetrie}) \quad (\text{D2})$$

$$d(x, z) \leq d(x, y) + d(y, z) \quad (\text{Dreiecksungleichung}) \quad (\text{D3})$$

Beispiel 1.1.2 Auf der Menge $X = \mathbb{R}$ ist z.B. $d(x, y) := |x - y|$ eine Metrik.

Beispiel 1.1.3 (Standardmetrik auf \mathbb{R}^n) Für $x, y \in \mathbb{R}^n$ definieren wir

$$d(x, y) := \sqrt{\sum_{i=1}^n (x_i - y_i)^2}.$$

Wir nennen diese Metrik die **EUKLIDISCHE METRIK** oder die **Standardmetrik** auf \mathbb{R}^n .

Beweis. ($\triangle \neq$ für die EUKLIDISCHE Norm, von mir) Für $x, y, z \in \mathbb{R}^n$ gilt

$$\begin{aligned} \sum_{i=1}^n |x_i - z_i|^2 &\leq \sum_{i=1}^n (|x_i - y_i| + |y_i - z_i|)^2 \\ &= \sum_{i=1}^n |x_i - z_i|^2 + 2 \sum_{i=1}^n |x_i - y_i||y_i - z_i| + \sum_{i=1}^n |y_i - z_i|^2 \\ &\leq \sum_{i=1}^n |x_i - z_i|^2 + 2 \left(\sqrt{\sum_{i=1}^n |x_i - y_i|^2} \right) \left(\sqrt{\sum_{i=1}^n |y_i - z_i|^2} \right) + \sum_{i=1}^n |y_i - z_i|^2 \\ &= \left(\sqrt{\sum_{i=1}^n |x_i - y_i|^2} + \sqrt{\sum_{i=1}^n |y_i - z_i|^2} \right)^2 \end{aligned}$$

Aufgrund der Stetigkeit und strikten Monotonie von $\sqrt{\cdot}$ folgt die Behauptung. \square

Beispiel 1.1.4 (ℓ^p -Metrik) Für den \mathbb{R}^n existiert nicht nur die im letzten Beispiel angegebene Metrik, sondern viele mehr. Zum Beispiel für $p \geq 1$ die sogenannte ℓ^p -Metrik (oder MINKOWSKI-Metrik):

$$d^p(x, y) := \sqrt[p]{\sum_{i=1}^n |x_i - y_i|^p}$$

Die ℓ^p -Metrik heißt auch ℓ^p -Abstand auf der Menge

$$\ell^p(\{1, \dots, n\}) := \left\{ f : \{1, \dots, n\} \rightarrow \mathbb{R} : \sum_{k=1}^n |f_k|^p < \infty \right\}.$$

Ersetzt man $\{1, \dots, n\}$ durch \mathbb{N} , so erhält man den später eingeführten normierten Vektorraum der p -summierbaren Folgen.

Für $p = 1$ können Sie die Dreiecksungleichung selbst beweisen, für $p > 1$ ist das etwas komplizierter. An der Stelle (*) benutzen wir die Höldersche Ungleichung aus der Analysis I:

$$\sum_{k=1}^n |a_k b_k| \leq \left(\sum_{k=1}^n |a_k|^p \right)^{\frac{1}{p}} \left(\sum_{k=1}^n |b_k|^p \right)^{\frac{1}{p}}.$$

Für $a_i, b_i \in \mathbb{R}$ mit $q := \frac{p}{p-1} > 1 \implies \frac{1}{p} + \frac{1}{q} = 1$ gilt

$$\begin{aligned} \sum_{i=1}^n |a_i + b_i|^p &= \sum_{i=1}^n |a_i + b_i| |a_i + b_i|^{p-1} \\ &\leq \sum_{i=1}^n |a_i| |a_i + b_i|^{p-1} + \sum_{i=1}^n |b_i| |a_i + b_i|^{p-1} \\ &\stackrel{(*)}{\leq} \left(\sum_{i=1}^n |a_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n |a_i + b_i|^{q(p-1)} \right)^{\frac{1}{q}} + \left(\sum_{i=1}^n |b_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n |a_i + b_i|^{q(p-1)} \right)^{\frac{1}{q}} \\ &= \left(\left(\sum_{i=1}^n |a_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |b_i|^p \right)^{\frac{1}{p}} \right) \left(\sum_{i=1}^n |a_i + b_i|^p \right)^{\frac{1}{q}}, \end{aligned}$$

weil $(p-1)q = p$ ist. Division mit $(\sum_{i=1}^n |a_i + b_i|^p)^{\frac{1}{q}}$ liefert wegen $1 - \frac{1}{q} = \frac{1}{p}$

$$\left(\sum_{i=1}^n |a_i + b_i|^p \right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^n |a_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |b_i|^p \right)^{\frac{1}{p}}, \quad (1)$$

falls $\sum_{i=1}^n |a_i + b_i|^p > 0$. Hierbei ist (1) die Dreiecksungleichung für die d_p -Norm

$$\|x\|_p := \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}.$$

Aber die Ungleichung gilt natürlich auch, wenn $\sum_{i=1}^n |a_i + b_i|^p = 0$. Mit

$$a_i := y_i - x_i, \quad b_i := z_i - y_i.$$

ergibt sich die Dreiecksungleichung für d^p .

Beispiel 1.1.5 (l^∞ -Metrik) Die sogenannte l^∞ -Metrik

$$d^\infty(x, y) := \max_{i=1}^n |x_i - y_i|.$$

ist eine weitere Metrik auf dem \mathbb{R}^n .

Beweis. Wir zeigen zunächst die Dreiecksungleichung:

$$\max_{i=1}^n |x_i - z_i| \leq \max_{i=1}^n \{|x_i - y_i| + |y_i - z_i|\} \leq \max_{i=1}^n |x_i - y_i| + \max_{i=1}^n |y_i - z_i|.$$

Nun zeigen wir den Grenzübergang. Es gilt

$$\begin{aligned} d_p(x, y) &\stackrel{\text{Def}}{=} \left(\sum_{k=1}^n |x_k - y_k|^p \right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^n d_\infty(x, y)^p \right)^{\frac{1}{p}} \\ &\leq (n \cdot d_\infty(x, y)^p)^{\frac{1}{p}} = \sqrt[p]{n} \cdot d_\infty(x, y), \end{aligned}$$

und andererseits, weil $\sum_{k=1}^n d_\infty(x, y)^p$ per Definition den Summanden $\max_{k=1}^n |x_k - y_k|$ enthalten muss,

$$d_p(x, y) \stackrel{\text{Def}}{=} \left(\sum_{k=1}^n |x_k - y_k|^p \right)^{\frac{1}{p}} \geq \left(\max_{k=1}^n |x_k - y_k|^p \right)^{\frac{1}{p}} = d_\infty(x, y).$$

Somit gilt

$$\begin{aligned} d_\infty(x, y) &\leq d_p(x, y) \leq \sqrt[n]{n} \cdot d_\infty(x, y) \\ \lim_{p \rightarrow \infty} d_\infty(x, y) &\leq \lim_{p \rightarrow \infty} d_p(x, y) \leq d_\infty(x, y), \end{aligned}$$

und daraus folgt die $\lim_{p \rightarrow \infty} d^p(x, y) = d^\infty(x, y)$ für alle $x, y \in \mathbb{R}^n$. \square

DEFINITION 1.1.6 (SPURMETRIK)

Ist (X, d) ein metrischer Raum, so ist jede Teilmenge $A \subset X$ auf natürliche Weise ein metrischer Raum mit der von d induzierten Metrik oder Spurmetrik.

$$d|_{A \times A}(x, y) := d_A(x, y) := d(x, y)$$

für $x, y \in A$.

Anders gesagt: Dann ist $(A, d|_{A \times A})$ ein metrischer Raum und d_A heißt die Spurmetrik. (Beweis: Wenn es für alle $x, y, z \in X \supset A$ gilt, so auch für alle $a, b, c \in A$.) Spurmetrik

Beispiel 1.1.7 (Diskrete Metrik) Ist X eine Menge, so liefert

$$d(x, y) = \begin{cases} 0 & \text{für } x = y \\ 1 & \text{sonst} \end{cases}$$

eine Metrik auf X , die sogenannte diskrete Metrik.

diskrete Metrik

Beweis. ($\triangle \neq$ für die diskrete Metrik) Fall 1: $x = z \neq y$. Dann gilt $d(x, z) = 0$ und aufgrund der positiven Definitheit ist die Bedingung auf jeden Fall erfüllt.

Fall 2: $x \neq z$. Dann gilt $x \neq y$ oder $y \neq z$, da y nicht gleichzeitig mit zwei verschiedenen Elementen übereinstimmen kann. In beiden Fällen ist einer der Terme auf der rechten Seite gleich Eins und der andere gleich Null und es herrscht Gleichheit. \square

DEFINITION 1.1.8 (BESCHRÄNKTHEIT)

Sei (X, d) ein metrischer Raum.

① Eine Teilmenge $A \subset X$ heißt beschränkt, wenn gilt:

- $\forall x \in X \exists M \in \mathbb{R}_{>0} \forall a \in A : d(x, a) \leq M$.

Ist $X \neq \emptyset$, so ist das äquivalent zu folgender Bedingung:

- $\exists x \in X \exists M \in \mathbb{R} : d(x, a) \leq M \forall a \in A$.

② Eine Abbildung $f : X \rightarrow Y$ einer beliebigen Menge X heißt beschränkt, wenn ihr Bild $f(X) \subset Y$ beschränkt ist.

BILD ZEICHNEN

DEFINITION 1.1.9 (DURCHMESSER)

Für eine nichtleere Teilmenge $A \subset Y$ heißt

$$\text{diam}(A) := \sup_{x, y \in A} d(x, y) \in [0, \infty]$$

der Durchmesser von A . Wir setzen $\text{diam}(\emptyset) := 0$.

Durchmesser

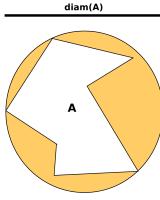


Abb. 2: Der Durchmesser der Menge A .

Bemerkung 1.1.10 (Vorgriff von mir) Eine alternative Definition einer beschränkten Teilmenge wäre die Existenz eines $\varepsilon > 0$, sodass A von der ε -Kugel überdeckt wird.

Lemma 1.1.11

Eine Teilmenge A des metrischen Raumes (Y, d) ist genau dann beschränkt, wenn sie endlichen Durchmesser hat.

Beweis. " \implies ": Sei o.E. $A \neq \emptyset$. Ist A beschränkt, so existiert $M \in \mathbb{R}$ und $y \in Y$ mit $d(y, y') < M$ für alle $y' \in A$. Also gilt für alle $y', y'' \in A$

$$d(y', y'') \leq d(y', y) + d(y, y'') \leq 2M.$$

Damit ist $\text{diam}(A) \leq 2M < \infty$.

" \impliedby ": Ist $M := \text{diam}(A) < \infty$ und $y \in A$, so gilt für alle $y' \in A$

$$d(y, y') \leq M,$$

also ist A beschränkt.

BILD ZEICHNEN.

□

DEFINITION 1.1.12 (MENGE DER BESCHRÄNKTEN ABBILDUNG)

Seien (Y, d) ein nichtleerer metrischer Raum und $X \neq \emptyset$ eine beliebige Menge. Dann ist

$$\mathcal{B}(X, Y) := \{f : X \rightarrow Y \mid f \text{ beschränkt}\}$$

die Menge der beschränkten Abbildungen von X in Y .

SATZ 1.1.1: SUPREMUMSMETRIK

Für $f, g \in \mathcal{B}(X, Y)$ definiert die Abbildung

$$d^{\sup}(f, g) := \sup_{x \in X} d(f(x), g(x))$$

Supremumsmetrik eine Metrik auf $\mathcal{B}(X, Y)$, die sogenannte **Supremumsmetrik**.

Beweis. Wohldefiniertheit: Sei $x_0 \in X$. Dann existiert zu $f, g \in \mathcal{B}(X, Y)$ ein M mit

$$d(f(x_0), f(x)) < M \quad \text{und} \quad d(g(x_0), g(x)) < M$$

für alle $x \in X$. Daher ist für alle x

$$d(f(x), g(x)) < 2M + d(f(x_0), g(x_0)),$$

und $\sup_{x \in X} d(f(x), g(x)) \in \mathbb{R}$. Also ist $d^{\sup} : \mathcal{B}(X, Y) \times \mathcal{B}(X, Y) \rightarrow \mathbb{R}$ definiert.

Metrik-Axiome (1), (2) sind trivial. Zur Dreiecksungleichung:

$$\begin{aligned} d^{\sup}(f, h) &= \sup_{x \in X} d(f(x), h(x)) \\ &\stackrel{\triangle \neq}{\leqslant} \sup_{x \in X} (d(f(x), g(x)) + d(g(x), h(x))) \\ &\leqslant \sup_{x \in X} d(f(x), g(x)) + \sup_{x \in X} d(g(x), h(x)) \\ &= d^{\sup}(f, g) + d^{\sup}(g, h). \end{aligned}$$

□

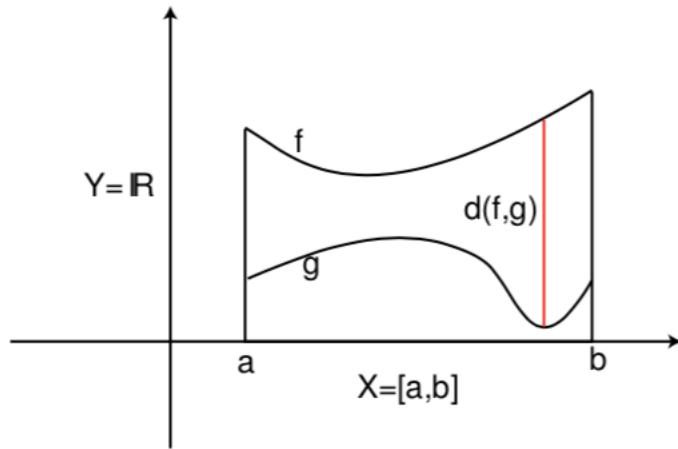


Abb. 3: Veranschaulichung der Metrik d^{\sup} . [Quelle: Ferus]

Bemerkung $(\mathcal{B}(X, Y), d^{\sup})$ und (\mathbb{R}^n, d^∞) isometrisch isomorph:

$$(\mathcal{B}(X, Y), d^{\sup}) \cong (\mathbb{R}^n, d^\infty).$$

Das heißt, es gibt eine Bijektion $\varphi : \mathcal{B}(X, Y) \rightarrow \mathbb{R}^n$, nämlich $\varphi : f \mapsto (f(1), \dots, f(n))$, für die $d^\infty(\varphi(f), \varphi(g)) = d^{\sup}(f, g)$ gilt.

AUFGABEN

- (1) Betrachten Sie $f : [0, \infty) \rightarrow [0, \infty)$, $x \mapsto \frac{x}{1+x}$ und die Metrik $d(x, y) := |x - y|$. Zeigen Sie,

- (i) dass f monoton wachsend ist.
- (ii) dass $d_f : \mathbb{R} \times \mathbb{R} \rightarrow [0, \infty)$, $(x, y) \mapsto f(d(x, y))$ eine Metrik definiert.

- (2) Sei $d_2 : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ eine Metrik (*) und $P \in \mathbb{R}^2$ fest. Zeigen Sie, dass die sogenannte SNCF-Metrik

$$d_F(x, y) := \begin{cases} d_2(x, y) & P \in \overline{xy} \\ d_2(x, P) + d_2(P, y) & \text{sonst} \end{cases}$$

eine Metrik auf \mathbb{R}^2 definiert.

Bemerkung: Man kann sich diese Metrik wie folgt vorstellen, sofern der Weg von einer französischen Stadt zu einer anderen nicht direkt über Paris verläuft, muss man über Paris fahren, da alle Weg nur von Paris aus verlaufen.

Außerdem ist die Metrik wohldefiniert.

- (3) Seien $P = \{\{a_1, a_2\}, \dots, \{a_{k-1}, a_k\}\}$ ein Weg von a_1 nach a_k und $\ell(P) = k$ die Anzahl der Kanten sowie $G = (V, E)$ endlich und zusammenhängend. Dann ist $d : E \times E \rightarrow [0, \infty)$, $(a, b) \mapsto \min\{\ell(w) \mid w \text{ ist ein Weg von } a \text{ nach } b\}$ eine Metrik.



- (4) Sei $\mathcal{G} := (K, E)$ ein endlicher, zusammenhängender Graph und $B : K \rightarrow (0, \infty)$ eine Abbildung. Sei $w = (k_1, k_2, \dots, k_n)$ ein Weg von a nach b in \mathcal{G} . Wir definieren die bewertete Länge $l_B(w) := \sum_{i=1}^n B(k_i)$. Zeigen Sie, dass die folgende Vorschrift eine Metrik auf \mathcal{G} definiert:

$$d(a, b) := \min\{l_B(w) \mid w \text{ ist ein Weg von } a \text{ nach } b\}$$

- (5) Sei $n : \mathbb{R} \rightarrow \mathbb{R}$ eine Abbildung. Wir setzen $d(x, y) := n|x - y|$ für $x, y \in \mathbb{R}$. Finden Sie eine möglichst einfache notwendige und hinreichende Bedingung an n dafür, dass d eine Metrik ist.

- (6) Seien (X, d) ein metrischer Raum, und $f : [0, \infty) \rightarrow [0, \infty)$, $x \mapsto \frac{x}{x+1}$ eine Funktion, sowie $d_f(x, y) : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} : (x, y) \mapsto f(d(x, y))$ eine Metrik.

- (i) Zeigen Sie: Eine Teilmenge $U \subseteq X$ ist genau dann offen bezüglich der Metrik d , wenn sie offen bezüglich der Metrik d_f ist.

- (ii) Beschreiben Sie alle beschränkten Teilmengen von X bezüglich der Metrik d_f .

- (iii) Geben Sie ein Beispiel eines metrischen Raumes (X, d) und einer Menge $M \subset X$ an, sodass M bezüglich d_f beschränkt, aber bezüglich d unbeschränkt ist.

- (7) Sei $\mathcal{C}(\mathbb{R}, \mathbb{R})$ der Vektorraum der stetigen Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$. Für $n \in \mathbb{N}$ sei

$$D_n : \mathcal{C}(\mathbb{R}, \mathbb{R}) \times \mathcal{C}(\mathbb{R}, \mathbb{R}) \rightarrow \mathbb{R} \quad (f, g) \mapsto \sup_{-n \leq x \leq n} |f(x) - g(x)|$$

- (i) Zeigen Sie, dass $D(f, g) := \sum_{n=1}^{\infty} 2^{-n} \frac{D_n(f, g)}{1+D_n(f, g)}$ eine Metrik auf $\mathcal{C}(\mathbb{R}, \mathbb{R})$ definiert.
- (ii) Für $B_\varepsilon(0) = \{g \in \mathcal{C}(\mathbb{R}, \mathbb{R}) \mid D(0, g) < \varepsilon\}$ mit $\varepsilon > 0$ gilt, dass es ein $k \in \mathbb{N} \setminus \{0\}$ und ein $\delta > 0$ gibt, sodass $\{g \in \mathcal{C}(\mathbb{R}, \mathbb{R}) \mid D_k(0, g) < \delta\} \subseteq B_\varepsilon(0)$ gilt.
- (iii) Für $k < \infty$ ist die Abbildung $D^k(f, g) := \sum_{n=1}^k 2^{-n} \frac{D_n(f, g)}{1+D_n(f, g)}$ keine Metrik auf $\mathcal{C}(\mathbb{R}, \mathbb{R})$.



- (8) Sei $X := \mathbb{R}^2$. Verifizieren Sie, dass die sog. **Dschungelfluss- oder Stacheldrahtmetrik**

$$d : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto \begin{cases} |x_2 - y_2|, & \text{wenn } x_1 = y_1, \\ |x_2| + |y_2| + |x_1 - y_1|, & \text{sonst.} \end{cases}$$

eine Metrik auf \mathbb{R}^2 ist.

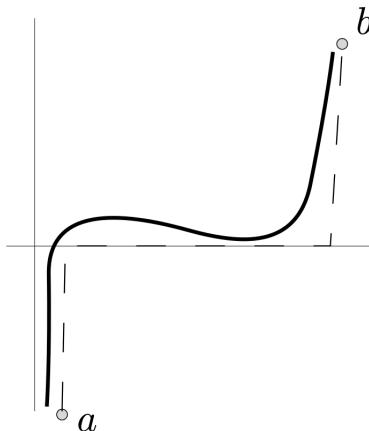


Abb. 4: Man sich einen Fisch vorstellen, der in einem Flusses entlang der x -Achse schwimmt, welcher unendlich viele Nebenflüsse hat, die parallel zu der y -Achse verlaufen. [Quelle]

- (9) Sind die folgenden Abbildungen Metriken auf \mathbb{R} ?

- (i) $d_1(x, y) := |\sin(x) - \sin(y)|$
- (ii) $d_2(x, y) := |\arctan(x) - \arctan(y)|$
- (iii) $d_3(x, y) := \min(d(x, y), 1)$, wobei $d(x, y)$ eine Metrik auf \mathbb{R} ist.

- (10) Unter welchen Voraussetzungen ist $d(x, y) := |f(x) - f(y)|$ eine Metrik auf \mathbb{R} ?

LÖSUNGEN

(1) hier

Beweis. (i) Ansatz 1: Für alle $x \in [0, \infty)$ gilt $f'(x) = \frac{1}{(x+1)^2} > 0$.

Ansatz 2: Seien $x, y \in [0, \infty)$. Dann gilt

$$x < y \implies x + xy < y + xy \implies x(1+y) < y(1+x) \implies \frac{x}{x+1} < \frac{y}{y+1} \implies f(x) < f(y).$$

(ii) Seien $x, y, z \in [0, \infty)$. Dann gilt

(a) " \Rightarrow ": $x = y \implies d_f(x, y) = \frac{d(x, y)}{d(x, y)+1} = \frac{0}{0+1} = 0$, sowie
 " \Leftarrow ": $d_f(x, y) = 0 \iff \frac{d(x, y)}{d(x, y)+1} = 0 \Rightarrow d(x, y) = 0 \iff x = y$, weil f injektiv ist.

(b) $d_f(x, y) = f(d(x, y)) = f(d(y, x)) = d_f(y, x)$.

(c) Für alle $x, y, z \in [0, \infty)$ gilt

$$\begin{aligned} d_f(x, y) = f(d(x, y)) &\stackrel{(a)}{\leqslant} f(d(x, z) + d(z, y)) = \frac{d(x, y) + d(z, y)}{1 + d(x, y) + d(z, y)} \\ &= \frac{d(x, y)}{1 + d(x, y) + d(z, y)} + \frac{d(z, y)}{1 + d(x, y) + d(z, y)} \\ &\leqslant \frac{d(x, y)}{1 + d(x, y)} + \frac{d(z, y)}{1 + d(z, y)} \\ &= d_f(x, y) + d_f(y, z) \end{aligned} \quad \square$$

(2) **Beweis.** Seien $x, y, z, P \in \mathbb{R}^2$, und aeG als 'auf einer Geraden' definiert dann gilt, dass

(i) " \Rightarrow ": $d(x, x) = d_2(x, x) \stackrel{(*)}{=} 0$.

" \Leftarrow ": Es gilt entweder $0 = d_F(x, y) = d_2(x, y) \stackrel{(*)}{\iff} x = y$ oder

$$0 = d_F(x, y) = d_2(x, P) + d_2(P, y) \stackrel{(*)}{\geqslant} d_2(x, y) \stackrel{(*)}{\geqslant} 0 \implies d_2(x, y) = 0 \stackrel{(*)}{\iff} x = y.$$

(ii) Klar durch die Symmetrie von d_2 (*).

(iii) Fall 1: x, y, P aeG.

Fall 1.1: x, y, z, P aeG. Klar weil (*).

Fall 1.2: z nicht aeG mit x, y, P . Dann gilt

$$\begin{aligned} d_F(x, y) = d_2(x, y) &\stackrel{(*)}{\leqslant} d_2(x, z) + d_2(z, y) && \text{(zwei Mal } \triangle \neq \text{)} \\ &\stackrel{(*)}{\leqslant} d_2(x, P) + d_2(P, z) + d_2(z, P) + d_2(P, y) = d_F(x, z) + d_F(z, y) \end{aligned}$$

Fall 2: x, y nicht aeG mit P .

Fall 2.1: x, z, P aeG aber y, z, P nicht aeG. Dann gilt

$$d_F(x, y) = d_2(x, P) + d_2(P, y) \leq d_2(x, z) + d_2(z, P) + d_2(P, y) = d_F(x, z) + d_F(z, y)$$

Analog argumentiert man mit y, z, P aeG aber x, y, P nicht aeG

Fall 2.2: Weder x, z, P noch y, z, P aeG. Dann gilt

$$d_F(x, y) = d_2(x, P) + d_2(P, y) \leq d_2(x, P) + 2d_2(P, z) + d_2(P, y) = d_F(x, z) + d_F(y, z)$$

Fall 2.2: x, z, P und y, z, P aeG. Dann gilt $z = P$ und somit

$$d_F(x, y) = d_2(x, P) + d_2(P, y) = d_F(x, z) + d_2(z, y)$$

- (3) **Beweis.**
- (i) " \Rightarrow ": $d(x, x) = \min_{P_{x,x}} \{\ell(P)\} = \ell(\text{leerer Pfad}) = 0$.
 - " \Leftarrow ": $d(x, y) = 0 \implies \min_{P_{x,y}} \{\ell(P)\} \stackrel{!}{=} 0 \iff \exists P_{x,y} : P = \{\text{emptyset}\} \iff x = y$.
 - (ii) $d(x, y) = \min_{P_{x,y}} \ell(P) = \min_{Q_{y,x}} \ell(Q) = d(y, x)$.
 - (iii) Seien $x, y, z \in E$.
 $d(x, y) + d(y, z) = \min \ell(P_x^y) + \min \ell(P_y^z)$.
 Wir hängen nun P_x^y und P_y^z aneinander, dann ist dieser Weg länger als P_x^z , also gilt

$$d(x, y) + d(y, z) \geq \min \ell(P_x^z) = d(x, z).$$

- (4) **Beweis.** Es gilt die drei definierenden Eigenschaften einer Metrik für alle $a, b, c \in E$ nachzuweisen:

(I) $d(a, b) = 0 \iff a = b$

" \Rightarrow ": Sei $a \neq b$. Das heißt jeder Weg zwischen a und b ist nicht leer. Da B nach $(0, \infty)$ abbildet, ist also jede Länge von Wegen zwischen a und b echt größer null, also auch das Minimum über die Mengen, also: $d(a, b) > 0$

" \Leftarrow ": Der leere Weg $w = \{\}$ führt für $a \in E$ von a nach a . Dieser Weg hat die Länge $\ell_B(w) = \sum_{i=1}^0 B(k_i) = 0$. Außerdem muss dies das Minimum aller Wege von a nach a sein, da B nur auf positive Werte abbildet, somit ist $d(a, a) = 0$.

Ferner gilt auch für alle $a, b \in E$: $d(a, b) \geq 0$, da die Funktion B eine strikt positiven Wertebereich besitzt und $\ell_B(w)$ somit nur positive Werte aufsummiert.

(II) Symmetrie: für alle $a, b \in E$ gilt $d(a, b) = d(b, a)$.

Weil \mathcal{G} zusammenhängend und ungerichtet ist, ist für $a, b \in E$ jeder Weg von a nach b auch ein Weg von b nach a , denn wenn (k_1, \dots, k_n) ein Weg von a nach b ist, ist (k_n, \dots, k_1) ein Weg von b nach a . Somit gilt

$$\begin{aligned} d(a, b) &= \min \{l_B(w) \mid w \text{ ist ein Weg von } a \text{ nach } b\} \\ &= \min \{l_B(w) \mid w \text{ ist ein Weg von } b \text{ nach } a\} = d(b, a) \end{aligned}$$

(III) Dreiecksungleichung: für alle $a, b, c \in E$ gilt $d(a, c) \leq d(a, b) + d(b, c)$.

Sei $a, b, c \in E$. Ist $w_0 = (k_1, \dots, k_\ell)$ ein Weg von a nach b mit der kürzesten Länge und $w_1 = (k_{\ell+1}, \dots, k_n)$ ein Weg von b nach c mit der kürzesten Länge. So ist $w_2 = (k_1, \dots, k_\ell, k_{\ell+1}, \dots, k_n)$, definiert als die Aneinanderreihung von w_0 und w_1 ebenfalls ein Weg von a nach c . Somit ist $\ell_B(w_2) \in \{l_B(w) \mid w \text{ ist ein Weg von } a \text{ nach } c\}$ und es gilt

$$\begin{aligned} d(a, c) + d(b, c) &= \ell_B(w_0) + \ell_B(w_1) \\ &= \ell_B(w_2) \geq \min \{l_B(w) \mid w \text{ ist ein Weg von } a \text{ nach } c\} = d(a, c) \end{aligned}$$

- (5) Behauptung: Für $x, y \in \mathbb{R}$ sind folgenden Bedingungen notwendig und hinreichend

$$n(0) = 0 \text{ und } n|_{\mathbb{R} \setminus \{0\}} > 0 \quad (1)$$

$$n(x) = n(-x) \quad (2)$$

$$n(x+y) \leq n(x) + n(y). \quad (3)$$

Beweis. Notwendig: Sei $x, y, z \in \mathbb{R}$.

Aus dem ersten Axiom einer Metrik folgt:

$$d(x, x) = 0 \implies n(x-x) = 0 \implies n(0) = 0$$

Außerdem folgt mit $d(x, y) \geq 0$

$$[x \neq y \Rightarrow d(x, y) \neq 0] \implies [x \neq y \Rightarrow n(x - y) > 0] \implies n|_{\mathbb{R} \setminus \{0\}} > 0$$

Aus dem zweitem Axiom einer Metrik folgt

$$d(x, y) = d(y, x) \implies n(x - y) = n(y - x) = n(-(x - y)) \implies n(z) = n(-z)$$

Aus dem dritten Axiom einer Metrik folgt:

$$d(x, y) \leq d(x, z) + d(z, y) \implies n(x - y) \leq n(x - z) + n(z - y)$$

Für $p_1 := x - z$ und $p_2 := z - y$ gilt also

$$n(p_1 + p_2) \leq n(p_1) + n(p_2).$$

Da x, y, z beliebig aus \mathbb{R} gewählt sind, sind auch p_1, p_2 als deren Differenzen beliebig in \mathbb{R} .

Hinreichend: Aus (1) $n(0) = 0$ folgt für $x = y$

$$d(x, y) = n(x - y) = n(0) = 0$$

Außerdem ist 0 der einzige Punkt, der auf 0 abbildet, da $n|_{\mathbb{R} \setminus \{0\}} > 0$ gilt, somit folgt für $x, y, z \in \mathbb{R}$

$$[n(z) = 0 \Rightarrow z = 0] \implies [d(x, y) = 0 \Rightarrow n(x - y) = 0 \Rightarrow x - y = 0]$$

Seien $x, y, z \in \mathbb{R}$. Dann gilt nach (2) $n(z) = n(-z)$. Also gilt für d

$$d(x, y) = n(x - y) = n(-(x - y)) = n(y - x) = d(y, x)$$

Nun fehlt lediglich die Dreiecksungleichung. Für $x, y, z \in \mathbb{R}$ gibt es $p_1, p_2 \in \mathbb{R}$, sodass $p_1 = x - z$ und $p_2 = z - y$ sind. Also folgt für d aus (3):

$$d(x, y) = n(p_1 + p_2) \leq n(p_1) + n(p_2) = d(x, z) + d(z, y)$$

(6) (i) Beweis. Damit U bezüglich $d(x, y)$ offen ist, muss gelten, dass

$$\begin{aligned} \forall u \in U \ \exists \varepsilon > 0 : \{x \in X : d(x, y) < \varepsilon\} &\subseteq U \\ \iff \forall u \in U \ \exists \varepsilon > 0 \ \forall x \in X : (d(x, y) < \varepsilon) &\implies x \in U. \end{aligned}$$

Damit U bezüglich $d_f(x, y)$ offen ist, muss analog gelten, dass

$$\forall u \in U \ \exists \tilde{\varepsilon} > 0 \ \forall x \in X : d_f(x, y) < \tilde{\varepsilon} \implies x \in U. \quad (*)$$

□

” \Rightarrow ”: Sei $\varepsilon > 0$ und $\tilde{\varepsilon} := \varepsilon$ und U bzgl. d offen. Da f streng monoton wachsend ist gilt $a < b \implies f(a) < f(b)$ für alle $a, b \in [0, \infty)$. Somit gilt für alle $x, y \in X$

$$d(x, y) < \varepsilon \implies \frac{d(x, y)}{1 + d(x, y)} < \frac{\varepsilon}{1 + \varepsilon} < \tilde{\varepsilon}$$

” \Leftarrow ”: Seien $\varepsilon > 0$ und $\tilde{\varepsilon} := \frac{\varepsilon}{1 - \varepsilon} > 0$ und U bzgl. d_f offen. Dann gilt aufgrund von (*) für alle $x, y \in X$

$$d_f(x, y) = \frac{d(x, y)}{1 + d(x, y)} < \varepsilon \implies d(x, y) < \varepsilon + \varepsilon \cdot d(x, y) \implies d(x, y) < \frac{\varepsilon}{1 - \varepsilon} = \tilde{\varepsilon}$$

- (ii) Damit $T \subset X$ eine beschränkte Teilmenge ist, muss gelten

$$\forall x \in X \exists M \in \mathbb{R} : d(x, \tilde{x}) \leq M \quad \forall \tilde{x} \in T$$

Für alle $x \in [0, \infty)$ gilt $0 < 1 \iff x < x + 1 \iff \frac{x}{x+1} < 1$, somit ist die Bedingung für alle Teilmengen von X mit $M := 1 \in \mathbb{R}$ erfüllt.

- (iii) Sei $X = \mathbb{R}$. Der metrische Raum (\mathbb{R}, d) mit $d(x, y) := |x - y|$ ist ein Beispiel für eine solche Metrik, denn Abstände von zwei Zahlen des Zahlenstrahls können beliebig groß werden, jedoch ist jede Teilmenge $M \subset \mathbb{R}$ bzgl. d_f nach Aufgabe 3b) beschränkt.

(7) hier

- (i) **Beweis.** Um zu zeigen, dass $D(f, g)$ wohldefiniert ist, zeigen wir $D(f, g) < \infty$.

Die unendliche Reihe $D(f, g)$ ist konvergent, da gilt

$$D(f, g) = \sum_{n=1}^{\infty} 2^{-n} \underbrace{\frac{D_n(f, g)}{1 + D_n(f, g)}}_{\in [0, 1]} \leq \sum_{n=1}^{\infty} 2^{-n} = 2 < \infty.$$

Es gilt die drei definierenden Eigenschaften einer Metrik nachzuweisen:

- (I) Für alle $f, g \in \mathcal{C}(\mathbb{R}, \mathbb{R})$, $n \in \mathbb{N}$ und $x \in \mathbb{R}$ gilt $2^{-n} > 0$ und $\sup_{-n \leq x \leq n} |f(x) - g(x)| \geq 0$.

Daraus folgt direkt $D(f, g) \geq 0$. Außerdem impliziert $f = g$

$$\sup_{-n \leq x \leq n} |f(x) - g(x)| = \sup_{-n \leq x \leq n} \{0\} = 0,$$

und damit

$$D(f, g) = \sum_{n=1}^{\infty} 2^{-n} \frac{D_n(f, g)}{1 + D_n(f, g)} = \sum_{n=1}^{\infty} 2^{-n} \frac{0}{1} = 0.$$

Umgekehrt impliziert $f \neq g$, dass es ein $x_1 \in \mathbb{R}$ gibt mit $f(x_1) \neq g(x_1)$.

Man wähle $n_1 \in \mathbb{N}$ so, dass $-n_1 \leq x_1 \leq n_1$ gilt. Dann folgt

$$\sup_{-n_1 \leq x \leq n_1} |f(x) - g(x)| \geq |f(x_1) - g(x_1)| > 0,$$

also schließlich

$$D(f, g) = \sum_{n=1}^{\infty} 2^{-n} \frac{D_n(f, g)}{1 + D_n(f, g)} \geq 2^{-n_1} \frac{D_{n_1}(f, g)}{1 + D_{n_1}(f, g)} > 0.$$

- (II) Für alle $f, g \in \mathcal{C}(\mathbb{R}, \mathbb{R})$, $n \in \mathbb{N}$ mit $x \in \mathbb{R}$ gilt

$$\sup_{-n \leq x \leq n} |f(x) - g(x)| = \sup_{-n \leq x \leq n} |g(x) - f(x)|.$$

Somit gilt

$$D(f, g) = \sum_{n=1}^{\infty} 2^{-n} \frac{D_n(f, g)}{1 + D_n(f, g)} = \sum_{n=1}^{\infty} 2^{-n} \frac{D_n(g, f)}{1 + D_n(g, f)} = D(g, f)$$

- (III) Es gilt für alle $f, g, h \in \mathcal{C}(\mathbb{R}, \mathbb{R})$, $n \in \mathbb{N}$

$$\begin{aligned} \sup_{-n \leq x \leq n} |f(x) - g(x)| &= \sup \{|f(x) - g(x)| : x \in \mathbb{R} \text{ mit } -n \leq x \leq n\} \\ &\leq \sup \{|f(x) + h(x)| + |h(x) - g(x)| : x \in \mathbb{R} \text{ mit } -n \leq x \leq n\} \\ &\leq \sup \{|f(x) + h(x)| + |h(x') - g(x')| : x, x' \in \mathbb{R} \text{ mit } -n \leq x \leq n\} \\ &= \sup \{|f(x) + h(x)| : x \in \mathbb{R} \text{ mit } -n \leq x \leq n\} \\ &\quad + \sup \{|h(x') - g(x')| : x' \in \mathbb{R} \text{ mit } -n \leq x \leq n\} \end{aligned}$$

Also gilt $D_n(f, g) \leq D_n(f, h) + D_n(h, g)$ und auch folgende Implikation:

$$\begin{aligned} D_n(f, g) &\leq D_n(f, h) + D_n(h, g) \\ \implies D_n(f, g) + D_n(f, g) \cdot (D_n(f, h) + D_n(h, g)) &\leq D_n(f, h) + D_n(h, g) + D_n(f, g) \cdot (D_n(f, h) + D_n(h, g)) \\ \implies \frac{D_n(f, g)}{1 + D_n(f, g)} &\leq \frac{D_n(f, h) + D_n(h, g)}{1 + D_n(f, h) + D_n(h, g)} \\ \implies \frac{D_n(f, g)}{1 + D_n(f, g)} &\leq \frac{D_n(f, h)}{1 + D_n(f, h) + D_n(h, g)} + \frac{D_n(h, g)}{1 + D_n(f, h) + D_n(h, g)} \\ \implies \frac{D_n(f, g)}{1 + D_n(f, g)} &\leq \frac{D_n(f, h)}{1 + D_n(f, h)} + \frac{D_n(h, g)}{1 + D_n(h, g)} \end{aligned}$$

Somit also auch

$$\begin{aligned} D(f, g) &= \sum_{n=1}^{\infty} 2^{-n} \frac{D_n(f, g)}{1 + D_n(f, g)} \\ &\leq \sum_{n=1}^{\infty} 2^{-n} \frac{D_n(f, h)}{1 + D_n(f, h)} + \sum_{n=1}^{\infty} 2^{-n} \frac{D_n(h, g)}{1 + D_n(h, g)} = D(f, h) + D(h, g) \end{aligned}$$

Damit sind alle Eigenschaften gezeigt. \square

- (ii) Beweis.** Aus der Analysis I ist bekannt, dass $\sum_{n=1}^{\infty} 2^{-n}$ konvergiert. Das bedeutet, dass die Folge der Reihenreste $(\sum_{n=i}^{\infty} 2^{-n})_{i \in \mathbb{N}}$ gegen Null konvergiert. Also kann man ein $k \in \mathbb{N}$ wählen, sodass

$$\sum_{n=k+1}^{\infty} 2^{-n} < \frac{\varepsilon}{2} \text{ gilt.}$$

Außerdem sei $\delta := \frac{\varepsilon}{4 - \varepsilon}$. Dann folgt aus $g \in \{g \in \mathcal{C}(\mathbb{R}, \mathbb{R}) | D_k(0, g) < \frac{\varepsilon}{4 - \varepsilon}\}$, dass $D_k(g, 0) < \delta = \frac{\varepsilon}{4 - \varepsilon}$. Weiterhin gilt für alle $n < k$ $\sup_{-n \leq x \leq n} |g(x)| \leq \sup_{-k \leq x \leq k} |g(x)|$ und somit folgt aus der in der Gruppenübung 2 bewiesenen Monotonie der Funktion $x \mapsto \frac{x}{1+x}$, dass für $n < k$ gilt, dass

$$\sum_{n=1}^k 2^{-n} \frac{D_n(0, g)}{1 + D_n(0, g)} \leq \sum_{n=1}^k 2^{-n} \frac{D_k(0, g)}{1 + D_k(0, g)} \text{ gilt.}$$

Insgesamt gilt also

$$\begin{aligned} \sum_{n=1}^{\infty} 2^{-n} \frac{D_n(0, g)}{1 + D_n(0, g)} &= \sum_{n=1}^k 2^{-n} \frac{D_n(0, g)}{1 + D_n(0, g)} + \sum_{n=k+1}^{\infty} 2^{-n} \frac{D_n(0, g)}{1 + D_n(0, g)} \\ &< \sum_{n=1}^k 2^{-n} \frac{D_k(0, g)}{1 + D_k(0, g)} + \frac{\varepsilon}{2} \\ &\leq \sum_{n=1}^k 2^{-n} \frac{\delta}{1 + \delta} + \frac{\varepsilon}{2} = \frac{\delta}{1 + \delta} \sum_{n=1}^k 2^{-n} + \frac{\varepsilon}{2} \\ &\leq \frac{\delta}{1 + \delta} \sum_{n=0}^{\infty} 2^{-n} + \frac{\varepsilon}{2} = \frac{\delta}{1 + \delta} \cdot \frac{1}{1 - \frac{1}{2}} + \frac{\varepsilon}{2} \\ &\quad (\text{Geometrische Reihe}) \\ &= \frac{2\delta}{1 + \delta} + \frac{\varepsilon}{2} = \frac{2\varepsilon}{(4 - \varepsilon)(1 + \frac{\varepsilon}{4 - \varepsilon})} + \frac{\varepsilon}{2} = \frac{2\varepsilon}{4 - \varepsilon + \varepsilon} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

Somit ist $D(0, g) < \varepsilon$ und $g \in B_{\varepsilon}(0)$ \square

- (iii) Beweis.** Sei $f, g \in \mathcal{C}(\mathbb{R}, \mathbb{R})$ mit $f(x) = g(x)$ für alle $x \in [-k, k]$ und $f(x) \neq g(x)$ für ein beliebiges $x \notin [-k, k]$. Solche Funktionen existieren, weil man eine

auf einem Intervall stetige Funktion auf verschiedene Arten stetig fortsetzen kann. Dann gilt $f \neq g$, jedoch auch für alle $n \leq k$

$$\sup_{-n \leq x \leq n} |f(x) - g(x)| = 0 \implies D^k(f, g) = 0.$$

Dies ist ein Widerspruch zur ersten definierenden Eigenschaft einer Metrik (X, d) :

$$d(x, y) = 0 \iff x = y \quad \forall x, y \in X$$



⑧ Wir überprüfen die drei Eigenschaften einer Metrik:

- i Weil in beiden Fällen nur Absolutbeträge reeller Zahlen addiert werden, ist $d(x, y) \geq 0$ für alle $x, y \in X$.

Ist $x = y$ gilt $x_1 = y_1$ und $x_2 = y_2$. Wir betrachten also den ersten Fall der Fallunterscheidung: es gilt $d(x, y) = |x_2 - y_2| = 0$.

Sei nun $d(x, y) = 0$. Ist $x_1 = y_1$, so muss $|x_2 - y_2| = 0$ gelten und somit $x_2 = y_2$, also $x = y = (0, 0)$. Sonst muss die Summe der positiven Zahlen $|x_2|$, $|y_2|$ und $|x_1 - y_1|$ Null seien, also müssen alle Summanden Null seien, es gilt $x_2 = y_2 = x_1 = y_1 = 0$ und somit $x = y = (0, 0)$.

- ii Für alle $x, y \in \mathbb{R}^2$ gilt $d(x, y) = d(y, x)$, weil der Betrag und die Summation symmetrisch ist.

- iii Seien $x, y, z \in \mathbb{R}^2$.

Fall 1: $x_1 = z_1$. Dann gilt

$$\begin{aligned} d(x, z) &= |x_2 - z_2| \leq |x_2| + |z_2| + |x_1 - z_1| (+|y_2 - z_2|) \\ &\leq |x_2| + |z_2| + 2|y_2| + |x_1 - y_1 + y_1 - z_1| \\ &\leq |x_2| + |y_2| + |x_1 - y_1| + |z_2| + |y_2| + |y_1 - z_1| = d(x, y) + d(y, z). \end{aligned}$$

und

$$d(x, z) = |x_2 - z_2| \leq |x_2 - y_2 + y_2 - z_2| \leq |x_2 - y_2| + |y_2 - z_2|$$

Fall 2: $x_1 \neq x_2 \neq z_1$. Dann gilt

$$d(x, z) = |x_2| + |z_2| + |x_1 - z_1| \stackrel{\triangle\neq}{\leq} |x_2| + |y_2| + |x_1 - y_1| + |y_2| + |z_2| + |y_1 - z_1|$$

oder ($y_1 = z_1$)

$$d(x, z) \leq$$

Fall 3: $x_1 = x_2 \neq z_1$. Dann gilt

$$d(x, z) = |x_2| + |z_2| + |x_1 - z_1|$$

Fall 4: $x_1 \neq x_2 = z_1$. Dann gilt

$$d(x, z) = |x_2| + |z_2| + |x_1 - z_1| = |z_1| + |z_2| + |x_1 - z_1|$$

- ⑨ i Es gilt $d(0, \pi) = |\sin(0) - \sin(\pi)| = 0$ aber $0 \neq \pi$, also ist die Metrik nicht positiv definit.

- ii Weil die Abbildung $x \mapsto \tan^{-1}(x)$ injektiv ist, gilt $d(x, y) = 0 \iff x = y$. Aufgrund der Symmetrie und der Dreiecksungleichung des Betrags ist die Symmetrie und die Dreiecksungleichungen klar.

- iii

- 10) Die Abbildung f muss injektiv (bzw. strikt monoton, wenn f stetig seien sollte) seien, weil sonst die positive Definitheit nicht gilt. Weil für den Betrag Symmetrie und die Dreiecksungleichung gilt, sind diese Eigenschaften immer erfüllt.

Zum Beispiel erfüllt die folgende Funktion diese Eigenschaften:

$$f(x) := \begin{cases} \tan^{-1}(x), & \text{wenn } x \in \mathbb{Q}, \\ \tan^{-1}(x+1), & \text{wenn } x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

1.2 Topologie in metrischen Räumen

- Wir erklären, was **Umgebungen** und was **offene Mengen** in einem metrischen Raum sind und lernen deren wesentliche Eigenschaften kennen.
- Begriffe wie Konvergenz oder Stetigkeit lassen sich allein mit dem Offenheitsbegriff ohne weiteren Rückgriff auf die Metrik definieren. Das ist der Ausgangspunkt der Verallgemeinerung metrischer Räume zu sogenannten **topologischen Räumen**.

DEFINITION 1.2.1 (TOPOLOGISCHER RAUM, UMGEBUNG)

Sei X eine Menge und $\mathcal{T} \subset \mathcal{P}(X)$ eine Teilmenge der Potenzmenge von X .

- ① Man nennt das Mengensystem \mathcal{T} eine **Topologie** für X und (X, \mathcal{T}) einen **topologischen Raum**, wenn gilt
 - (T1) Vereinigungen beliebig vieler und Durchschnitte endlich vieler Mengen aus \mathcal{T} sind wieder in \mathcal{T} .
 - (T2) Es gilt $X \in \mathcal{T}$ und $\emptyset \in \mathcal{T}$.
 Die Mengen aus \mathcal{T} nennt man dann die **offenen Mengen** von (X, \mathcal{T}) .
- ② Ein beliebige Menge $U \subset X$ heißt eine **Umgebung** von $x \in X$, falls es eine offene Menge $V \in \mathcal{T}$ gilt, für die $x \in V \subset U$.

Bemerkung 1.2.2 (von mir) Wir lassen bewusst $X = \emptyset$ zu.

Beispiel 1.2.3 Die Standardtopologie auf \mathbb{R} ist die Menge aller offenen Mengen.

Beispiel 1.2.4 (von mir) Die einfachsten Topologien auf einer beliebigen Menge X sind die **indiskrete Topologie** $\mathcal{T} := \{\emptyset, X\}$ und die **diskrete Topologie** $\mathfrak{T} := \mathcal{P}(X)$.

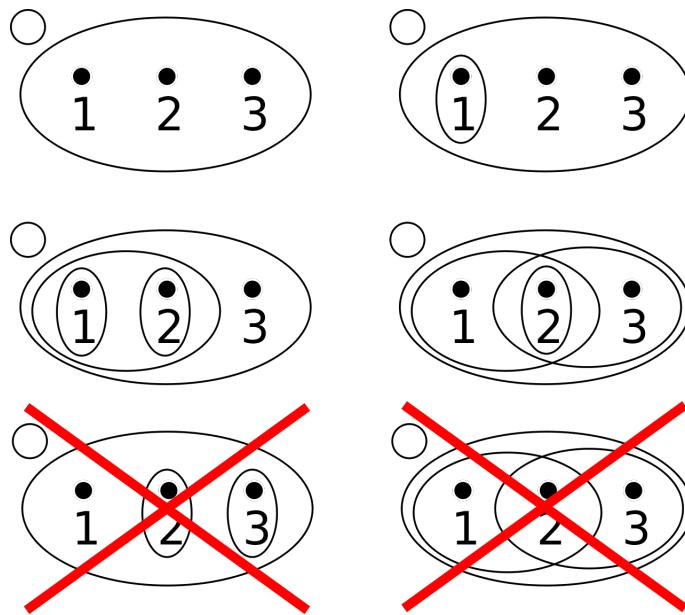


Abb. 5: Beispiele und Gegenbeispiele zu Topologien - die sechs Abbildungen stellen Teilmengen der Potenzmenge von $\{1, 2, 3\}$ dar (der kleine Kreis links oben ist jeweils die leere Menge). Die ersten vier sind Topologien: im Beispiel unten links fehlt $\{2, 3\}$, unten rechts $\{2\}$ zur Topologie-Eigenschaft. Quelle

DEFINITION 1.2.5 (UMGEBUNG, OFFEN, ABGESCHLOSSEN)

Seien (X, d) ein metrischer Raum und $a \in X$.

① Für $\varepsilon > 0$ heißt

$$U_\varepsilon(a) := \{x \in X \mid d(x, a) < \varepsilon\}$$

offene Kugel die offene Kugel vom Radius ε um a oder die (offene) ε -Umgebung von a .

Umgebung ② Eine Teilmenge $Y \subset X$ heißt eine Umgebung von a , wenn es $\varepsilon > 0$ gibt, so dass $U_\varepsilon(a) \subset Y$.

offen ③ Eine Teilmenge $Y \subset X$ heißt offen, wenn Y eine Umgebung jedes seiner Punkte $x \in Y$ ist, bzw. falls für jedes $x \in Y$ gilt: Y ist eine Umgebung von x .

abgeschlossen ④ Eine Teilmenge $Y \subset X$ heißt abgeschlossen, wenn $X \setminus Y$ offen ist.

Lemma 1.2.6 (Äquivalenz von Ana-I-Offenheit und neuer Definition)

In Analysis I definierten wir Offenheit im \mathbb{R}^n als

$$X \subset \mathbb{R}^n \text{ offen} \iff \forall x \in X \exists \varepsilon > 0 : U_\varepsilon(x) \subset X$$

Dies ist konsistent mit ③ aus der vorangegangenen Definition:

$$X \subset \mathbb{R}^n \text{ Ana-I-offen} \iff X \subset \mathbb{R}^n \text{ (iii)-offen.}$$

Beweis. " \implies ": klar, weil $U_\varepsilon(x)$ insbesondere eine Umgebung von x ist.

" \impliedby ": Für $x \in X$ existiert eine Umgebung $U \subset X$. In U existiert daher ein $U_\varepsilon(x)$ mit $\varepsilon > 0$. \square

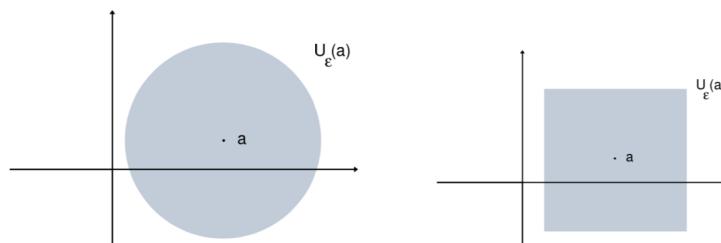


Abb. 6: Die ε -Kugeln Umgebungen für die Standardmetrik und für d^∞ auf \mathbb{R}^2 .

Beispiel 1.2.7 ① Für die Supremumsmetrik auf $\mathcal{B}(X, Y)$ besteht $U_\varepsilon(f)$ aus allen Funktionen $g : X \rightarrow Y$ mit $\sup_{x \in X} d(f(x), g(x)) < \varepsilon$.

② Sei $(X, d) = \mathbb{R}$ mit der Standardmetrik.

Die Mengen \emptyset , $(0, 1)$, $(1, \infty)$, $(-\infty, \infty)$ sind offen

Die Mengen \emptyset , $[0, 1]$, $[1, \infty[$, $(-\infty, \infty)$ sind abgeschlossen,

Die Menge $[0, 1]$ ist weder offen noch abgeschlossen.

③ In $X = (0, 1) \cup (2, 3)$ sind $(0, 1)$ und $(2, 3)$ offen und abgeschlossen. (vgl. Spurmetrik)

- ④ In jedem topologischen Raum (X, d) sind X und \emptyset sowohl offen als auch abgeschlossen.
- ⑤ Bezüglich der diskreten Metrik sind alle Teilmengen offen, also alle Teilmengen auch abgeschlossen.
- ⑥ Die offenen Kugeln $U_\varepsilon(a)$ sind offen.

Beweis. Zu ⑥: Für alle $x \in X$ gilt $U_\xi(x) = \{x\}$ für ein ξ mit $0 < \xi < 1$.

Oder: Sei z die diskrete Metrik. Dann gilt $B_1^z(0) = \{x \in \mathbb{R}^n : z(x, 0) < 1\} = \{0\}$ ist zugleich offen und abgeschlossen, da gilt, dass $\mathbb{R}^n = \bigcup_{x \in \mathbb{R}^n \setminus \{0\}} \{x\}$ als Vereinigung offener Mengen offen ist. Bezuglich z gilt also $\mathbb{R}^n = \{x \in \mathbb{R}^n : z(x, 0) \leq 1\}$ ist beschränkt.

Zu ⑦: TODO SELBST ÄHNLICH WIE IN ANA I □

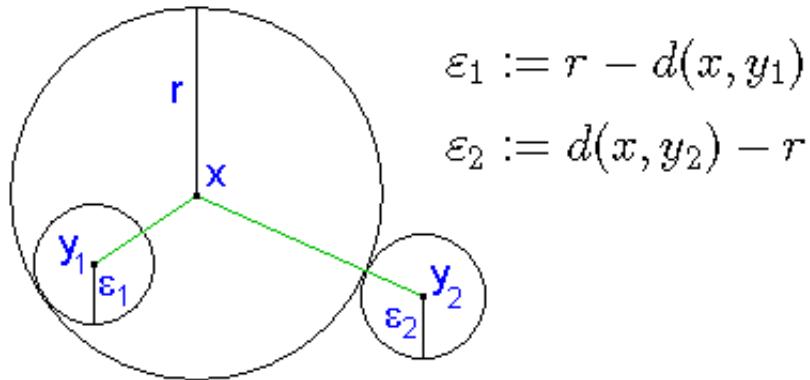


Abb. 7: Veranschaulichung, dass offen/abg. ε -Kugeln offen/abg. sind. [Quelle]

SATZ 1.2.1: DIE MENGE DER STETIGE FUNKTIONEN IST ABGESCHLOSSEN

In $(\mathcal{B}([a, b], \mathbb{R}), d^{\sup})$ ist die Menge der stetigen Funktionen

$$C^0([a, b]) := \{f : [a, b] \rightarrow \mathbb{R} \mid f \text{ stetig}\} \subset \mathcal{B}([a, b], \mathbb{R})$$

abgeschlossen.

Beweis. Zunächst sind stetige Funktionen auf einem kompakten Intervall beschränkt. Also ist wirklich $C^0([a, b]) \subset \mathcal{B}([a, b], \mathbb{R})$.

Wir zeigen, dass $\mathcal{B}([a, b], \mathbb{R}) \setminus C^0([a, b])$ offen ist. Sei $f \in \mathcal{B}([a, b], \mathbb{R}) \setminus C^0([a, b])$ bzw. sei $f \in \mathcal{B}([a, b], \mathbb{R})$ unstetig an der Stelle $x_0 \in [a, b]$. Wir suchen ein $\varepsilon > 0$, sodass $U_\varepsilon(f) \subset \mathcal{B}([a, b], \mathbb{R}) \setminus C^0([a, b])$ ist, d.h., dass jedes $g \in U_\varepsilon(f)$ unstetig ist.

Ist $f \in \mathcal{B}([a, b], \mathbb{R}) \setminus C^0([a, b])$, so gilt

$$\begin{aligned} \exists \varepsilon > 0 \quad \forall \delta > 0 \quad \exists x \in [a, b] : |x - x_0| < \delta \quad \text{und} \quad |f(x) - f(x_0)| \geq \varepsilon \\ \iff \exists \varepsilon > 0 \quad \forall \delta > 0 \quad \exists x \in U_\delta(x_0) \subset [a, b] : f(x) \notin U_\varepsilon(f(x_0)). \end{aligned}$$

Wir zeigen nun, dass jede beschränkte Funktion $g \in U_{\frac{\varepsilon}{3}}(f)$ unstetig ist.

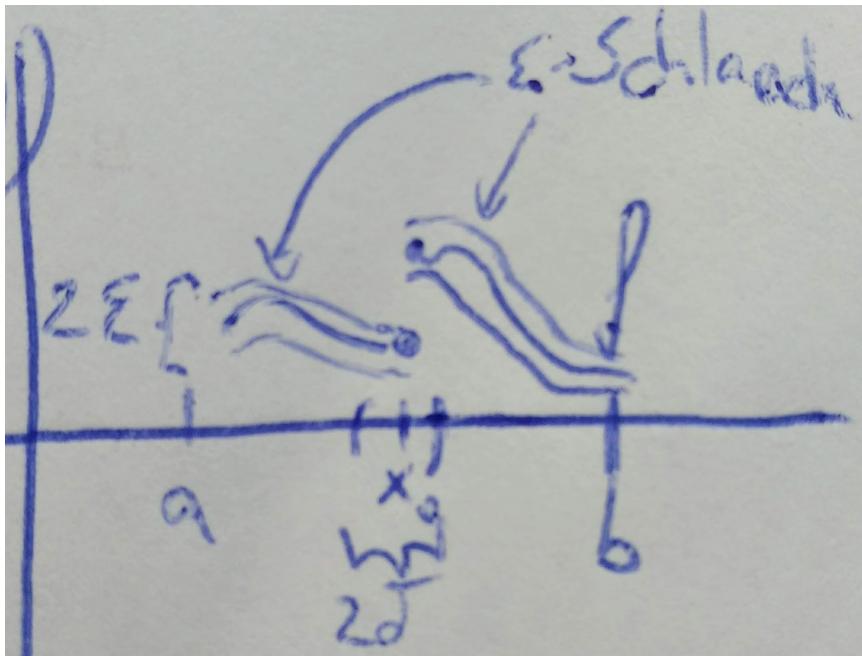
Sei nun $g \in U_{\varepsilon/3}(f)$, dann gilt $g \in \mathcal{B}([a, b], \mathbb{R})$ und $\sup |f - g| < \varepsilon/3$. Dann existiert zu jedem $\delta > 0$ ein $x \in [a, b]$ mit $|x - x_0| < \delta$ und

$$\varepsilon \leq |f(x) - f(x_0)| \stackrel{\Delta \neq}{\leq} \underbrace{|f(x) - g(x)|}_{<\varepsilon/3} + |g(x) - g(x_0)| + \underbrace{|g(x_0) - f(x_0)|}_{<\varepsilon/3}.$$

Also existiert zu jedem $\delta > 0$ ein $x \in [a, b]$ mit

$$|x - x_0| < \delta \quad \text{und} \quad |g(x) - g(x_0)| \geq \frac{\varepsilon}{3}.$$

Also ist g unstetig und $U_{\varepsilon/3}(f)$ besteht nur aus unstetigen Funktionen. Daher ist die Menge der unstetigen Funktionen in $\mathcal{B}([a, b], \mathbb{R})$ offen und das Komplement $C^0([a, b])$ abgeschlossen. \square



Bemerkung 1.2.8 Die Aussage gilt nur mit einem kompakten Intervall $[a, b]$, da stetige Funktionen auf offenen Intervallen wie z.B. auf ganz \mathbb{R} im Allgemeinen nicht beschränkt sind.

Bemerkung (Vorgriff) Wir zeigten eine Variante von "Der gleichmäßige Grenzwert stetiger Funktionen $[a, b] \rightarrow \mathbb{R}$ ist stetig" (vgl. Satz 1.7.4).

Das sieht man so: $C^0([a, b])$ ist abgeschlossen, also liegt der Grenzwert jeder konvergenten Folge in $C^0([a, b])$ wieder in $C^0([a, b])$. Beachte: Grenzwerte in $(\mathcal{B}([a, b], \mathbb{R}), d^{\sup})$ sind gleichmäßige Grenzwerte.

DEFINITION 1.2.9 (HAUSDORFF'SCH)

Ein topologischer Raum (X, \mathcal{T}) heißt haussdorff'sch, wenn zu je zwei verschiedenen Punkten $x, y \in X$ disjunkte Umgebungen von x bzw y existieren. Salopp gesagt kann man also je zwei Punkte mit offenen Mengen trennen.

Bemerkung 1.2.10 (von mir) HAUSDORFF'sch ist nur eine von vielen **Trennungsaxiomen** der Topologie.

SATZ 1.2.2

Metrische Räume sind HAUSDORFF'sche Räume.

Beweis. Sei $\varepsilon := d(x, y) > 0$. Wir setzen definieren $U := U_{\frac{\varepsilon}{2}}(x)$ und $V := U_{\frac{\varepsilon}{2}}(y)$. Dann gilt $x \in U$ und $y \in V$. Wäre $z \in U \cap V$, so wäre

$$d(x, y) \leq d(x, z) + d(z, y) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Das ist ein Widerspruch zur Definition von ε . Also gibt es kein $z \in U \cap V$. \square

Beispiel 1.2.11 (Nicht HAUSDORFF'sche topologische Räume (von mir))

- Die **kofiniten Topologie** \mathcal{T} auf einer unendlichen Menge X : in \mathcal{T} sind genau **alle** Teile Mengen offen, die leer sind, oder deren Komplemente endlich sind.¹ Somit sind die abgeschlossenen Mengen genau die endlichen Mengen.²
- Betrachte die Teilmenge $\mathfrak{T} := \{[a, \infty), a \in \mathbb{R}\} \cup \{\emptyset\} \subset \mathbb{R}^3$
Sie definiert eine Topologie auf \mathbb{R} , da die leere Menge per Definition in \mathfrak{T} enthalten ist, und $\bigcup_{a \in \mathbb{R}} [a, \infty) = \mathbb{R} \in \mathfrak{T}$ bzw. $\bigcup_{a \in A} [a, \infty) = [\min(A), \infty) \in \mathfrak{T}$ für eine beliebige Teilmenge $A \subset \mathbb{R}$ gilt. Ferner gilt $\bigcup_{x \in E} [a, \infty) = [\max(E), \infty) \in \mathfrak{T}$ für eine endliche Teilmenge $E \subset \mathbb{R}$. Die Topologie ist nicht HAUSDORFF'sche, weil sich je zwei nichtleere offene Mengen schneiden. Man sagt auch, \mathfrak{T} ist ein **irreduzierbarer Topologischer Raum**.

SATZ 1.2.3: METRISCHE RÄUME SIND TOPOLOGISCHE RÄUME

Seien (X, d) ein metrischer Raum und \mathcal{T} die Menge aller offenen Teilmengen von X . Dann ist \mathcal{T} eine Topologie auf X bzw. (X, \mathcal{T}) ein topologischer Raum.

Beweis. Sei $(U_i)_{i \in I}$ eine Familie offener Mengen in X . Ist $x \in \bigcup_{i \in I} U_i$, so existiert ein $i \in I$ mit $x \in U_i$. Weil U_i offen ist, existiert dazu ein $\varepsilon > 0$ mit

$$U_\varepsilon(x) \subset U_i \subset \bigcup_{j \in I} U_j.$$

Also ist $\bigcup_{i \in I} U_i$ offen.

Ist andererseits $x \in \bigcap_{i \in I} U_i$, so existiert zu jedem $i \in I$ ein $\varepsilon_i > 0$ mit $U_{\varepsilon_i}(x) \subset U_i$. Wir wählen zu jedem i ein solches $\varepsilon_i > 0$. Ist die Indexmenge I endlich, so ist

$$\varepsilon := \min\{\varepsilon_i \mid i \in I\}$$

positiv, und es gilt $U_\varepsilon(x) \subset U_{\varepsilon_i}(x) \subset U_i$ für jedes $i \in I$. Daher ist

$$U_\varepsilon(x) \subset \bigcap_{i \in I} U_i.$$

 \square

¹Da zum Beispiel die Nullstellenmenge von Polynomen über einem Körper \mathbb{K} , endlich oder ganz \mathbb{K} ist, ist die sog. **ZARISKI-Topologie** auf \mathbb{K} genau die kofiniten Topologie.

ZARISKI-Topologie

²Die kofiniten Topologie ist die kleinste Topologie, in welcher alle einteiligen Mengen abgeschlossen sind. Ist die Grundmenge X jedoch endlich, ist die kofiniten Topologie genau die **diskrete Topologie**, in welcher (ähnlich zur **diskreten Metrik**) alle Mengen offen und zugleich **abgeschlossen** sind, da jede Teilmenge $U \subset X$ und deren Komplement $U^C \subset X$ endlich sind, also $\mathcal{T} = \mathcal{P}(X)$ gilt.

³Mit dieser Topologie kann auch die Definition der **Unterhalbstetigkeit** herleiten: eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ ist genau dann unterhalbstetig, wenn sie bezüglich \mathfrak{T} stetig ist.

Für unendliches I klappt das letzte Argument des Beweises nicht und ist die Aussage auch nicht richtig:

Beispiel 1.2.12 Es gilt $\mathbb{R} \supset \bigcap_{n \in \mathbb{N}_{>0}} (-\frac{1}{n}, \frac{1}{n}) = \{0\}$, welches ein Gegenbeispiel liefert, da es ein unendlicher Schnitt offener Menge darstellt, der bezüglich der Standardmetrik auf \mathbb{R} nicht offen ist.

Korollar 1.2.13

Der Durchschnitt von beliebig vielen und die Vereinigung von endlich vielen abgeschlossenen Mengen sind wieder abgeschlossen.

Folgt direkt aus den

Lemma 1.2.14 (DEMORGANSchen Regeln)

$$\left(\bigcup_{i \in I} B_i \right)^C = \bigcap_{i \in I} B_i^C \quad \text{und} \quad \left(\bigcap_{i \in I} B_i \right)^C = \bigcup_{i \in I} B_i^C.$$

Beweis. Einfach, vgl z.B. [hier](#). □

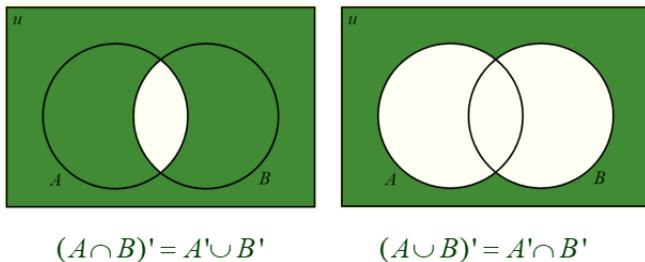


Abb. 8: Die DEMORGAN'schen Gesetze für zwei Mengen $A, B \subset u$. [Quelle](#).

DEFINITION 1.2.15 (SPUR- BZW. UNTERRAUMTOPOLOGIE)

Sei (X, \mathcal{T}) ein topologischer Raum und $Y \subset X$ eine Teilmenge. Dann ist $\mathcal{T}_y := \{U \cap Y : U \in \mathcal{T}\}$ eine Topologie auf Y , die sogenannte Spur- oder Unterraumtopologie, also ist (Y, \mathcal{T}_y) ein topologischer Raum.

Beweis. (von mir) Die Spurtopologie ist eine Topologie auf Y , da die leere Menge in \mathcal{T} enthalten ist und somit $\emptyset \cap Y = \emptyset \in \mathcal{T}_Y$ gilt. Nun gilt $X \in \mathcal{T}$ und somit $X \cap Y = Y \in \mathcal{T}_Y$. Sei nun $(Y_i)_{i \in I}$ eine Familie in \mathcal{T}_Y . Dann können wir Y_i als $U_i \cap Y$ für $U_i \in \mathcal{T}$ für jedes $i \in I$ schreiben:

$$\bigcup_{i \in I} Y_i = \bigcup_{i \in I} (U_i \cap Y) \stackrel{(*)}{=} Y \cap \underbrace{\bigcup_{i \in I} U_i}_{\in \mathcal{T}} \in \mathcal{T}_y,$$

wobei wir in $(*)$ ein Distributivgesetz verwendet haben, welches in der nächsten Abbildung für zwei Mengen U und V veranschaulicht wird. Sei nun $(Y_i)_{i \in \{1, \dots, n\}}$ für $n \in \mathbb{N}$ eine endliche Familie in \mathcal{T}_Y . Dann gilt (vgl. oben)

$$\bigcap_{i=1}^n Y_i = \bigcap_{i=1}^n (U_i \cap Y) = U_1 \cap \dots \cap U_n \cap Y = Y \cap \underbrace{\bigcap_{i=1}^n U_i}_{\in \mathcal{T}} \in \mathcal{T}_Y. \quad \square$$

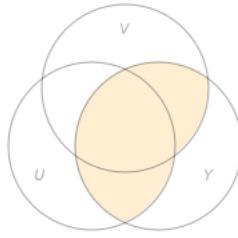


Abb. 9: Das Distributivgesetz $\bigcup_{i \in I} (X_i \cap Y) = Y \cap \bigcup_{i \in I} X_i$. [Quelle: WA]

Beispiel 1.2.16 (von mir) Die Spurtopologie auf $\mathbb{Z} \subset \mathbb{R}$ (analog für \mathbb{N}) ist die **diskrete Topologie**. Bspw. ist die Menge $\{a\} = \left(a - \frac{1}{k+1}, a + \frac{1}{k+1}\right) \cap \mathbb{Z}$ für ein beliebiges $k \in \mathbb{N}$ offen in \mathbb{Z} .

**SATZ 1.2.4: DIE SPURTOPOLOGIE IN EINEM METRISCHEN RAUM
IST DIE TOPOLOGIE DES METRISCHEN TEILRAUMS**

Seien (X, d) ein metrischer Raum und $A \subset X$ eine beliebige Teilmenge versehen mit der induzierten Metrik d_A . Dann sind die offenen Teilmengen von (A, d_A) genau die Durchschnitte offener Teilmengen von (X, d) mit A : $B \subset A$ ist genau dann offen in (A, d_A) , wenn eine offene Teilmenge Y in (X, d) existiert, sodass $B = A \cap Y$ gilt.

Anders ausgedrückt: Seien (X, d) ein metrischer Raum und (A, d_A) ein metrischer Teilraum. Dann ist die von d_A induzierte Topologie \mathcal{T}_{d_A} auf A gleich der Spurtopologie \mathcal{T}_A .

Der Satz gilt auch mit "abgeschlossen" statt "offen".

Beweis. TODO: Bilder

„ \implies “ bzw. $\mathcal{T}_{d_A} \subset \mathcal{T}_A$.

Sei $B \subset A$ offen bzw. sei $B \subset \mathcal{T}_{d_A}$ und

$$U_\varepsilon^Z(x) := \{y \in Z : d_A(x, y) = d(x, y) < \varepsilon\}$$

für eine Teilmenge $Z \in X$. Dann existiert zu jedem $x \in B$ ein $\varepsilon_x > 0$ mit

$$B \supset U_{\varepsilon_x}^A(x) = A \cap U_{\varepsilon_x}^X(x).$$

Die Menge

$$Y := \bigcup_{x \in B} U_{\varepsilon_x}^X(x)$$

ist als Vereinigung offener Teilmengen von X offen in X und es gilt $B = A \cap Y$, „ \supset “ folgt aus dem Obigen, „ \subset “ aus der Definition von $Y \supset B$. Damit ist $B \in \mathcal{T}_A$.

„ \Leftarrow “ bzw. $\mathcal{T}_A \subset \mathcal{T}_{d_A}$ (Selbst.)

Sei $B \subset A$ eine Menge mit der Eigenschaft $B = Y \cap A$ für ein offene Teilmenge $Y \in X$ und $x \in B$. Dann gilt auch $x \in Y$. Da Y offen in X ist, existiert ein $\varepsilon > 0$, sodass $x \in U_\varepsilon^X(x) \subset Y$ gilt. Nun gilt jedoch auch $x \in U_\varepsilon^A(x) = U_\varepsilon^X(x) \cap A \subset Y \cap A = B$. Somit ist $B \subset A$ in (A, d_A) offen.

Nun zeigen wir die Behauptung mit „abgeschlossen“ mittels des eben Bewiesenen:
 „ \implies “: Sei $C \subset A$ eine abgeschlossene Teilmenge. Dann ist $B := C^C \cap A$ offen in A

und nach dem eben Bewiesenen existiert eine in X offene Teilmenge $Y \subset X$ gilt $B = A \cap Y$. Sei nun $D := Y^C$ ein Teilmenge, welche als Komplement einer offenen Menge abgeschlossen ist. Weil $C \subset A$ gilt, folgt

$$C = B^C \cap A = (A \cap Y)^C \cap A = (A^C \cup Y^C) \cap A = Y^C \cap A = D \cap A.$$

„ \Leftarrow “: Wir nehmen an, es existiert eine abgeschlossene Teilmenge $D \subset X$, sodass $C = D \cap A$ gilt. Sei nun $B := C^C \cap A$, dann gilt

$$B = C^C \cap A = (D \cap A)^C \cap A = (D^C \cap A^C) \cap A = D^C \cap A.$$

Nun ist D^C als Komplement einer abgeschlossenen Menge offen, also ist B nach dem eben Bewiesenen offen in S und dann ist $C = B^C \cap A$ abgeschlossen in A . \square

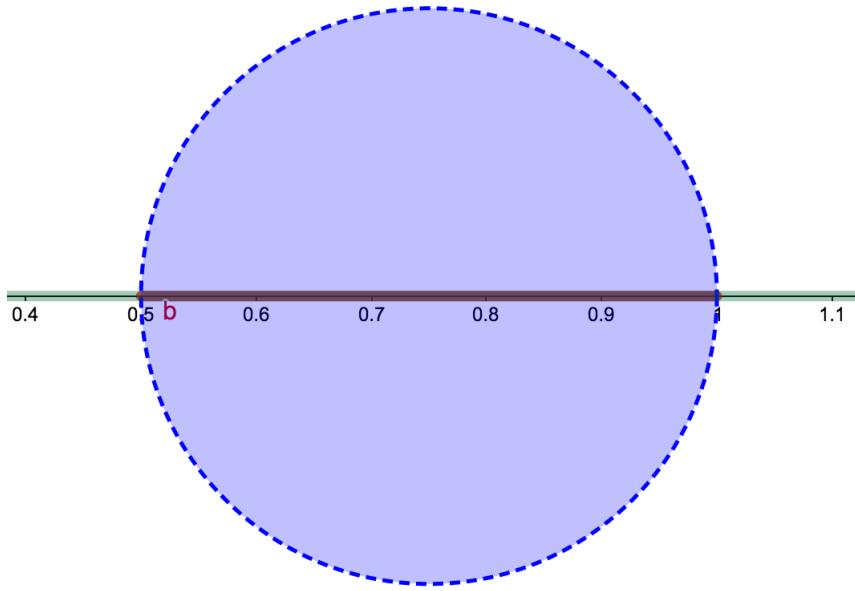


Abb. 10: Hier ist X die x - y -Ebene, A die ***x-Achse*** und ***b*** das offene Intervall $(0.5, 1)$. Die in X offene Menge Y ist die ***offene Kreisscheibe***

Beispiel 1.2.17 Sei $X := (0, 3]$ mit der Standardmetrik. Überlegen Sie, welche Attribute auf welche Teilmengen von X zutreffen, welche nicht:

	$(0, 1]$	$(2, 3]$	$(1, 2]$	$(0, 3]$
offen	Nein	Ja	Nein	Ja
abgeschlossen	Ja	Nein	Nein	Ja

Um diese intuitiven Resultat zu bestätigen, benutzt man den vorangegangen Satz: $(0, 1] = [0, 1] \cap (0, 3]$ für die Abgeschlossenheit der ersten Menge und $(2, 4)$ für die Offenheit der zweiten. Das Komplement der ersten Menge ist $(1, 3]$ und somit wie $(2, 3]$ nicht abgeschlossen, also kann die erste Menge nicht offen seien. Das Komplement dritte Menge ist $(0, 1) \cup [2, 3]$ ist als Vereinigung einer in X offen und einer in X abgeschlossenen Menge weder offen noch abgeschlossen, also kann die Menge selbst auch weder noch seien. Die gesamte Menge ist nach Definition immer offen und abgeschlossen.

DEFINITION 1.2.18 (INNERES, ABGESCHLOSSENE HÜLLE, RAND)

Seien (X, d) ein metrischer Raum und $Y \subset X$ eine Teilmenge.

- ① Das **Innere von Y** ist die Menge der **innere Punkte** von Y :

$$\overset{\circ}{Y} := Y^0 := \{y \in Y \mid \exists \varepsilon > 0 : U_\varepsilon(y) \subset Y\}.$$

- ② Die **abgeschlossene Hülle** oder **Abschluss** von Y ist die Menge

$$\bar{Y} := X \setminus (X \setminus Y)^0 = ((Y^C)^\circ)^C.$$

- ③ Der **Rand von Y** ist die Menge $\partial Y := \bar{Y} \setminus \overset{\circ}{Y}$.

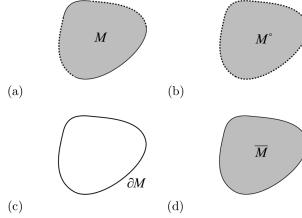


Abb. 11: (a) Eine Teilmenge $M \subset \mathbb{R}^2$, (b) ihr Inneres, (c) ihr Rand und (d) ihr Abschluss. [Quelle]

Bemerkung Das sind alles topologische Begriffe! In einem topologischen Raum (X, \mathcal{T}) ist

$$\overset{\circ}{Y} := \{y \in Y : y \text{ ist ein innerer Punkt von } Y\} = \{y \in Y \mid \exists U \in \mathcal{T} : y \in U \subset Y\}.$$

Die anderen Definitionen können so übernommen werden, da kein Abstandsbegriff benötigt wird.

SATZ 1.2.5: CHARAKTERISIERUNGEN DES INNEREN, ABSCHLUSSES UND DES RANDS

Seien (X, d) ein metrischer Raum und $Y \subset X$ eine Teilmenge. Dann gilt

- ① $\overset{\circ}{Y}$ ist die größte offene Menge in Y :

$$\overset{\circ}{Y} = \bigcup_{\substack{U \text{ offen} \\ U \subset Y}} U.$$

Insbesondere ist $\overset{\circ}{Y}$ als Vereinigung offener Mengen selber offen.

- ② \bar{Y} ist die kleinste abgeschlossene Menge, die Y enthält:

$$\bar{Y} = \bigcap_{\substack{A \text{ abg.} \\ A \supseteq Y}} A.$$

Insbesondere ist \bar{Y} als Durchschnitt abgeschlossener Mengen selber abgeschlossen.

- ③ Die Randpunkte von Y sind charakterisiert dadurch, dass jede ihrer Umgebungen Punkte von Y und Punkte von $X \setminus Y$ enthält:

$$\partial Y = \{x \in X \mid \forall \varepsilon > 0 U_\varepsilon(x) \cap Y \neq \emptyset \neq U_\varepsilon(x) \cap (X \setminus Y)\}.$$

∂Y ist abgeschlossen.

Dies lässt sich auch in topologischen Räumen machen!

Beweis. ① Sei $V := \bigcup_{\substack{U \text{ offen} \\ U \subset Y}} U$.

” \subseteq “: Ist $y \in \overset{\circ}{Y}$, so existiert $\varepsilon > 0$ mit $U := U_\varepsilon(y) \subset Y$, und U ist offen. Also ist $y \in U \subset V$.

” \supseteq “: Ist $y \in V$, so existiert ein offenes $U \subset Y$ mit $y \in U$. Da U offen ist, existiert ein $\varepsilon > 0$ mit $U_\varepsilon(y) \subset U \subset Y$. Also ist y ein innerer Punkt, es gilt $y \in \overset{\circ}{Y}$.

② Mit den DEMORGANSchen Regeln erhält man

$$\overline{Y} = ((Y^C)^C)^C = \left(\bigcup_{\substack{Y^C \supseteq U \\ U \text{ offen}}} U \right)^C = \bigcap_{\substack{Y^C \supseteq U \\ U \text{ offen}}} U^C = \bigcap_{\substack{Y \supseteq U^C \\ U \text{ offen}}} U^C = \bigcap_{\substack{A \text{ abgeschl.} \\ A \supseteq Y}} A.$$

③ Nach Definition ist

$$\partial Y := \overline{Y} \setminus Y^\circ = \left((Y^C)^\circ \right)^C \setminus Y^\circ = \left((Y^C)^\circ \cup Y^\circ \right)^C.$$

Hierbei verwenden wir $A \setminus B = B^C \cup A$. Diese Menge ist das Komplement der Vereinigung der inneren Punkte von Y und Y^C . \square

Beispiel 1.2.19 Seien

$$U := \{(x, y, 0) \in \mathbb{R}^3 \mid x^2 + y^2 < 1\} \subset \mathbb{R}^3$$

$$E := \{(x, y, 0) \in \mathbb{R}^3 \mid x, y \in \mathbb{R}\} \subset \mathbb{R}^3$$

ein Kreis in der x - y -Ebene und die x - y -Ebene selbst.

Wir betrachten den \mathbb{R}^3 mit der Standardmetrik d . Dann ist das Innere $\overset{\circ}{U} = \emptyset$ und

$$\partial U = \bar{U} = \{(x, y, 0) \in \mathbb{R}^3 \mid x^2 + y^2 \leq 1\}.$$

Betrachtet man hingegen U als Teilmenge von (E, d_E) , so ist

$$\overset{\circ}{U} = U,$$

$$\bar{U} = \{(x, y, 0) \in \mathbb{R}^3 \mid x^2 + y^2 \leq 1\},$$

$$\partial U = \{(x, y, 0) \in \mathbb{R}^3 \mid x^2 + y^2 = 1\}.$$

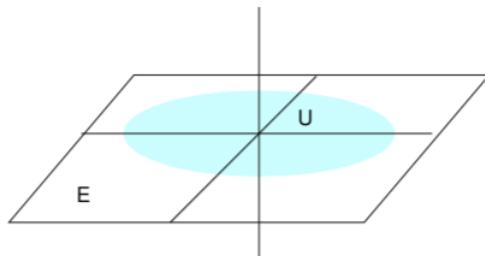


Abb. 12: Die Ebene E und die Kreisscheibe U . [Quelle: Ferus]

Aufgaben

(1) Skizzieren Sie für $n \in \{1, 2, 3\}$ und $p \in \{1, 2, \infty\}$ die Kugel $U_1^p := \{y \in \mathbb{R}^n \mid d_p(x, y) < 1\}$.

(2) Betrachten Sie die folgenden Mengen

$$\begin{aligned} A &:= \{(x, y) \in \mathbb{R}^2 \mid x = 2\}, \\ B &:= \{x \in \mathbb{R}^2 : d_2(0, x)^2 < 4\}, \\ C &:= \{x \in \mathbb{R}^2 : d_2(0, x)^2 \leq 4\}, \\ D &:= \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 4\}, \\ E &:= \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 4\}, \\ F &:= \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 4, z \in [0, 2)\}, \\ G &:= \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 \leq z^2, z \in [0, 2)\}, \\ H &:= \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = z^2, z \in [0, 2)\}, \\ I &:= \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq 16\}. \end{aligned}$$

- (i) Skizzieren Sie die Mengen.
 - (ii) Bestimmen Sie jeweils den Rand und das Innere.
 - (iii) Bestimmen Sie, ob die Mengen offen, abgeschlossen oder beschränkt sind.
- (3) Sei (X, d) ein metrischer Raum. Zeigen Sie, dass
- (i) jede endliche Teilmenge E von X abgeschlossen ist.
 - (ii) der Grenzwert einer konvergenten Folge $(a_k)_{k \in \mathbb{N}}$ eindeutig ist.
- Gelten die Aussagen, wenn X ein topologischer Raum mit der Hausdorff-Eigenschaft ist?
- (4) Sei (X, \mathcal{T}) ein topologischer Raum, und $A \subseteq B \subseteq X$ Teilmengen, welche wir mit der Spurtopologien \mathcal{T}_A bzw. \mathcal{T}_B versehen. Zeigen Sie, dass,
- (i) die Topologie \mathcal{T}_A mit der Spurtopologie, welche von \mathcal{T}_B induziert wird, übereinstimmt.
 - (ii) falls A offen in X ist, und $U \subseteq A$ eine Teilmenge ist, so gilt $U \in \mathcal{T}_A \iff U \in \mathcal{T}$. Bleibt dies richtig, wenn A nicht offen ist?
- (5) Sei (X, \mathcal{T}) ein topologischer Raum, A eine abgeschlossene und O eine offene Menge. Zeigen Sie, dass
- (i) $A \setminus O$ abgeschlossen ist.
 - (ii) $O \setminus A$ offen ist.



(6) Sei (X, \mathcal{T}) ein topologischer Raum und $A, B \subset X$ Teilmengen. Beweisen oder widerlegen Sie die folgenden Gleichheiten

- (i) $\overline{A \cup B} = \overline{A} \cup \overline{B}$
- (ii) $\overline{A \cap B} = \overline{A} \cap \overline{B}$

iii) $\overline{A \setminus B} = \overline{A} \setminus \overline{B}$

7) Sei $f : X \rightarrow Y$ eine Abbildung und $\mathcal{F} := \{W \subseteq Y \mid f^{-1}(W)$ ist offen in $(X, d)\}$.

- 1) Zeigen Sie, dass \mathcal{F} eine Topologie auf Y ist.
- 2) Beschreiben Sie die Spurtopologie auf der Teilmenge $\tilde{Y} := Y \setminus \{f(x) \mid x \in X\}$.



8) (von hier) Sind beliebige Schnitt und Vereinigungen von Topologien wieder eine Topologie?

Lösung auch [hier](#).

9) Zeigen Sie, dass die kofiniten Topologie eine Topologie auf jeder unendlichen Menge X definiert, welche nicht HAUSDORFF'sch ist.

10) (von hier) Zeigen Sie, dass die **koabzählbare Topologie**

$$\mathcal{T} := \{U \subset X : U = \emptyset \text{ oder } U^C \text{ ist abzählbar}\}$$

welche nah mit der **kofiniten Topologie** verwandt ist, eine Topologie auf einer unendlichen Menge X definiert (sonst ist \mathcal{T} wieder die diskrete Topologie), jedoch (X, \mathcal{T}) nicht HAUSDORFF'sch ist.

Lösung auch [hier](#).

11) (von hier) Seien $U_1 \subset U_2 \subset \dots \subset U_n \subset \dots \subset X$ eine aufsteigend ineinander enthaltene Familie von Mengen. Dann heißt das Mengensystem

$$\mathcal{T} := \{\emptyset, X\} \cup \{U_i : i \in I\} \cup \left\{ \bigcup_{k=1}^{\infty} U_k \right\}$$

verschachtelte die **verschachtelte Topologie** auf X . Verifizieren Sie die Topologeeigenschaften **Topologie** und zeigen Sie, für die **verschachtelte Intervalltopologie auf $X := (0, 1)$** (wähle $U_n := (0, \frac{n}{n+1})$ für $n \in \mathbb{N}_{>0}$), dass (X, \mathcal{T}) kein HAUSDORFF-Raum ist.

Weitere Beispiele in den reellen Zahlen sind $U_k := (-k, k)$ oder $V_k := [-k, k]$ für $k \in \mathbb{Z}$ sowie $W_n := \{1, \dots, n \mid n \in \mathbb{N}\}$ oder $S_n := \{n, n+1, \dots \mid n \in \mathbb{N}\}$ (die Ausgangs- bzw. Finalsegmenttopologie).

12) Geben Sie ein Beispiel und eine Begründung für einen topologischen Raum, der kein metrischer Raum ist (man sagt, der top. Raum ist nicht **metrisierbar**).

Lösungen

- ① Für $n = 1$ ist U_1^p immer durch die Fläche zwischen $y = x \pm 1$ und nicht auf diesen Gerade liegt, gegeben.

Für $n = 2$ gilt (o.B.d.A $x = 0$)

$$U_1^1(0) = \{y \in \mathbb{R}^2 \mid |y_1| + |y_2| < 1\}$$

(Eine offene Raute mit den Eckpunkten $(0| \pm 1)$ und $(\pm 1| 0)$)

$$U_1^2(0) = \{y \in \mathbb{R}^2 \mid \sqrt{y_1^2 + y_2^2} < 1\}$$

(Eine offene Kreisscheibe um den Ursprung mit Radius 1)

$$U_1^\infty(0) = \{y \in \mathbb{R}^2 \mid \max\{|y_1|, |y_2|\} < 1\}$$

(Ein offenes Quadrat mit den Eckpunkt $(\pm 1| \pm 1)$)

Für $n = \infty$ gilt (o.B.d.A $x = 0$)

$$U_1^1(0) = \{y \in \mathbb{R} : |y| < 1\} \quad (\text{Offene Einheitskreisscheibe um den Ursprung})$$

$$U_1^2(0) = \{(y_1, y_2) \in \mathbb{R}^2 : |y_1|, |y_2| < 1\}$$

(Ein offenes Einheitsquadrat um den Ursprung)

$$U_1^\infty(0) = \{(y_1, y_2, y_3) \in \mathbb{R}^3 : |y_1|, |y_2|, |y_3| < 1\}$$

(Ein offener Einheitswürfel um den Ursprung)

- ② i) A ist eine vertikale Gerade bei $x = 2$ im \mathbb{R}^2 .

B ist eine offene Kreisscheibe mit Radius 2 und Mittelpunkt $(0, 0)$.

C ist eine abgeschlossene Kreisscheibe mit Radius 2 und Mittelpunkt $(0, 0)$.

D ist ein Kreis mit Radius 2 und Mittelpunkt $(0, 0)$.

E ist dieser Kreis in der $z = 0$ Ebene.

F ist ein Zylinder mit Grundfläche E, der keinen Deckel bei $z = 2$ besitzt.

G ist ein ausgefüllter Kegel der durch $x = y$ begrenzt wird und keinen Deckel bei $z = 2$ besitzt.

H ist die Oberfläche eines Kegels, der durch $x = y$ begrenzt wird und keinen Deckel bei $z = 2$ besitzt.

I ist eine Sphäre mit Radius 4 um den Ursprung.

- ii) Für die Mengen A, D, E, F und H gilt, dass ihr Inneres leer ist und ihr Rand die Menge selbst ist. Für alle anderen Mengen gilt

$$\partial B = D, \quad \mathring{B} = B \quad \partial C = D \quad \mathring{C} = B \quad \partial G = H \quad \mathring{G} = \{(x, y, z) \in$$

$$\mathbb{R}^3 : x^2 + y^2 < z^2, z \in [0, 2)\} \quad \partial I = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 16\}$$

$$\mathring{I} = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 < 16\}$$

- iii) G ist weder offen noch abgeschlossen.

- ③ i) Sei $x \in X \setminus E := E^C$. Wir suchen ein $\varepsilon > 0$, sodass $U_\varepsilon(x) \subset E^C$. Wähle $\varepsilon := \min_{e \in E} d(e, x)$. Dann gilt für alle $e \in E$, dass $0 \leq d(e, x) < \varepsilon$ ist. Dann ist E abgeschlossen, denn es gilt

$$U_\varepsilon(x) \cap E = \{y \in X : d(x, y) \leq \varepsilon\} \cap E = \emptyset \implies U_\varepsilon(x) \subseteq E^C \implies E^C \text{ offen}$$

- ii) Sei $(a_k)_{k \in \mathbb{N}} \subset X$ mit $\stackrel{k \rightarrow \infty}{\rightarrow} a \in X$ und $\stackrel{k \rightarrow \infty}{\rightarrow} b \in X$. Widerspruchsbeweis: Sei $a \neq b$.

Dann gilt $d(a, b) > 0$. Dann gilt

$$\exists K_1 \in \mathbb{N} \forall k > K_1 : d(a_k, a) < \frac{d(a, b)}{2} \quad \text{und} \quad \exists K_2 \in \mathbb{N} \forall k > K_1 : d(a_k, b) < \frac{d(a, b)}{2}.$$

Somit gilt für alle $k > \max\{k_1, k_2\}$

$$d(a, b) \leq d(a_k, a) + d(a_k, b) < \frac{d(a, b)}{2} + \frac{d(a, b)}{2} = d(a, b),$$

was einen Widerspruch darstellt. Alternativ kann man die Aussage auch direkt beweisen.

$$0 \leq d(a, b) \leq d(a_k, a) + d(a_k, b) \xrightarrow{k \rightarrow \infty} 0 \stackrel{\text{Sandwich}}{\implies} d(a, b) = 0 \iff a = b,$$

was auch einen Widerspruch darstellt.

- (iii) Angenommen es existiert eine Folge $(a_k)_{k \in \mathbb{N}} \subset X$ mit $\xrightarrow{k \rightarrow \infty} a \in X$ und $\xrightarrow{k \rightarrow \infty} b \in X$ und $a \neq b$. Da (X, \mathcal{T}) haussdorff'sch ist, gilt

$$\exists U_a, U_b \subset \mathcal{T} : a \in U_a, b \in U_b, U_a \cap U_b = \emptyset$$

Da a_k gegen a, b konvergiert, müssen U_a und U_b ab einem bestimmten Index (vgl. (b)) alle Folgenglieder enthalten, sie sind jedoch disjunkt.

(4)

Sei nun X ein topologischer Raum mit der Hausdorff-Eigenschaft.

- (1) Wir betrachten $x \in E^C$. Da (X, \mathcal{T}) haussdorff'sch ist, gilt

$$\forall e \in E \exists U_x^e, U_e \in \mathcal{T} \text{ mit } x \in U_x^e, e \in U_e, \text{ sodass } U_x^e \cap U_e = \emptyset.$$

Somit gilt

$$\left(\bigcap_{e \in E} U_x^e \right) \cap U_e = \emptyset \quad \forall e \in E \implies \left(\bigcap_{e \in E} U_x^e \right) \cap E = \emptyset \implies \bigcap_{e \in E} U_x^e \subseteq E^C$$

weil $|E| < \infty$ ist, und der endliche Schnitt offener Mengen wieder offen ist. Dann gilt

$$\bigcup_{x \in E^C} \left(\bigcap_{e \in E} U_x^e \right) = E^C.$$

Weil $\bigcap_{e \in E} U_x^e$ offen ist, und somit auch seine endliche Vereinigungen, ist E^C offen und somit E abgeschlossen.

- (2) **Beweis.** (i) Wir zeigen stattdessen, dass $X \setminus (A \setminus O)$ offen ist.

Beweis. Es gilt

$$X \setminus (A \setminus O) = \{x \in X : x \notin (A \setminus O)\} = (X \setminus A) \cup O,$$

denn es gilt

$$A \setminus O = \{x \in A : x \notin O\} = \{x \in A : x \notin A \cap O\}.$$

$X \setminus A$ ist offen in X , weil A abgeschlossen ist, und O ist offen, und weil die Vereinigung offener Mengen offen ist, ist $X \setminus (A \setminus O)$ offen. \square

- (ii) Nach der selben Argumentation ist $O \setminus A = (X \setminus A) \cap O$, welches ein Schnitt in X offener Mengen ist. \square



- (3) **TODO:** einfache Lösungen für Beweise finden

(4) (i) **Beweis.** Zu (1): Es gilt

$$\emptyset \in \mathcal{F} \iff \underbrace{\emptyset \subseteq Y}_{\text{axiomatisch erfüllt}} \wedge f^{-1}(\emptyset) \text{ ist offen in } (X, d).$$

Es gilt $f^{-1}(\emptyset) = \emptyset$, da sonst ein x existiert mit $f(x) = \emptyset$, was einen Widerspruch darstellt. \emptyset ist per Definition in (X, d) offen, da (X, d) ein metrischer Raum ist.

Analog gilt trivialerweise $Y \subseteq Y$ und da $f^{-1}(Y) = X$ ist, ist $f^{-1}(Y)$ in (X, d) offen, da (X, d) ein metrischer Raum ist.

Zu (2): Sei nun $U_i \in \mathcal{F}$ für alle $i \in I$, wobei I eine beliebige Indexmenge ist. Dann gilt für alle

$$x \in f^{-1} \left(\bigcup_{i \in I} U_i \right) \iff x \in f^{-1}(U_i) \text{ für ein } i \in I \iff x \in \bigcup_{i \in I} f^{-1}(U_i)$$

Da $\bigcup_{i \in I} f^{-1}(U_i)$ eine Vereinigung offener Mengen ist, ist $f^{-1}(\bigcup_{i \in I} U_i)$ auch offen und es gilt $\bigcup_{i \in I} U_i \in \mathcal{F}$.

Zu (3): Sei I ein beliebiger Indexmenge, dann gilt für alle $x \in W$

$$\begin{aligned} x \in f^{-1} \left(\bigcap_{i \in I} U_i \right) &\iff f(x) \in \bigcap_{i \in I} U_i \\ &\iff f(x) \in U_i \ \forall i \in I \\ &\iff x \in f^{-1}(U_i) \ \forall i \in I \\ &\iff x \bigcap_{i \in I} f^{-1}(U_i). \end{aligned} \quad \square$$

(ii) Sei $y \in \tilde{Y}$ beliebig. Dann gilt $y \in Y$ und $y \notin \{f(x) : x \in X\} \iff f^{-1}(\{y\}) = \emptyset$. Somit ist $\{y\}$ offen in $(\tilde{Y}, \mathcal{F}_{\tilde{Y}})$. Dann gilt nach Satz 24 bzw. 2a, dass eine offene Teilmenge $V \subseteq Y$ mit $\{y\} = \tilde{Y} \cap V$ existiert. Wählen wir $V := \{y\}$, dann ist $\{y\} = \tilde{Y} \cap \{y\}$ erfüllt, und da Eigenschaft (2) von Topologien (s. 2b 1)) gilt, ist, weil y beliebig gewählt wurde, jede Vereinigung von Elementen aus \tilde{Y} in $\mathcal{F}_{\tilde{Y}}$. Somit gilt $\mathcal{F}_{\tilde{Y}} \supseteq \mathcal{P}(\tilde{Y})$, da nach der Definition einer Topologie aber immer $\mathcal{P}(\tilde{Y}) \supseteq \mathcal{F}_{\tilde{Y}}$ gilt, muss $\mathcal{P}(\tilde{Y}) = \mathcal{F}_{\tilde{Y}}$ gelten.



(5) **Beweis. Vereinigung:** Schon die Vereinigung zweier Topologien muss keine Topologie sein: Wähle zum Beispiel die Topologien $\mathcal{T}_1 := \emptyset, \{a\}, \{a, b\}, X$ und $\mathcal{T}_2 := \emptyset, \{b, c\}, X$ auf der Menge $X := \{a, b, c\}$. Nun ist jedoch $\{a, b\} \cap \{b, c\} = \{b\}$ als Schnitt von Mengen in $\mathcal{T}_1 \cup \mathcal{T}_2$ nicht in $\mathcal{T}_1 \cup \mathcal{T}_2$.

Schnitt: Wir zeigen, dass wenn $(\mathcal{T}_k)_{k \in I}$ eine Familie von Topologien auf X ist, auch $\mathcal{T} := \bigcap_{k \in I} \mathcal{T}_k$ eine Topologie auf X ist. Dafür genügt es, die Behauptung für $I = \{1, 2\}$ zu zeigen, indem wir die beiden Topologieeigenschaften überprüfen:

(T1) Weil jede Topologie die leere und die gesamte Menge enthalten, enthält sie auch der Schnitt

(T2) Sei $(U_i)_{i \in I} \subset \mathcal{T}_1 \cap \mathcal{T}_2$ eine Familie offener Mengen und $U := \bigcup_{i \in I} U_i$. Dann gilt $U_i \in \mathcal{T}_1$ und $U_i \in \mathcal{T}_2$ für alle $i \in I$. Daraus folgt $U \in \mathcal{T}_1$ und $U \in \mathcal{T}_2$ und somit $U \in \mathcal{T}_1 \cap \mathcal{T}_2$. Genau so argumentiert man für endliche Schnitte. \square

- 6 Es handelt sich hierbei um eine Topologie auf X , da die leere Menge per Definition und die gesamte Menge X in \mathcal{T} sind, weil das Komplement der gesamten Menge die leere Menge ist, welche endlich ist. Sei $(U_i)_{i \in I} \subset \mathcal{T}$ eine Familie in \mathcal{T} . Mit den **DEMORGAN'schen Regeln** gilt $(\bigcup_{i \in I} U_i)^C = \bigcap_{i \in I} U_i^C =: \tilde{U}$. Ist $U_i^C = \emptyset$ für ein $i \in I$, so ist $\tilde{U} = \emptyset \in \mathcal{T}$. Wenn nicht, so ist \tilde{U} als Schmitt endlicher Mengen endlich und somit ist \mathcal{T} . Sei nun $(U_i)_{i \in \{1, \dots, n\}}$ eine endliche Familie. Dann gilt (vgl. oben) $(\bigcap_{i=1}^n U_i)^C = \bigcup_{i=1}^n U_i^C =: \hat{U}$. Ist nun $U_i = \emptyset$ für ein $i \in \{1, \dots, n\}$, so ist $U_i^C = X$ und somit $\hat{U} = X \in \mathcal{T}$. Ist das nicht der Fall, dann ist \hat{U} als Vereinigung endlicher Mengen endlich und somit in \mathcal{T} .

Sei nun $x \in X$ und U eine Umgebung von x . Nun existiert keine Umgebung irgendeines anderen Punktes $y \in X$ in U^C , da U^C endlich ist, also ist dieser topologischer Raum nicht HAUSDORFF'sch.

- 7 **Beweis.** Wir überprüfen die Topologieeigenschaften :

- (T1) Per Definition gilt $\emptyset \in \mathcal{T}$ und, da $X^C = \emptyset$ ist, auch $X \in \mathcal{T}$.
 (T2) Seien $(U_i)_{i \in I} \subset \mathcal{T}$ eine Familie offener Mengen und $U := \bigcup_{i \in I} U_i$. Mit den **DEMORGAN'schen Regeln** zeigt man wie bei der kofiniten Topologie, dass $U \in \mathcal{T}$ ist, da Schmit abzählbarer Mengen abzählbar sind.

Da die endliche Vereinigung abzählbarer Mengen abzählbar ist, folgt auch die zweite Behauptung.

Das (X, \mathcal{T}) nicht Hausdorff ist, zeigt man genau wie in der vorangegangen Aufgabe. \square

- 8 **Beweis.** Die erste Eigenschaft (T1) ist trivialerweise erfüllt.

Seien zuerst $(M_i)_{i \in \{1, \dots, n\}} \subset \mathcal{T}$ eine endliche Familie offener Mengen und $M := \bigcap_{i=1}^n M_i$. Ist $M_i \neq \emptyset$ (dann gilt $M = \emptyset \in \mathcal{T}$), gilt $\bigcap_{i=1}^n M_i = U_\varphi \in \mathcal{T}$, wobei φ der kleinste Index ist, sodass $M_i = U_\varphi$ gilt.

Seien nun $(M_i)_{i \in I} \subset \mathcal{T}$ eine Familie offener Mengen und $M := \bigcup_{i \in I} M_i$ und $U := \bigcup_{i \in I} U_i$. Ist $M_i = X$ für ein $i \in I$, so gilt $M = X \in \mathcal{T}$. Ist I eine endliche Indexmenge, so gilt $M = U_{\max}$, wobei max der größte Vorkommende Index ist. Ist I nicht endlich, so gilt $M = U$ aufgrund der aufsteigenden Struktur der U_k .

Der topologische Raum $((0, 1), \mathcal{T}_{(0,1)})$ ist nicht HAUSDORFF: Wähle $x \in (0, 1)$, sodass $x < \frac{n}{n+1}$ für ein $n \in \mathbb{N}$ ist. Aus Symmetriegründen genügt es, $y > x$ zu wählen. Dann aber enthält jede Umgebung von y auch x . \square

- 9 Als Beispiel kann man die kofinite oder -abzählbare Topologie auf einer unendlichen Mengen X betrachten, da alle nicht HAUSDORFF'schen Räume nicht metrisierbar sind. **TODO: WARUM?**

Genau ist die indiskrete (oder auch triviale) Topologie $\mathcal{T} = \{\emptyset, X\}$ nicht HAUSDORFF'sch (nicht mal **T₀**) und somit nicht metrisierbar, weil (für $|X| > 1$) die „Distanz“ zweier Punkte immer Null ist, man kann also nur eine **Halb- bzw. Pseudometrik** auf X definieren.

1.3 Konvergenz

- Wir definieren Konvergenz in metrischen Räumen und konkretisieren das an sehr verschiedenen Beispielen und untersuchen den Zusammenhang zu Offenheit und Abgeschlossenheit von Mengen.
- Als wichtige Sätze lernen wir das Schachtelungsprinzip und den BANACH'schen Fixpunktsatz kennen.

Sei von nun an immer (X, d) ein metrischer Raum.

DEFINITION 1.3.1 (KONVERGENTE FOLGE, GRENZWERT, DIVERGENT)

Sei $(x_k)_{k \in \mathbb{N}}$ eine Folge in (X, d) . Die Folge heißt **konvergent gegen $a \in X$** , und wir schreiben $\lim_{k \rightarrow \infty} x_k = a$ oder $x_k \rightarrow a$, wenn $\lim_{k \rightarrow \infty} d(x_k, a) = 0$. Dann heißt a der **Limes** oder **Grenzwert** der Folge.

Die Folge heißt **konvergent**, wenn es ein $a \in X$ gibt, so dass (x_k) gegen a konvergiert. Eine nicht konvergente Folge heißt **divergent**.

Bemerkung Die rein topologische Charakterisierung von $\lim_{k \rightarrow \infty} x_k = a$ ist, dass in jeder Umgebung von a fast alle Glieder der Folge (x_k) liegen oder, äquivalent (metrisch), wenn in jedem $U_\varepsilon(a)$ mit $\varepsilon > 0$ fast alle Glieder der Folge (x_k) liegen.

Für \mathbb{R} mit der Standardmetrik ist diese Konvergenz also die übliche aus Analysis I.

Lemma 1.3.2 (Äquivalente Metriken)

Sind auf einer Menge X zwei Metriken d und \tilde{d} mit der Eigenschaft

$$C \cdot d(x, y) \leq \tilde{d}(x, y) \leq \frac{1}{C} \cdot d(x, y) \quad \forall x, y \in X$$

und eine Konstante $C > 0$ gegeben, so fallen die Begriffe Offenheit, Konvergenz (, Stetigkeit) und Umgebung in (X, d) und (X, \tilde{d}) zusammen.

Beweis. Klar. □

SATZ 1.3.1: KOORDINATENWEISE KONVERGENZ

Wir betrachten eine Folge $(x_k)_{k \in \mathbb{N}}$ im \mathbb{R}^n mit der Standard- oder d_p -Metrik für $p \in [1, \infty]$ und schreiben $x_k = (x_{k1}, \dots, x_{kn})$. Weiter sei $a = (a_1, \dots, a_n) \in \mathbb{R}^n$. Dann gilt

$$\lim_{k \rightarrow \infty} x_k = a \iff \lim_{k \rightarrow \infty} x_{kj} = a_j \text{ für alle } j.$$

Beweis. Wir wollen das vorangegangene Lemma anwenden und müssen also nur zeigen, dass sich die Metriken mit Konstanten gegeneinander abschätzen lassen.

Fall 1: $1 \leq p < \infty$. Für alle $k \in \mathbb{N}$ und $j \in \{1, \dots, n\}$ gilt

$$|x_{kj} - a_j| \leq \underbrace{\left(\sum_{i=1}^n |x_{ki} - a_i|^p \right)^{1/p}}_{d(x_k, a)} \leq n^{1/p} \sup_{1 \leq i \leq n} |x_{ki} - a_i|.$$

Fall 2: $p = \infty$. Dann gilt

$$\begin{aligned} d^\infty(x, y) &= \left(\max_{i=1}^n |x_i - y_i|^p \right)^{\frac{1}{p}} \leq d^p(x, y) \\ &= \left(\sum_{i=1}^n \underbrace{|x_i - y_i|}_{{\leq \max_{j=1}^n |x_j - y_j|}}^p \right)^{1/p} \leq n^{1/p} d^\infty(x, y). \quad (2) \end{aligned}$$

□

SATZ 1.3.2: FOLGEN-ABGESCHLOSSENHEIT

Sei A eine Teilmenge des metrischen Raumes (X, d) . Dann sind die beiden folgenden Aussagen äquivalent:

- ① A ist abgeschlossen.
- ② Für jede konvergente Folge $(x_k)_{k \in \mathbb{N}} \subset A$ gilt $\lim x_k \in A$.

Beweis. „ \Rightarrow “. Sei A abgeschlossen. Sei (x_k) eine Folge in A und $\lim x_k =: x \in X$. Wäre $x \in A^C$, so läge x also in der offenen Menge A^C , und diese wäre eine Umgebung von x . Dann lägen fast alle x_k in A^C . Es liegt aber kein x_k in dieser Menge.

„ \Leftarrow “: Der Grenzwert jeder konvergenten Folge in A liege wieder in A . Wir zeigen, dass A^C offen ist. Andernfalls existiert nämlich einen Punkt $x \in A^C$, so dass kein $U_\varepsilon(x)$ ganz in A^C liegt, bzw., sodass jede Umgebung von x nicht ganz in A^C liegt. Dann existiert zu jedem $k \in \mathbb{N}$ ein $x_k \in A$ mit $x_k \in U_{\frac{1}{k+1}}(x)$. Daher konvergieren die x_k gegen $x \in X \setminus A$. Widerspruch! □

Bemerkung 1.3.3 Die Rückrichtung gilt in allen erstabzählbaren topologischen Räumen, die Hinrichtung immer.

KOROLLAR 1.3.4 (PRODUKTMENGEN)

Seien $A_1, \dots, A_n \subset \mathbb{R}$ abgeschlossen. Dann ist auch $\bigtimes_{k=1}^n A_k \subset \mathbb{R}^n$ abgeschlossen.

Beweis. Das folgt unmittelbar aus Satz 1.3.2 in Verbindung mit Satz 1.3.1. □

DEFINITION 1.3.5 (GLEICHMÄSSIGE KONVERGENZ)

Seien $X \neq \emptyset$ eine Menge und (Y, d) metrischen Raum. Eine Folge $(f_k)_{k \in \mathbb{N}}$ von gleichmäßig konvergenten Abbildungen $f_k : X \rightarrow Y$ heißt gleichmäßig konvergent gegen $f : X \rightarrow Y$, wenn $\lim_{k \rightarrow \infty} d^{\text{sup}}(f_k, f) = 0$ gilt, d.h.

$$\begin{aligned} \forall \varepsilon > 0 \ \exists k_0 \in \mathbb{N} \ \forall k \geq k_0 \ \underbrace{\forall x \in X \ d(f_k(x), f(x)) \leq \varepsilon}_{\iff \sup_{x \in X} d(f_k(x), f(x)) \leq \varepsilon.} \\ \iff d^{\text{sup}}(f_k, f) \leq \varepsilon. \end{aligned}$$

Beispiel 1.3.6 (Ungleichmäßige Konvergenz)

Seien $f_k : [0, 1] \rightarrow [0, 1]$, $x \mapsto x^k$ und

$$f(x) := \begin{cases} 1 & \text{für } x = 1, \\ 0 & \text{sonst.} \end{cases}$$

Dann konvergiert für jedes $x \in [0, 1]$ die Folge punktweise $(f_k(x))_{k \in \mathbb{N}}$ gegen $f(x)$.

Aber für jedes k gilt $\sup_{x \in [0,1]} |f_k(x) - f(x)| = 1$, und daher ist die Konvergenz nicht gleichmäßig.

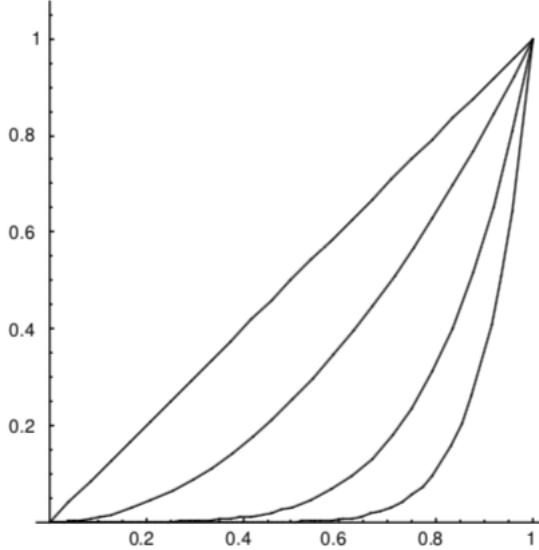


Abb. 13: Die Funktionen x, x^2, x^3, x^4 auf dem Intervall $[0, 1]$. [Quelle: Ferus]

Nun sehen wir, dass die Abgeschlossenheit von $\mathcal{C}_c([a, b], \mathbb{R})$ in $\mathcal{B}([a, b], \mathbb{R})$ zu dem folgenden Satz äquivalent ist:

SATZ 1.3.3: GLM. GRENZWERTE STETIGER FUNKTIONEN SIND STETIG

Sei $(f_k : [a, b] \rightarrow \mathbb{R})_{k \in \mathbb{N}} \subset \mathcal{C}_c([a, b], \mathbb{R})$ eine Folge stetiger Funktionen. Ist diese Folge gleichmäßig konvergent gegen $f : [a, b] \rightarrow \mathbb{R}$, so ist f stetig.

Beweis. Als stetige Funktionen auf einem kompakten Intervall sind die f_k beschränkt. Also liegt die Folge in $\mathcal{B}([a, b], \mathbb{R})$. Auch die Grenzfunktion f liegt in diesem Raum, also ist sie auch beschränkt. Die Folge (f_k) konvergiert also in $\mathcal{B}([a, b], \mathbb{R})$ gegen f , es gilt $\lim_{k \rightarrow \infty} d^{\sup}(f_k, f) = 0$. Nach Satz 1.2.1 ist $C^0([a, b])$ abgeschlossen, und nach Satz 1.3.2 liegt der Grenzwert f dann auch in $C^0([a, b])$. \square

Alternativer Beweis. (Meiner) Weil die Folge $(f_k)_{k \in \mathbb{N}}$ gleichmäßig gegen f konvergiert, gilt

$$\forall \varepsilon > 0 \exists k_0 \in \mathbb{N} : \sup_{x \in [a, b]} d(f_k(x), f(x)) < \frac{\varepsilon}{3} \quad \forall k \geq k_0.$$

Weil f auf einem kompakten Intervall definiert ist, ist f sogar gleichmäßig stetig: es gilt

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x, y \in [a, b] : d(x, y) < \delta \implies d(f(x), f(y)) < \frac{\varepsilon}{3}.$$

Seien nun $x, y \in [a, b]$ mit $d(x, y) < \delta$ für ein $\delta > 0$ und $k \geq k_0$. Dann gilt

$$\begin{aligned} d(f(x), f(y)) &\leq d(f(x), f_k(x)) + d(f_k(x), f_k(y)) + d(f_k(y), f(y)) \\ &\leq 3 \sup_{x \in [a, b]} d(f_k(x), f(x)) < \varepsilon. \end{aligned}$$

\square

Beispiel 1.3.7 (Ein schönes Gegenbeispiel von stackexchange.com) Sei $(f_n : \mathbb{R} \rightarrow \mathbb{R})_{n \in \mathbb{N}}$ eine Familie gleichmäßig stetiger Funktionen, welche auf jedem Kompaktum $[a, b]$ mit $a < b$ gleichmäßig gegen f konvergiert. Dann muss f nicht gleichmäßig stetig sein: Betrachte die Funktionenfolge

$$g_k(x) := \begin{cases} -k, & x \leq k, \\ x, & x \in (-k, k), \\ k, & x \geq k \end{cases} \quad \text{und} \quad f_k(x) := g_k(x) \cdot x.$$

Die Funktionen f_n sind alle gleichmäßig stetig, und sie konvergieren gleichmäßig auf jedem Kompaktum in \mathbb{R} , aber die Grenzfunktion $f(x) := x^2$ ist nicht gleichmäßig stetig.

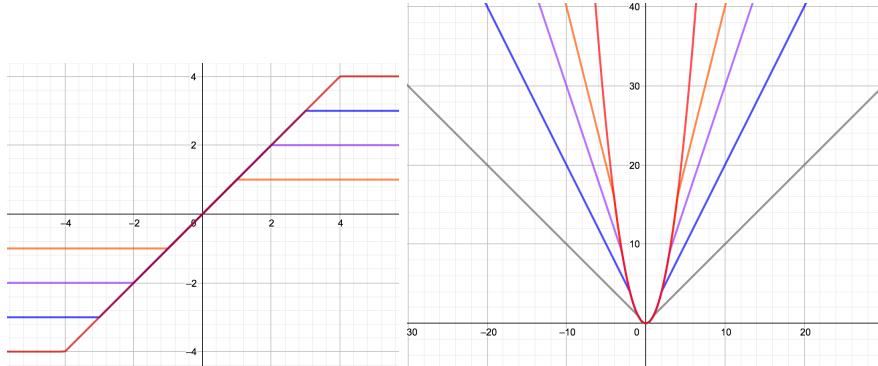


Abb. 14: Die Funktionen g_k und f_k für $k \in \{1, \dots, 4\}$ sowie die Grenzfunktion f . [Quelle: GeoGebra]

DEFINITION 1.3.8 (**d-CAUCHY-FOLGE**)

CAUCHY-FOLGE Eine Folge (x_k) in (X, d) heißt **d-CAUCHY-FOLGE**, wenn gilt

$$\forall \varepsilon > 0 \exists k_0 \in \mathbb{N} \forall k, l > k_0 : d(x_k, x_l) < \varepsilon.$$

DEFINITION 1.3.9 (**VOLLSTÄNDIGKEIT**)

vollständig Ein metrischer Raum (X, d) heißt **vollständig**, wenn jede CAUCHY-Folge in (X, d) konvergiert.

Korollar 1.3.10

Jede konvergente Folge ist eine CAUCHY-Folge.

Beweis. (Idee)

$$d(x_k, x_\ell) \stackrel{\triangle}{=} \underbrace{d(x_k, x)}_{\leq \frac{\varepsilon}{2}} + \underbrace{d(x, x_\ell)}_{\leq \frac{\varepsilon}{2}} < \varepsilon.$$

□

Korollar 1.3.11

Der metrische Raum (\mathbb{R}^n, d_p) ist für jedes $p \in [1, \infty]$ vollständig.

Beweis. Sei $(x_k)_{k \in \mathbb{N}} \subset \mathbb{R}^n$ eine CAUCHY-Folge, so ist wegen

$$|x_{ki} - x_{\ell i}| \leq \left(\sum_{j=1}^n |x_{kj} - x_{\ell j}|^p \right)^{1/p} = d^p(x_k, x_\ell)$$

für alle i auch $(x_{ki})_{k \in \mathbb{N}}$ eine CAUCHY-Folge, also konvergent. Aber koordinatenweise Konvergenz bedeutet Konvergenz im (\mathbb{R}^n, d^p) . □

Lemma 1.3.12 (Vollständigkeit)

Ist der metrische Raum (X, d) vollständig, so auch $(\mathcal{B}([a, b], X), d^{\sup})$.

Beweis. (Tut) Seien $I := [a, b]$ und $Y := (\mathcal{B}([a, b], X), d^{\sup})$ und $(f_n)_{n \in \mathbb{N}}$ eine CAUCHY-Folge in Y , dann existiert für alle $\varepsilon > 0$ ein $N_\varepsilon \in \mathbb{N}$, sodass für alle $n, m \geq N_\varepsilon$

$$\begin{aligned} d^{\sup}(f_n, f_m) \leq \varepsilon &\iff \sup_{x \in I} d(f_n(x), f_m(x)) \leq \varepsilon \\ &\iff d(f_n(x), f_m(x)) \leq \varepsilon \quad \forall x \in I. \end{aligned}$$

gilt. Also ist $(f_n(x))_{n \in \mathbb{N}}$ für jedes $x \in I$ eine CAUCHY-Folge in X .

Wir definieren $f : I \rightarrow X$, $x \mapsto \lim_{n \rightarrow \infty} f_n(x)$. Weil X vollständig ist, konvergiert $(f_n(x))_{n \in \mathbb{N}}$ für jedes $x \in I$. Somit ist f wohldefiniert, da in metrischen Räumen Grenzwerte eindeutig sind.

Wir wollen nun $f_n \rightarrow f \in Y$ zeigen. Dafür zeigen wir zunächst, dass $f \in Y$ ist. Da f_n eine CAUCHY-Folge ist, gilt insbesondere ein N_ε , sodass für alle $n \geq N_\varepsilon$

$$\begin{aligned} d^{\sup}(f_n, f_{N_\varepsilon}) \leq 1 &\iff d(f_n(x), f_{N_\varepsilon}(x)) \leq 1 \quad \forall x \in I \\ &\stackrel{n \rightarrow \infty}{\implies} d(f(x), f_{N_\varepsilon}(x)) \leq 1 \quad \forall x \in I \end{aligned} \tag{3}$$

gilt. Somit gilt für alle $x \in I$

$$\begin{aligned} d(f(x), f_{N_\varepsilon}(a)) &\leq d(f(x), f_{N_\varepsilon}(x)) + d(f_{N_\varepsilon}(x), f_{N_\varepsilon}(a)) \\ &\stackrel{(3)}{\leq} 1 + d(f_{N_\varepsilon}(x), f_{N_\varepsilon}(a)) \\ &\implies \sup_{x \in I} d(f(x), f_{N_\varepsilon}(a)) \leq 1 + \underbrace{\sup_{x \in I} d(f_{N_\varepsilon}(x), f_{N_\varepsilon}(a))}_{< \infty, \text{ weil } f_{N_\varepsilon} \text{ beschränkt ist}} < \infty. \end{aligned}$$

Somit gilt $\sup_{x \in I} f(x) < \infty$, also ist f beschränkt und somit $f \in Y$. Da (f_n) ein CAUCHY-Folge ist, gilt

$$\begin{aligned} \forall \varepsilon > 0 \ \exists N_\varepsilon \in \mathbb{N} \ \forall n, m \in \mathbb{N} \ \forall x \in I : d(f_m(x), f_n(x)) \leq \varepsilon \\ \stackrel{m \rightarrow \infty}{\implies} \forall x \in I : d(f(x), f_n(x)) \leq \varepsilon \\ \iff d^{\sup}(f, f_n) \leq \varepsilon \iff f_n \rightarrow f \end{aligned} \quad \square$$

Beispiel 1.3.13 TODO: Hier fehlt ein Gegenbeispiel für die Rückrichtung.

SATZ 1.3.4: VOLLSTÄNDIGKEIT UND ABGESCHLOSSENHEIT

Ist (X, d) vollständig und $A \subset X$ versehen mit der induzierten Metrik d_A , so ist (A, d_A) genau dann vollständig, wenn A abgeschlossen in X ist.

Beweis. Zu (\implies). Sei (A, d_A) vollständig und sei $(a_k)_{k \in \mathbb{N}}$ eine Folge in A , die gegen $x \in X$ konvergiert. Wir müssen zeigen, dass $x \in A$.

Als konvergente Folge ist (a_k) eine d -CAUCHY-Folge. Dann ist sie aber auch eine d_A -CAUCHY-Folge und nach der vorausgesetzten Vollständigkeit konvergent gegen ein $a \in A$. Das bedeutet $\lim d_A(a_k, a) = \lim d(a_k, a) = 0$. Also konvergiert (a_k) auch in X gegen a . Dann ist aber wegen der Eindeutigkeit des Grenzwertes $x = a \in A$.

Zu (\Leftarrow). Seien nun A abgeschlossen und (a_k) eine CAUCHY-Folge in (A, d_A) . Das ist dann auch eine CAUCHY-Folge in (X, d) , also konvergent gegen ein $x \in X$. Weil A abgeschlossen ist, liegt x in A und ist der Grenzwert von (a_k) in (A, d_A) . \square

SATZ 1.3.5: SCHACHTELUNGSPRINZIP

Seien (X, d) ein vollständiger metrischer Raum und $A_0 \supset A_1 \supset \dots$ eine "absteigende" bzw. "fallende" Folge von abgeschlossenen nichtleeren Mengen mit $\lim_{k \rightarrow \infty} \text{diam}(A_k) = 0$. Dann existiert genau ein $x^* \in X$, das in allen A_k liegt.

Beweis. Eindeutigkeit. (# ≤ 1) Sind $x_0, x_1 \in \bigcap_{k \in \mathbb{N}} A_k$, so gilt $x_0, x_1 \in A_k$ für alle $k \in \mathbb{N}$ und daher $d(x_0, x_1) \leq \text{diam } A_k$. Daher ist $d(x_0, x_1) = 0$, also $x_0 = x_1$.

Existenz. (# ≥ 1) Wähle aus jedem A_k ein a_k . Wir wollen zunächst zeigen, dass die Folge (a_k) eine CAUCHY-Folge ist. Sei also $\varepsilon > 0$. Da $\text{diam } A_k \rightarrow 0$ existiert ein $k_0 \in \mathbb{N}$ mit $\text{diam } A_k \leq \varepsilon$ für alle $k \geq k_0$. Sind $k, l \geq k_0$, so sind nach $A_0 \supset A_1 \supset \dots$ $a_k, a_l \in A_{k_0}$. Damit ist aber $d(a_k, a_l) \leq \varepsilon$ und (a_k) ist eine CAUCHY-Folge. Weil (X, d) vollständig ist, ist die Folge $(a_k)_{k \in \mathbb{N}}$ konvergent gegen ein $x^* \in X$, und es bleibt zu zeigen, dass $x^* \in A_k$ für alle k . Aber die Folge $(a_l)_{l \geq k}$ ist, wiederum wegen $A_0 \supset A_1 \supset \dots$, eine Folge in der abgeschlossenen Menge A_k , und deshalb liegt ihr Grenzwert in A_k . \square

Beispiel 1.3.14 (Gegenbeispiel bei fehlender Vollständigkeit.) Seien $X := \mathbb{Q}$, welcher nicht vollständig ist und $A_n := [q - \frac{1}{n}, q + \frac{1}{n}] \cap \mathbb{Q}$ für eine irrationale Zahl q und $n \in \mathbb{N}$. Dann ist $(\bigcap_{n \in \mathbb{N}} A_n) \cap \mathbb{Q} = \emptyset$.

Lemma 1.3.15 (CAUCHY-Folgenkriterium)

Seien $(x_k)_{k \in \mathbb{N}}$ eine Folge im metrischen Raum (X, d) und $a_k := d(x_k, x_{k+1})$. Gilt $\sum_{k=0}^{\infty} a_k < \infty$, ist a_k also konvergent, so ist $(x_k)_{k \in \mathbb{N}}$ eine CAUCHY-Folge.

Beweis. Nach der $\ell - 1$ -fachen Anwendung der Dreiecksungleichung ist

$$d(x_k, x_{k+\ell}) \leq \sum_{j=k}^{k+\ell-1} d(x_j, x_{j+1}) = \sum_{j=k}^{k+\ell-1} a_j \leq \sum_{j=k}^{\infty} a_j \xrightarrow{k \rightarrow \infty} 0. \quad \square$$

Beispiel 1.3.16 (Gegenbeispiel für Rückrichtung des CF-Kriteriums)

Die harmonische Folge $(x_n := \frac{1}{n})_{n \in \mathbb{N}}$ ist eine Cauchy-Folge, weil für o.B.d.A. $\ell \geq k > 1$ gilt: $d(x_\ell, x_k) < \frac{1}{\ell}$, jedoch gilt $\sum_{k=1}^{\infty} x_k = \infty$.

DEFINITION 1.3.17 (KONTRAKTION(SMODUL))

Eine Abbildung $f : (X, d_X) \rightarrow (Y, d_Y)$ zwischen metrischen Räumen heißt kontrahierend, wenn ein $\lambda \in (0, 1)$ existiert, so dass für alle $x_1, x_2 \in X$

$$d_Y(f(x_1), f(x_2)) \leq \lambda \cdot d_X(x_1, x_2).$$

gilt. Wichtig ist die echte Ungleichung $\lambda < 1$. Die Zahl λ heißt dann auch ein Kontraktionsmodul von f .

Bemerkung 1.3.18 (von mir) Die obige Abbildung f also genau dann kontrahierend, wenn $f(X) \subset X$ gilt und f LIPSCHITZ-stetig mit einer LIPSCHITZ-Konstanten $L < 1$ ist.

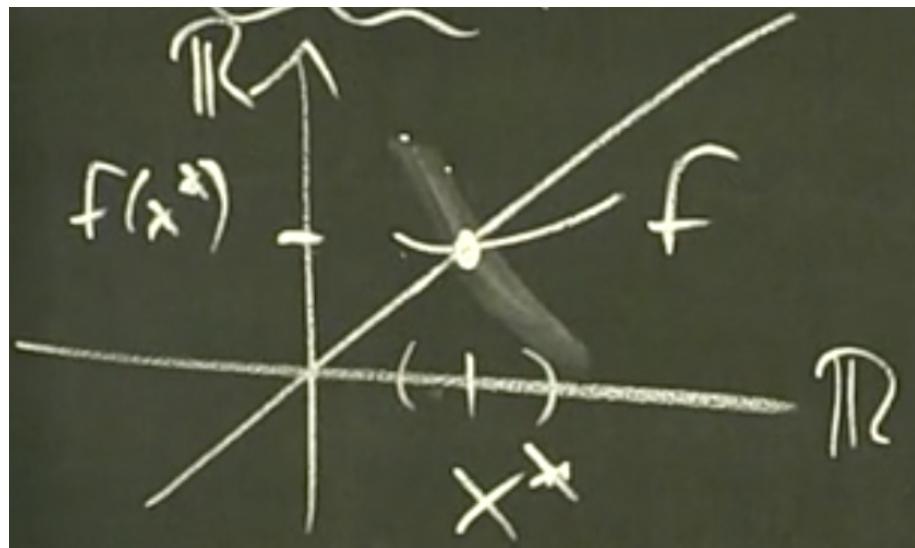
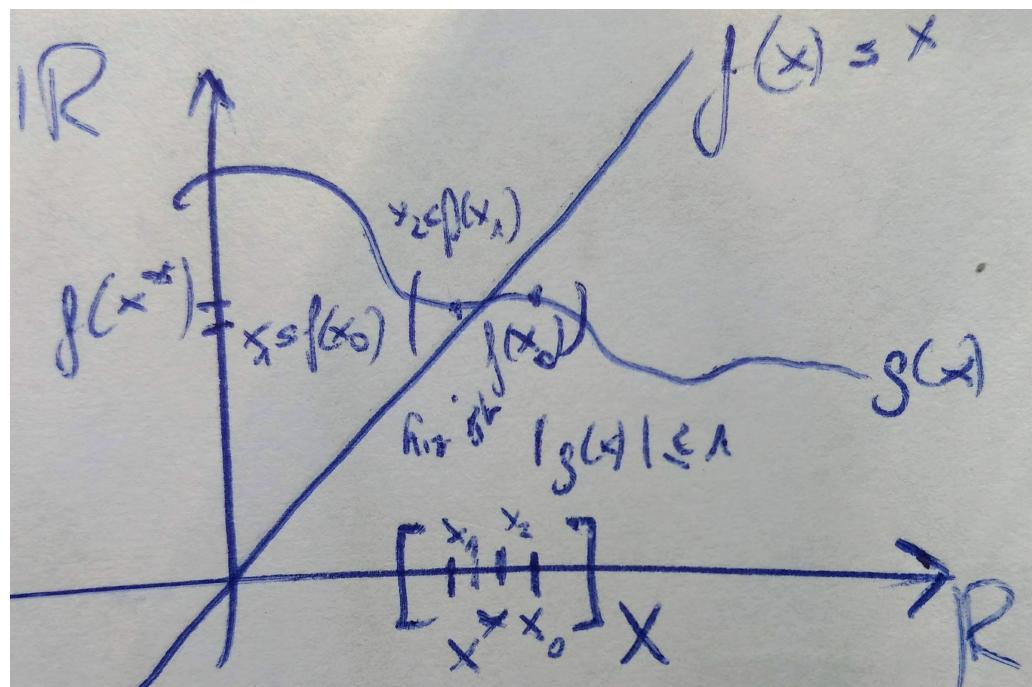


Abb. 15: Eine Abbildung $f : \mathbb{R} \rightarrow \mathbb{R}$, welche in einer Umgebung um einen Fixpunkt x^* kontrahierend ist. [Quelle: Vorlesung]



SATZ 1.3.6: BANACH'SCHER FIXPUNKTSATZ

Seien (X, d) ein nicht-leerer, vollständiger metrischer Raum und $f : X \rightarrow X$ eine kontrahierende Abbildung.

① Dann hat f genau einen Fixpunkt: Es gibt genau ein $x^* \in X$ mit $f(x^*) = x^*$.

② Ist $x_0 \in X$ und definiert man rekursiv $x_{k+1} = f(x_k)$ für alle $k \in \mathbb{N}$, so ist die Iterationsfolge $(x_k)_{k \in \mathbb{N}}$ konvergent gegen x^* .

Iterationsfolge

Beweis. Einzigkeit. Seien x_1^* und x_2^* zwei Fixpunkte von f . Dann gilt

$$0 < d(x_1^*, x_2^*) = d(f(x_1^*), f(x_2^*)) \leq \lambda d(x_1^*, x_2^*),$$

was einen Widerspruch darstellt.

Existenz. Seien $x_0 \in X$ und dazu $(x_k)_{k \in \mathbb{N}}$ definiert wie im Satz. Wir zeigen, dass $(x_k)_{k \in \mathbb{N}}$ eine CAUCHY-Folge ist. Zunächst ist

$$d(x_k, x_{k+1}) \leq \lambda d(x_{k-1}, x_k) \leq \lambda^2 d(x_{k-2}, x_{k-1}) \leq \dots \leq \lambda^k d(x_0, x_1)$$

Weil die Reihe $\sum_{k=0}^{\infty} \lambda^k d(x_0, x_1)$ als geometrische Reihe konvergiert, ist nach dem CAUCHY-Folgenkriterium (x_k) also eine CAUCHY-Folge und wegen der Vollständigkeit konvergent gegen ein $x^* \in X$.

Es bleibt zu zeigen, dass x^* ein Fixpunkt von f ist. Beachten Sie dazu, dass für alle k

$$\begin{aligned} d(f(x^*), x^*) &\leq d(f(x^*), x_{k+1}) + d(x_{k+1}, x^*) \leq d(f(x^*), f(x_k)) + d(x_{k+1}, x^*) \\ &\leq \lambda \underbrace{d(x^*, x_k)}_{\rightarrow 0} + \underbrace{d(x_{k+1}, x^*)}_{\rightarrow 0}. \end{aligned}$$

Es folgt $d(f(x^*), x^*) = 0$, also $f(x^*) = x^*$. \square

Alternativer Beweis. (für die Einzigkeit) Hat man gezeigt, dass $(x_k)_{k \in \mathbb{N}}$ eine CAUCHY-Folge in X ist, kann man auch so zeigen, dass x^* ein Fixpunkt ist. Weil f eine Kontraktion ist, ist sie insbesondere (LIPSCHITZ)-stetig (\star) und es gilt

$$x^* = \lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} f(x_{n-1}) \stackrel{(\star)}{=} f\left(\lim_{n \rightarrow \infty} x_{n-1}\right) = f(x^*). \quad \square$$

Korollar 1.3.19

Die Behauptung des Satzes gilt auch unter der schwächeren(!) Voraussetzung, dass nicht f , aber für ein $m \in \mathbb{N}$ die m -te Iteration $f^m := f \circ \dots \circ f$ kontrahierend ist.

Beweis. (selbst, von hier, (Thm. 2.4).) Nach dem obigen Satz hat die Abbildung f^m eine Fixpunkt $x^* \in X$. Nun gilt

$$f^m(f(x^*)) = f(f^m(x^*)) = f(x^*),$$

also ist $f(x^*)$ ein Fixpunkt von f^m . Weil der Fixpunkt von f^n nach dem obigen Satz eindeutig ist, muss $f(x^*) = x^*$ sein, also hat f genau einen Fixpunkt. \square



Eine tolle Anwendung des Fixpunktsatzes von BANACH ist der Satz von PICARD-LINDELÖF im Kontext von gewöhnlichen Differentialgleichungen:

DEFINITION 1.3.20 (GEWÖHNLICHE DIFFERENTIALGLEICHUNG, AWP)

Seien $I, J \subset \mathbb{R}$ Intervalle, $x_0 \in I$, $y_0 \in J$ sowie $f : I \times J \rightarrow \mathbb{R}$ eine stetige Abbildung. Dann nennt man für alle $x \in I$

$$y'(x) = f(x, y(x))$$

gewöhnliche eine gewöhnliche Differentialgleichung und wir suchen ein differenzierbare Differentialgleichung Lösungsfunktion $y : I \rightarrow J$, welche diese Gleichung erfüllt.

Man nennt dies ein **Anfangswertproblem**, wenn zusätzlich die Bedingung $y(x_0) = y_0$ gegeben ist. Die Kurzschreibweise dafür lautet $y' = f(x, y)$, $y(x_0) = y_0$.

Ein Kriterium für die Existenz und Eindeutigkeit einer Lösung y liefert der

SATZ 1.3.7: PICARD-LINDELÖF

Seien $I := [x_0 - a, x_0 + a]$, $J := [y_0 - b, y_0 + b]$ und f LIPSCHITZ-stetig im zweiten Argument, d.h. es existiert ein $L > 0$, sodass für alle $x \in I$ gilt $|f(x, y) - f(x, \tilde{y})| \leq L \cdot |\tilde{y} - y|$. Setze $M := \max_{x \in I, y \in J} |f(x, y)|$ und $\alpha := \min \left\{ a, \frac{b}{M} \right\}$.

Ist $\alpha < \frac{1}{L}$, so existiert eine eindeutige Lösung des Anfangswertproblems auf $\tilde{I} := [x_0 - \alpha, x_0 + \alpha]$.

Beweis. Wir wollen den Fixpunktsatz von BANACH auf die Fixpunktgleichung $y = \Phi(y)$ anwenden. Dabei sei $\Phi(\varphi)(x) := y_0 + \int_{x_0}^x f(t, \varphi(t))dt$, denn das ist wegen des Hauptsatzes der Integral- und Differentialgleichung äquivalent zu dem Anfangswertproblem und erfüllt die Bedingung $\varphi(x_0) = y_0$.

Betrachte $X := C(\tilde{I}, J)$ mit der Supremumsmetrik d^{\sup} . Das Paar (X, d^{\sup}) ist nach einer Verallgemeinerung von Lemma 1.3.12 ein vollständiger metrischer Raum als abgeschlossener Teilraum von $C(\tilde{I}, \mathbb{R})$. Die Abbildung $\Phi : X \rightarrow X$ ist wohldefiniert, denn für alle $\varphi \in X$ ist $\Phi(\varphi) : \tilde{I} \rightarrow \mathbb{R}$ differenzierbar und somit insbesondere stetig und es gilt

$$|\Phi(\varphi)(x) - y_0| = \left| \int_{x_0}^x f(t, \varphi(t))dt \right| \leq \int_{x_0}^x |f(t, \varphi(t))|dt \leq \alpha \cdot M \leq b,$$

also ist $\Phi(\varphi)(x) \in J$ für alle $x \in \tilde{I}$ und somit $\Phi(\varphi) \in X$.

Zuletzt zeigen wir, dass Φ kontrahierend ist.

Seien $\varphi, \psi \in X$, dann gilt

$$\begin{aligned} d^{\sup}(\Phi(\varphi), \Phi(\psi)) &= \sup_{x \in \tilde{I}} |\Phi(\varphi)(x) - \Phi(\psi)(x)| = \sup_{x \in \tilde{I}} \left| \int_{x_0}^x f(t, \varphi(t)) - f(t, \psi(t)) dt \right| \\ &\leq \sup_{x \in \tilde{I}} \int_{x_0}^x |f(t, \varphi(t)) - f(t, \psi(t))| dt \leq \sup_{x \in \tilde{I}} \int_{x_0}^x L |\varphi(t) - \psi(t)| dt \\ &\leq L \cdot \sup_{x \in \tilde{I}} \int_{x_0}^x d^{\sup}(\varphi, \psi) dt \leq \underbrace{L \cdot \alpha}_{<1} d^{\sup}(\varphi, \psi). \end{aligned} \quad \square$$

AUFGABEN

LÖSUNGEN

1.4 Kompaktheit und Vollständigkeit

- Mit der Kompaktheit lernen wir einen zentralen Begriff der Topologie, der Analysis und der Geometrie kennen.
- Die Definition mittels offener Überdeckungen ist logisch nicht ganz einfach, dafür aber an die vielfältigen Verwendungen der Kompaktheit angepasst.
- Im \mathbb{R}^n existiert eine einfache Charakterisierung kompakter Mengen (Satz von HEINE-BOREL), die aber in unendlich-dimensionalen (und vielen anderen) Räumen nicht greift.

Seien nun stets wenn nicht anders behauptet (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge.

DEFINITION 1.4.1 (OFFENE ÜBERDECKUNG)

offene Überdeckung Eine offene Überdeckung von A ist eine Familie $(U_i)_{i \in I}$ von offenen Teilmengen $U_i \subset X$, so dass $A \subset \bigcup_{i \in I} U_i$ gilt.

DEFINITION 1.4.2 ((ÜBERDECKUNGS)KOMPAKTHEIT)

kompakt Die Menge A heißt kompakt, wenn jede offene Überdeckung von A eine endliche Überdeckung von A enthält, d.h. wenn gilt: Ist $(U_i)_{i \in I}$ eine offene Überdeckung von A , so existiert eine endliche Teilmenge $K \subset I$, so dass $A \subset \bigcup_{i \in K} U_i$. Man

Teilüberdeckung nennt $(U_i)_{i \in K}$ dann auch eine endliche Teilüberdeckung von A .

Korollar 1.4.3

Sei $(a_n)_{n \in \mathbb{N}} \subset X$ eine gegen $a \in X$ konvergente Folge. Dann ist $A := \{a\} \cup \{a_n : n \in \mathbb{N}\}$ kompakt.

Beweis. Existiert eine offene Überdeckung $(U_i)_{i \in I}$ von A , so liegt a in einer der offenen Mengen. Wegen der Konvergenz der Folge liegen in dieser Menge dann aber fast alle a_k , und man braucht nur noch endlich viele weitere U_i , um den Rest "einzufangen". \square

Die Menge $A := \{a_n : n \in \mathbb{N}\}$ ist natürlich nicht kompakt, man nehme als Beispiel $a_k := \frac{1}{k}$, dann gilt $A = (0, 1]$.

Lemma 1.4.4

Eine Teilmenge A eines metrischen Raumes (X, d) ist genau dann kompakt, wenn sie als (triviale) Teilmenge von (A, d_A) kompakt ist.

Beweis. (Große Übung) TODO \square

SATZ 1.4.1: PRODUKTE ABGESCHLOSSENER BESCHRÄNKTER REELLER INTERVALLE SIND KOMPAKT IN \mathbb{R}^n

Seien $I_1, \dots, I_n \subset \mathbb{R}$ abgeschlossene und beschränkte Intervalle mit $I_k = [a_k, b_k]$. Dann ist der abgeschlossene n -dimensionale Quader $Q := \times_{k=1}^n I_k$ kompakt in \mathbb{R}^n .

Beweis. Sei $(U_i)_{i \in I}$ eine offene Überdeckung von Q . Wir nehmen an, dass keine endliche Teilfamilie von $(U_i)_{i \in I}$ ganz Q überdeckt.

Wir zerlegen Q durch Halbieren aller Seiten in 2^n abgeschlossene Teilquader vom halben Durchmesser. Dann existiert wenigstens eines dieser Teilquader, wir nennen es Q_1 , welches nicht durch endlich viele der U_i zu überdecken ist.

Wir zerlegen Q_1 durch Halbieren aller Seiten in 2^n abgeschlossene Teilquader vom halben Durchmesser. Dann existiert wenigstens einen dieser Teilquader, wir nennen es Q_2 , welches nicht durch endlich viele der U_i zu überdecken ist.

Durch Fortsetzung dieses Verfahrens finden wir eine Folge von abgeschlossenen fallend ineinander enthaltenen nicht leeren Quadern $Q \supset Q_1 \supset Q_2 \supset \dots$ mit $0 < \text{diam}(Q_k) \leq C \cdot 2^{-k} \rightarrow 0$ für eine Konstante $C > 0$, die von der längsten Diagonale im ursprünglichen Quader abhängt, deren keines sich durch endlich viele der U_i überdecken lässt.

Nach dem Schachtelungsprinzip existiert nur ein $x \in \bigcap_{k=1}^n Q_k \subset Q$. Nach Voraussetzung existiert ein $i_0 \in I$ mit $x \in U_{i_0} \neq \emptyset$. Weil U_{i_0} offen ist, existiert ein $\varepsilon > 0$ mit $x \in U_\varepsilon(x) \subset U_{i_0}$. Dann liegt aber jeder Quader Q_k vom Durchmesser $< \varepsilon$ ganz in U_{i_0} , im Widerspruch zur Konstruktion von Q , da die verbleibenden endlichen Q_k mit endlichen vielen Indizes "eingefangen" werden können! \square

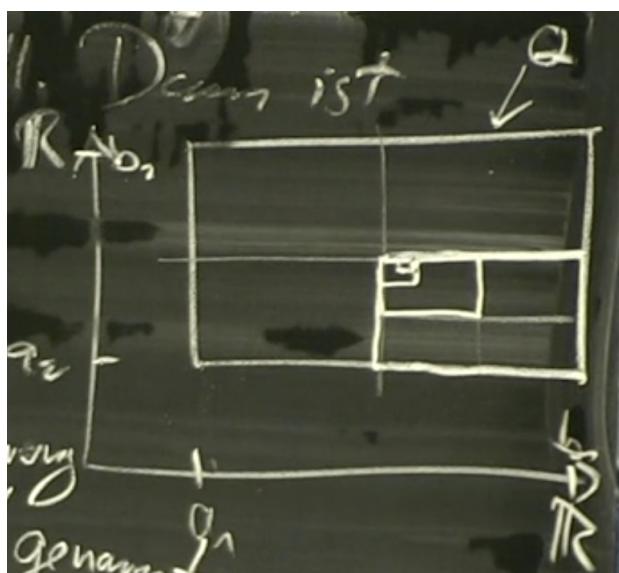


Abb. 16: Veranschaulichung der 2^n -Teilung des Quaders in $n = 2$. [Quelle: Vorlesung]

Für $n = 1$ liefert der Satz:

Korollar 1.4.5

Intervalle $[a, b]$ mit $-\infty < a \leq b < +\infty$ sind kompakt in \mathbb{R} .

SATZ 1.4.2: RÜCKRICHTUNG HEINE-BOREL

Eine kompakte Teilmenge eines metrischen Raumes ist abgeschlossen und beschränkt.

Beweis. Sei $A \subset X$ kompakt. Zum Beweis müssen wir geeignete offene Überdeckungen von A konstruieren und ausnutzen, dass sie endliche Teilüberdeckungen besitzen.

Zur Beschränktheit. Sei o.B.d.A $X \neq \emptyset$. Sei $x \in X$. Dann ist die Familie $(U_k)_{k \in \mathbb{N} \setminus \{0\}}$ eine offene Überdeckung von A und wegen der Kompaktheit von A existiert ein $N \in \mathbb{N} \setminus \{0\}$, sodass $A \subset \bigcup_{k=1}^N U_k$ gilt.

Alternative von mir: Sei $\varepsilon > 0$. Wir wählen als offene Überdeckung von A die Familie $U := (U_\varepsilon(a))_{a \in A} \subset X$. Weil A kompakt ist, existiert zu jeder und damit auch dieser offenen Überdeckung eine endliche Teilüberdeckung $(U_i)_{i \in K} \subset U$. Setze nun $D := \sum_{k \in K} \underbrace{\text{diam}(U_k)}_{<\infty}$. Dann gilt $\text{diam}(A) \leq D < \infty$.

Zur Abgeschlossenheit. Sei $x \in X \setminus A$. Zu jedem $a \in A$ sei U_a eine offene Kugel um a mit Radius $\frac{1}{2} d(a, x)$, dann gilt $x \notin U_a$. Offenbar bildet $(U_a)_{a \in A}$ eine offene Überdeckung von A , und nach Voraussetzung existiert also ein $n \in \mathbb{N}$ und a_0, \dots, a_n , so dass $A \subset \bigcup_{k=0}^n U_{a_k}$ gilt.

Sei $\varepsilon := \frac{1}{2} \min_{k=0}^n \{d(x, a_k)\}$. Dann gilt $\varepsilon > 0$, weil $x \notin A$ liegt, und $U_\varepsilon(x) \cap U_{a_k} = \emptyset$ für alle $k \in \{0, \dots, n\}$. Daher ist $U_\varepsilon(x) \subset X \setminus A$, also $X \setminus A$ offen und A abgeschlossen. \square

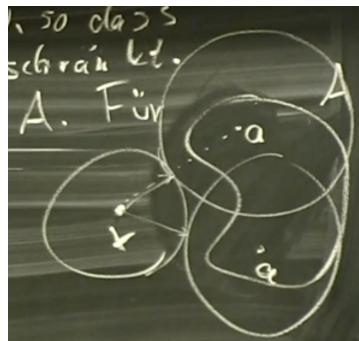


Abb. 17: Visualisierung für die Argumentation der Abgeschlossenheit

SATZ 1.4.3

Jede abgeschlossene Teilmenge einer kompakten Teilmenge ist kompakt.

Beweis. Seien (X, d) ein metrischer Raum, $A \subset X$ eine kompakte Teilmenge und $B \subset A$ eine abgeschlossene Teilmenge von A . Sei $(U_i)_{i \in I}$ eine offene Überdeckung von B . Wir suchen eine endliche Teilüberdeckung.

Durch Hinzunahme der offenen Menge $U := X \setminus B$ erhält man eine offene Überdeckung von A . Weil A kompakt ist, existiert eine endliche Teilmenge $K \subset I$ mit $A \subset (U \cup \bigcup_{k \in K} U_k)$. Weil B und U disjunkt sind, gilt $B \subset \bigcup_{k \in K} U_k$, und wir haben für B eine endliche Teilüberdeckung von $(U_i)_{i \in I}$ gefunden. \square

SATZ 1.4.4: HEINE-BOREL

Eine Teilmenge $A \subset \mathbb{R}^n$ mit der d^p -Metrik ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.

Beweis. " \implies gilt nach Satz 1.4.2.

" \Leftarrow ": Klar nach den Sätzen 1.4.1 und 1.4.3, da man für jede abgeschlossenen und beschränkte Menge A einen kompakten Quader finden kann, der A enthält. Also ist A als abgeschlossenen Teilmenge eines Kompaktaums kompakt. \square

Dieser Satz ist falsch in allgemeinen metrischen Räumen:

Beispiel 1.4.6 (abgeschlossen & beschränkt impliziert nicht kompakt)

Sei M eine unendliche Menge und $(X, d) = (\mathcal{B}(M, \mathbb{R}), d^{\text{sup}})$. Für $m \in M$ sei $f_m \in \mathcal{B}(M, \mathbb{R})$ definiert durch

$$f_m(n) := \begin{cases} 1 & \text{für } m = n, \\ 0 & \text{sonst.} \end{cases} \implies d^{\text{sup}}(f_m, f_n) = \begin{cases} 0 & \text{für } m = n, \\ 1 & \text{sonst.} \end{cases} \quad (4)$$

Die Menge $A := \{f_m \mid m \in M\}$ hat daher den Durchmesser $\text{diam}(M) = 1$ und ist beschränkt.

Betrachtet man A als Teilmenge von $\mathcal{B}(M, \mathbb{R})$ mit der durch d^{sup} induzierten Metrik, so ist A trivialerweise auch abgeschlossen, da jede konvergente Folge in A ab einem bestimmten Index konstant sind.

Nach (4) ist $U_{\frac{1}{2}}(f_m) \cap A = \{f_m\}$, und die offene Überdeckung $\left(U_{\frac{1}{2}}(f_m)\right)_{m \in M}$ von A besitzt deshalb keine endliche Teiliüberdeckung. Also ist A als Teilmenge von (A, d_A^{sup}) nicht kompakt, wohl aber abgeschlossen und beschränkt.

Derselbe Beweis klappt für jede unendliche Menge mit der diskreten Metrik.

Nach Satz 1.4.4 ist A auch als Teilmenge von $(\mathcal{B}(M, \mathbb{R}), d^{\text{sup}})$ nicht kompakt. Es ist aber beschränkt, s. oben, und auch abgeschlossen, wie wir mittels Satz 1.3.2 noch zeigen wollen. Eine konvergente Folge in A ist eine CAUCHY-Folge. Nach (4) und dem Cauchykriterium mit $\varepsilon < 1$ sind dann aber fast alle Folgenglieder gleich, also gleich dem Limes, der damit ebenfalls in A liegt.

DEFINITION 1.4.7 (FOLGENKOMPAKTHEIT)

Sei (X, d) ein metrischer Raum. Dann heißt $A \subset X$ folgenkompakt, falls jede Folge in A eine konvergente Teilfolge besitzt. Weil A dann auch abgeschlossen ist, liegt der Limes dann in A .

SATZ 1.4.5: BOLZANO-WEIERSTRASS

Seien (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge. Dann gilt

$$A \text{ überdeckungskompakt} \iff A \text{ folgenkompakt}$$

Beweis. " \implies ": Seien A überdeckungskompakt und $(x_k)_{k \in \mathbb{N}}$ eine Folge in A .

Fall 1: Es existiert ein $a \in A$, sodass jede Umgebung von a unendlich viele Folgenglieder x_k enthält.

Wir konstruieren induktiv (mit $n_0 \in \mathbb{N}$ beliebig) $n_0 < n_1 < n_2 < \dots$ in \mathbb{N} , sodass $U_{\frac{1}{k}}(a)$ für alle $k \in \mathbb{N}$ den Punkt x_{n_k} enthält.

Sei $\varepsilon > 0$. Wir wählen $k \geq \frac{1}{\varepsilon}$, dann gilt $U_{\frac{1}{k}}(a) \subset U_\varepsilon(a)$. Weil die $U_{\frac{1}{k}}(a)$ fallend ineinander enthalten sind, liegen alle x_{n_j} mit $j \geq k$ (unendlich viele) in $U_{\frac{1}{k}}(a)$ also auch in $U_\varepsilon(a)$. Somit gilt $\lim_{k \rightarrow \infty} x_{n_k} = a$.

Fall 2: Für alle $a \in X$ existiert eine offene Umgebung U_a von a , die nur endlich viele x_k enthält (Negation von Fall 1). Dann ist $(U_a)_{a \in A}$ eine offene Überdeckung von A . Weil A kompakt ist, wird A auch schon von endlich vielen der U_a überdeckt, d.h.

$$\exists n \in \mathbb{N} \ \exists a_0, \dots, a_n \in A : A \subset U_{a_0} \cup \dots \cup U_{a_n}.$$

Also liegen nur endlich viele der x_n in A , was einen Widerspruch dazu darstellt, dass x_n in komplett in A liegt und unendliche viele Folgenglieder besitzt.

” \Leftarrow “: Etwas trickreich (vgl. Klaus Jänich, Topologie, Springer Hochschultext, 2. Aufl. p.97). \square

Korollar 1.4.8

Jeder kompakte metrische Raum ist vollständig.

Beweis. Ist (X, d) kompakt und $(x_k)_k$ eine CAUCHY-Folge in X , so hat diese nach dem Satz von BOLZANO-WEIERSTRASS eine gegen x^* konvergente Teilfolge $(x_{k_j})_{j \in \mathbb{N}}$. Weil X kompakt und somit abgeschlossen ist, liegt $x^* \in X$. Nun zeigen wir, dass dann die gesamte Folge $(x_k)_k$ gegen x^* konvergiert.

Sei $\varepsilon > 0$. Wir wählen k so groß, dass $d(x_{n_k}, x^*) \leq \frac{\varepsilon}{2}$ gilt. Dann wähle n so groß, dass $d(x_n, x_{n_k}) \leq \frac{\varepsilon}{2}$ gilt, was geht, da x_n eine CAUCHY-Folge ist. Also gilt

$$d(x_n, x^*) \leq (x_n, x_{n_k}) + d(x_{n_k}, x^*) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \quad \square$$

Korollar 1.4.9 (Teilfolgen von CAUCHY-Folgen, von mir)

Konvergiert eine Teilfolge einer CAUCHY-Folge, so auch die gesamte CAUCHY-Folge.

AUFGABEN

- (1) [HW07 von Math104 von des USC Berkeley] Sei (X, d) ein metrischen Raum.

(i) Sei $(S_i)_{i \in I} \subset X$ eine Familie von abgeschlossenen Teilmengen. Zeigen Sie, dass auch ihr Durchschnitt $S := \bigcap_{i \in I} S_i$ kompakt in X ist.

(ii) Sei $(S_i)_{i \in \{1, \dots, n\}} \subset X$ eine endliche Familie kompakter Teilmengen. Zeigen Sie, dass deren Vereinigung $S := \bigcup_{i=1}^n S_i$ kompakt in X ist.

- (2) (**Idee von hier**). [Gegenbeispiel für den Satz von HEINE-BOREL]

Sei $\mathcal{K} := \mathbb{R}$ (hier funktioniert auch jeder unbeschränkter metrischer Raum) und d die induzierte Metrik. Definiere nun eine neue Metrik \tilde{d} auf \mathcal{K} auf die folgende Weise für $x, y \in \mathcal{K}$:

$$\tilde{d}(x, y) = \begin{cases} d(x, y), & \text{wenn } d(x, y) < 1, \\ 1, & \text{sonst.} \end{cases}$$

- (i) Verifizieren Sie, dass \tilde{d} ein Metrik auf \mathcal{K} definiert.
(ii) Zeigen Sie, dass die von d induzierte Topologie mit der von \tilde{d} induzierten Topologie übereinstimmt und folgern Sie, dass der metrische Raum (\mathcal{K}, \tilde{d}) nicht kompakt ist.
(iii) Wie kann man das als Gegenbeispiel für den Satz von HEINE-BOREL nutzen?

- (3) (**Idee von hier**). Zeigen Sie, dass der metrische Raum (\mathbb{N}, d) mit der diskreten Metrik d beschränkt aber nicht kompakt ist und illustrieren Sie, warum man so ein Gegenbeispiel für den Satz von HEINE-BOREL konstruieren kann.

Tipp: Wie sehen die offenen Mengen in (\mathbb{N}, d) aus?

- (4) [**von mir**] Zeigen Sie, dass jede Menge mit der diskreten Metrik einen vollständigen metrischen Raum bildet.

- (5) (**vgl. hier**) Zeigen Sie, dass die von der Dschungelflussmetrik induzierte Topologie zusammen mit dem \mathbb{R}^2 einen nicht kompakten aber vollständigen topologischen Raum bildet. *Tipp: Nutzen Sie, dass \mathbb{R} mit der Standardmetrik vollständig ist. Ein topologischer Raum (X, \mathcal{T}) heißt kompakt, wenn X kompakt ist. Also ist eine Teilmenge eines topologischen Raumes genau dann kompakt, wenn sie als topologischer Raum mit der Teilraumtopologie kompakt ist.*

- (6) Geben Sie ein Gegenbeispiel für den Satz von DINI, wenn auf die Voraussetzung verzichtet wird, dass die Grenzfunktion stetig ist.



- (7) Seien (X, \mathcal{T}) ein Hd-Raum und $(K_n)_{n \in \mathbb{N}}$ eine Familie von kompakten Teilmengen. Zeigen Sie, dass

(i) $\bigcap_{n \in \mathbb{N}} K_n$ kompakt ist.

(ii) Für ein $k \in \mathbb{N}$ die Menge $\bigcup_{n=1}^k K_n$ kompakt ist.

(iii) Geben Sie ein Beispiel dafür an, dass im Allgemeinen $\bigcup_{n \in \mathbb{N}} K_n$ nicht

kompakt ist.

- ⑧ Sei (X, \mathcal{T}) ein Hausdorff'scher topologischer Raum. Wir sagen X besitzt die endliche Durchschnittseigenschaft, wenn für jede Familie $(A_i)_{i \in I}$ von abgeschlossenen Mengen mit $\bigcap_{i \in I} A_i = \emptyset$ es bereits endlich viele $i_1, \dots, i_k \in I$ geben muss, so dass $\bigcap_{i=i_1}^{i_k} A_i = \emptyset$ gilt. Zeigen Sie, dass
- ① X genau dann kompakt ist, wenn X die endliche Durchschnittseigenschaft besitzt.
 - ② falls alle A_i mit $i \in I$ kompakt sind, aus $\bigcap_{i \in I} A_i = \emptyset$ bereits $\bigcap_{i=i_1}^{i_k} A_i = \emptyset$ für geeignete $i_1, \dots, i_k \in I$ folgend muss.
- ⑨ Zeigen Sie, dass die Funktion $f : [0, \infty) \rightarrow \mathbb{R}$, $x \mapsto \frac{e^{-x}}{3} - \sqrt{x}$ eine Nullstelle besitzt, und leiten Sie aus dem Banachschen Fixpunktsatz ein Verfahren ab, die Nullstelle bis auf eine Genauigkeit von 10^{-2} zu bestimmen.

LÖSUNGEN

- (1) (i) Da kompakte Mengen nach Satz 1.4.2 insbesondere beschränkt sind, sind alle S_i und damit auch nach Korollar 1.2.13 S abgeschlossen. Für alle $i \in I$ gilt $S \subset S_i$ und somit ist S nach Satz 1.4.3 als abgeschlossene Teilmenge einer kompakten Menge abgeschlossen.
- (ii) Seien $(U_i)_{i \in I}$ eine offene Überdeckung von S und $U := \bigcup_{i \in I} U_i$. Für jedes $j \in \{1, \dots, n\}$ gilt $S_j \subset S \subset U$ also ist U auch eine offene Überdeckung von S_j . Weil S_j kompakt ist, existiert eine endliche Teilmenge I_j , sodass $S_j = \bigcup_{i \in I_j} U_i$ gilt. Nun ist auch $I_0 := \bigcup_{k=1}^n I_k$ als endliche Vereinigung endlicher Mengen endlich und es gilt $S \subset \bigcup_{i \in I_0} U_i$, also existiert eine endliche Teilüberdeckung und S ist kompakt.
- (2) (i) (a) Weil d eine Metrik ist, ist \tilde{d} stets positiv. Gilt $x = y$, so gilt $d(x, y) = 0 < 1$ und somit $\tilde{d}(x, y) = 0 \neq 1$. Gilt $\tilde{d}(x, y) = 0 \neq 1$, so gilt $\tilde{d}(x, y) = d(x, y) = 0$ und somit $x = y$.
- (b) Diese Metrik ist offensichtlich symmetrisch, weil es d ist.
- (c) Seien $x, y, z \in \mathcal{K}$.

Fall 1: $d(x, z) \geq 1$. Dann gilt $\tilde{d}(x, z) = 1$ und somit

$$\tilde{d}(x, y) \leq 1 \leq \tilde{d}(x, z) + \tilde{d}(y, z).$$

Fall 2: $d(y, z) \geq 1$ Ist aus Symmetriegründen analog zu Fall 1.

Fall 3: $d(x, z) < 1$ und $d(y, z) < 1$ Dann gilt

$$\tilde{d}(x, y) \leq d(x, y) \leq d(x, z) + d(y, z) = \tilde{d}(x, z) + \tilde{d}(y, z).$$

- (ii) **Beweis.** Wir müssen zeigen, dass die Mengen, welche bezüglich d offen sind, genau die sind, die bezüglich \tilde{d} offen sind.
- (iii) Der metrische Raum (\mathcal{K}, \tilde{d}) ist nicht kompakt. Nun müssen noch zeigen, dass er jedoch abgeschlossen und beschränkt ist. Er ist beschränkt, weil $\tilde{d}(x, y) \leq 1$ für alle $x, y \in \mathcal{K}$ gilt.

- (3) **Beweis.** Weil wir die diskrete Metrik benutzen, sind alle offenen Mengen der Form

$$B(x, \varepsilon) := \begin{cases} \{x\}, & \text{wenn } \varepsilon < 1, \\ X, & \text{wenn } \varepsilon \geq 1 \end{cases}$$

Nun ist \mathbb{N} beschränkt, da für alle $n \in \mathbb{N}$ und alle $\varepsilon \geq 1$ $\mathbb{N} \subset B(n, \varepsilon)$ gilt.

\mathbb{N} ist auch bezüglich der diskreten Metrik abgeschlossen, weil bezüglich der diskreten Metrik alle Mengen offen und abgeschlossen sind.

Wir nehmen an, \mathbb{N} sei kompakt. Die Familie $(B(n, \frac{1}{2}))_{n \in \mathbb{N}}$ ist eine offene Überdeckung von \mathbb{N} , und nach Voraussetzung existiert eine endliche Teilüberdeckung $C := (B(n, \frac{1}{2}))_{n \in \{1, \dots, N\}}$, sodass $\mathbb{N} \subset C$ gilt. Da $\frac{1}{2} < 1$ ist, sind alle Mengen aus der Umgebung eelementige Mengen. Also ist \mathbb{N} endlich, was ein Widerspruch ist, also ist \mathbb{N} nicht kompakt. \square

Genauer gilt sogar: Eine Teilmenge eines metrischen Raumes ausgestattet mit der diskreten Metrik ist genau dann kompakt, wenn sie endlich ist.

(4)

- (5) Da wir schon zeigten, dass (\mathbb{R}^2, d) ein metrischer Raum ist, müssen wir wegen Satz 1.2.3 nur noch zeigen, dass $(\mathbb{R}^2, \mathcal{T})$ nicht kompakt aber vollständig ist.

(i) Hint here

- (ii) Sei $S = \{(x_n, y_n)\}_{n \in \mathbb{N}} \subset \mathbb{R}^2$ eine d -CAUCHY-Folge. Nun genügt es nach Korollar 1.4.9 zu zeigen, dass eine Teilfolge von S konvergiert.

[Weiter hier.](#)

- (6) Die Funktionenfolge $(f_n : [0, 1] \rightarrow X, x \mapsto 1 - x^n)_{n \in \mathbb{N}}$, wobei X ein kompakter topologischer Raum ist, konvergiert punktweise und monoton aber nicht gleichmäßig (ähnlich wie in Bsp. 1.3.6) gegen die unstetige Funktion

$$f : [0, 1] \rightarrow X, x \mapsto \begin{cases} 1, & x \in [0, 1), \\ 0, & x = 1. \end{cases}$$

und es gilt

$$\begin{aligned} \lim_{k \rightarrow \infty} \sup_{x \in [0, 1]} |f_k(x) - f(x)| &= \lim_{k \rightarrow \infty} \sup_{x \in [0, 1]} \begin{cases} |1 - x^k - 1|, & x \in [0, 1), \\ |1 - 1^k - 0|, & x = 1. \end{cases} \\ &= \lim_{k \rightarrow \infty} \sup_{x \in [0, 1]} \begin{cases} x^k, & x \in [0, 1), \\ 0, & x = 1. \end{cases} \\ &= \lim_{\substack{k \rightarrow \infty \\ x \in [0, 1]}} x^k = 1 \neq 0. \end{aligned}$$



- (7) (i) **Beweis.** Wir zeigen zunächst das

Lemma 1: Kompakte Mengen in Hd-Räumen sind abgeschlossen.

Beweis. Seien (X, \mathcal{T}) ein Hd-Raum, $K \subset X$ beliebig und kompakt, und $x \in K^C := X \setminus K$.

Idee: Wenn jedes $x \in K^C$ eine offene Umgebung $U_x \subset K^C$ besitzt, so gilt $K^C = \bigcup_{x \in K^C} U_x$, welches also Vereinigung offener Mengen offen ist.

Für jedes $y \in K$ existieren aufgrund der Hd-Eigenschaft von X offene Umgebungen $U_y, U_x^y \in \mathcal{T}$ mit $x \in U_x^y, y \in U_y : U_x^y \cap U_y = \emptyset$. Die Familie $(U_y)_{y \in K}$ ist eine offene Überdeckung der kompakten Menge K . Deswegen existieren nach Definition ?? $y_1, \dots, y_n \in K$, sodass $K \subset \bigcup_{i=1}^n U_{y_i}$ für $n \in \mathbb{N}$ gilt.

Behauptung: $\left(\bigcap_{i=1}^n U_{y_i} \right) \cap K = \emptyset$

Beweis. Wir zeigen die stärkere Aussage $\left(\bigcap_{i=1}^n U_{y_i} \right) \cap \left(\bigcup_{i=1}^n U_{y_i} \right) = \emptyset$. Dies gilt, denn es gilt

$$\left(\bigcap_{i=1}^n U_{y_i} \right) \cap \left(\bigcup_{i=1}^n U_{y_i} \right) = \bigcup_{i=1}^n \left(\underbrace{\bigcap_{j=1}^n U_x^{y_j} \cap U_{y_i}}_{\subset U_x^{y_i}} \right) = \bigcup_{i=1}^n (\emptyset) = \emptyset \quad \square$$

Aus der Behauptung folgt $\bigcap_{i=1}^n \underbrace{U_x^{y_i}}_{\text{offen}} \subseteq K^C$ für welche $n??$. Somit ist x ein innerer Punkt von K , und, weil x beliebig war, K^C offen, also K abgeschlossen. \square

Aus dem Lemma folgt, dass K_n für alle $n \in \mathbb{N}$ abgeschlossen ist. Da (X, \mathcal{T}) hd ist, ist $\bigcap_{n \in \mathbb{N}} K_n$ abgeschlossen und $\bigcap_{n \in \mathbb{N}} K_n \subset K_1$. Aus Satz 52 folgt dann sofort, dass $\bigcap_{n \in \mathbb{N}} K_n$ kompakt ist. \square

- (ii) **Beweis.** Sei $(A_j)_{j \in I}$ eine offene Überdeckung der Menge $\bigcup_{i=1}^n K_i$, wobei I eine beliebige Indexmenge ist. Dann ist $(A_j)_{j \in I}$ auch eine offene "Überdeckung von jedem K_i mit $i \in \{1, \dots, n\}$. Dann existieren für jede $i \in I$ j_1, \dots, j_{m_i} mit $m \in \mathbb{N}$, sodass $K_i \subset \bigcup_{j=j_1}^{j_{m_i}} A_j$. Dann gilt

$$\bigcup_{i=1}^n K_i = \bigcup_{i=1}^n \left[\bigcup_{j=j_1}^{j_{m_i}} A_j \right],$$

und das liefert eine endliche Teilmenge. \square

- (iii) Seien (\mathbb{R}, d_2) eine metrische Raum und \mathcal{T} die induzierte Topologie und $(\mathbb{R}, \mathcal{T})$ ein Hd-Raum. Dann gilt für alle $n \in \mathbb{N}$, dass $[n, n+1]$ als abgeschlossenes und beschränktes Intervall nach Heine-Borel kompakt ist, und dass $\bigcup_{n \in \mathbb{N}} [n, n+1] = [0, \infty)$ unbeschränkt und somit nach Heine-Borel nicht kompakt ist.

- (8) (i) **Beweis.** " \Rightarrow ": Seien X kompakt und $(A_i)_{i \in I}$ eine Familie abgeschlossener Mengen mit $\bigcap_{i \in I} A_i = \emptyset$. Dann gilt

$$X = \emptyset^C = \left(\bigcap_{i \in I} A_i \right)^C \stackrel{\text{d.M.}}{=} \bigcup_{i \in I} \underbrace{(A_i)^C}_{\text{offen}}.$$

Somit liefert $(A_i)_{i \in I}$ eine offene Überdeckung von X , und deswegen existieren $i_1, \dots, i_n \in I$, sodass $X = \bigcup_{j=1}^n (A_{i_j})^C$ gilt.

Somit gilt

$$\emptyset = X^C = \left(\bigcup_{j=1}^n (A_{i_j})^C \right)^C = \bigcap_{j=1}^n A_{i_j} = \bigcap_{i=i_1}^{i_n} A_i$$

" \Leftarrow ": Seien X eine Menge mit der Durchschnittseigenschaft (D.Eig.) und $(B_i)_{i \in I}$ eine offene Überdeckung von X . Dann gilt

$$\emptyset = X^C = \left(\bigcup_{i \in I} B_i \right)^C = \bigcap_{i \in I} \underbrace{(B_i)^C}_{\text{abgeschl.}} \stackrel{\text{D.Eig.}}{\implies} \exists i_1, \dots, i_n \in I : \bigcap_{j=1}^n (B_{i_j})^C = \emptyset$$

Somit gilt analog zu (i) $X = \emptyset^C = \bigcup_{j=1}^n B_{i_j}$. Somit ist $\bigcup_{j=1}^n B_{i_j}$ eine endliche Teilüberdeckung und somit ist X kompakt. \square

- (ii) **Beweis.** Sei $(A_i)_{i \in I} \subset X$ ein Familie kompakter Mengen mit $\bigcap_{i \in I} A_i = \emptyset$. Man betrachte nun ein festes, beliebiges $i = i_0$ die kompakte Menge A_{i_0} , und definiert $B_i := A_{i_0} \cap A_i$.

Gilt $B_i = \emptyset$ für ein $i \in I$, so sind wir fertig, da der Schnitt $\bigcap_{k \in \{i, i_0\}} B_i$ schon leer ist.

Sind alle B_i nicht leer und es gilt $\bigcap_{i \in I} B_i = \bigcap_{i \in I} A_i = \emptyset$. Die B_i sind abgeschlossen in dem kompakten Teilraum A_{i_0} also können wir die Aussage der vorherigen Teilaufgabe auf die Familie $(B_i)_{i \in I} \subset A_{i_0}$ anwenden.

Es existiert also eine endliche Teilmenge $J \subset I$, sodass $\bigcap_{i \in J} B_i = \emptyset$ gilt, also existiert eine endliche Teilmenge $J' = J \cup \{i_0\} \subset I$, sodass $\bigcap_{i \in J'} A_i = \emptyset$ gilt, was zu zeigen war. \square

- (9) Es gilt $f(0) = \frac{1}{3} > 0$ und $f(\frac{1}{2}) = \frac{1}{3\sqrt{e}} - \frac{1}{\sqrt{2}} = \frac{\sqrt{2}-3\sqrt{e}}{3\sqrt{2e}} < 0$, somit besitzt f als auf \mathbb{R} differenzierbare Funktion nach dem ZWS mindestens eine Nullstelle $x^* \in (0, \frac{1}{2})$. Nun gilt $f(x) = 0 \implies \frac{e^{-x}}{3} - \sqrt{x} = 0 \implies \frac{e^{-2x}}{9} = x$. Sei also $\tilde{f} : X := [0, \frac{1}{2}] \rightarrow \mathbb{R}$, $x \mapsto \frac{e^{-2x}}{9}$ die (streng monoton fallende) kontrahierende Fixpunktfunction. Um

die Kontraktivität zu zeigen, nutzen wir den MWS, der für auf X stetige Funktionen besagt, dass für ein $\xi \in X$ und alle $x, y \in X$

$$|\tilde{f}(x) - \tilde{f}(y)| \leq |\tilde{f}'(\xi)| \cdot |x - y| = \left| \frac{-2 \cdot e^{-2\xi}}{9} \right| \cdot |x - y| \leq \left| \frac{-2 \cdot e^{(-2)\cdot 0}}{9} \right| \cdot |x - y| = \frac{2}{9} \cdot |x - y| \text{ gilt.}$$

Das Kontraktionsmodul von \tilde{f} ist also $\lambda := \frac{2}{9} \in (0, 1)$. Wir wählen $x_0 := 0 \in X$ und definieren für alle $k \in \mathbb{N}$ $x_{k+1} := \tilde{f}(x_k)$. Nach Satz 46 gilt $(x_k)_{k \in \mathbb{N}} \rightarrow x^*$, und nach Aufgabe 3b aus der dritten Hausaufgabe gilt $d(x_n, x^*) \leq \frac{\lambda^n}{1-\lambda} d(x_1, x_0)$ für $n \in \mathbb{N}$ und eine beliebiges festes $\lambda \in (0, 1)$. Nach unserem Verfahren gilt $x_1 = \tilde{f}(0) = \frac{1}{9}$, also gilt

$$d(x_n, x^*) \leq \frac{\lambda^n}{1-\lambda} \cdot \frac{1}{9} = \frac{\left(\frac{2}{9}\right)^n}{1-\frac{2}{9}} \cdot \frac{1}{9} = \frac{1}{7} \cdot \left(\frac{2}{9}\right)^n.$$

Um die gewollte Genauigkeit zu erhalten, lösen wir $\frac{1}{7} \cdot \left(\frac{2}{9}\right)^n \leq \frac{1}{100}$, und erhalten $n \geq 2$. Nach dem gewählten Verfahren gilt $x_2 = \tilde{f}(x_1) \approx 0.089$ und $f(x) = 0 \implies x \approx 0.092$.

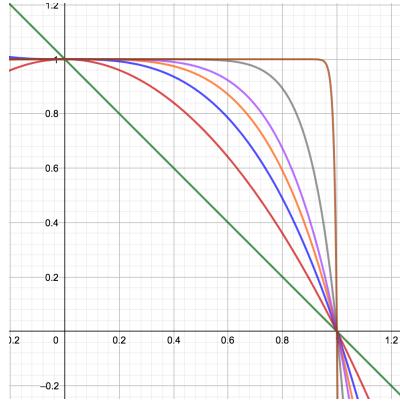


Abb. 18: Die Funktionenfolge $(f_n : [0, 1] \rightarrow X, x \mapsto 1 - x^n)_{n \in \mathbb{N}}$ für $n \in \{1, 2, 3, 4, 5, 10, 100\}$

1.5 Zusammenhang

- Die Rolle der Intervalle in \mathbb{R} wird in metrischen Räumen übernommen von den sogenannten zusammenhängenden Mengen, die wir jetzt kennenlernen.

Jede Menge X mit mindestens zwei Elementen lässt sich trivialerweise schreiben als Vereinigung zweier nicht-leerer disjunkter Teilmengen. Aber nicht jeder metrische Raum lässt sich schreiben als Vereinigung zweier nichtleerer disjunkter offener Mengen.

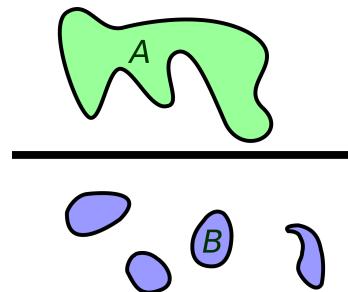


Abb. 19: A ist zusammenhängend, B, A^C und B^C sind es nicht. [Quelle]

DEFINITION 1.5.1 (ZUSAMMENHANG)

Sei (X, d) ein metrischer Raum.

- ① X heißt **zusammenhängend**, wenn es nicht die Vereinigung zweier nichtzusammenhängender leerer disjunkter offener Mengen ist:
Für alle offenen $U, V \subset X$ mit $U \cap V = \emptyset$ und $U \cup V = X$ gilt $U = \emptyset$ oder $V = \emptyset$.
Das ist äquivalent zur Forderung, dass \emptyset und X die einzigen zugleich offen und abgeschlossen Teilmengen sind.
- ② Eine Teilmenge $A \subset X$ heißt **zusammenhängend**, wenn (A, d_A) zusammenhängend ist.
- ③ X heißt **wegzusammenhängend**, falls es zu je zwei Punkten $p, q \in X$ einen **wegzusammenhängenden** Weg von p nach q , d.h. eine stetige Abbildung $\gamma : [0, 1] \rightarrow X$ mit $\gamma(0) = p$ und $\gamma(1) = q$ gibt.⁴ ⁵

SATZ 1.5.1: ZUSAMMENHÄNGENDE TEILMENGEN

Seien (X, d) ein metrischer Raum und $A \subset X$. Dann ist A genau dann zusammenhängend, wenn für alle offenen Mengen $U, V \subset X$ mit $U \cap V = \emptyset$ und $U \cup V \supset A$ gilt: $U \cap A = \emptyset$ oder $V \cap A = \emptyset$.

Beweis. Zu (\implies). Sei (A, d_A) zusammenhängend und seien $U, V \subset X$ wie im Satz. Dann sind $U \cap A$ und $V \cap A$ in A offene Teilmengen von (A, d_A) mit leerem Durchschnitt, deren Vereinigung A ist. Also ist $U \cap A = \emptyset$ oder $V \cap A = \emptyset$.

Zu (\impliedby). Wir wollen zeigen, dass (A, d_A) zusammenhängend ist. Seien also $U, V \subset A$ offen in A mit $U \cap V = \emptyset$ und $U \cup V = A$. Dann existiert nach Satz 1.2.4 offene

Mengen \tilde{U}, \tilde{V} von X mit $\tilde{U} \cap A = U, \tilde{V} \cap A = V$.

Aber um die Voraussetzungen anwenden zu können, müssen \tilde{U} und \tilde{V} disjunkt sein. Deshalb müssen wir die Erweiterungen \tilde{U} und \tilde{V} von U und V sorgfältig konstruieren.

Wir wählen zu jedem $x \in U$ ein $\varepsilon_x > 0$ mit $U_{\varepsilon_x}(x) \cap A \subset U$. Das geht, weil U in A offen ist. Wir definieren $\tilde{U} := \bigcup_{x \in U} U_{\frac{1}{2}\varepsilon_x}(x)$ und \tilde{V} entsprechend.

Natürlich sind das offene Teilmengen von (X, d) , und sie sind auch disjunkt: Wäre $z \in \tilde{U} \cap \tilde{V}$, so gäbe es $x \in U$ und $y \in V$ mit

$$z \in U_{\frac{1}{2}\varepsilon_x}(x) \cap U_{\frac{1}{2}\varepsilon(y)}(y).$$

Sei o.B.d.A $\varepsilon(y) \leq \varepsilon_x$. Dann ist aber

$$d(x, y) \leq d(x, z) + d(z, y) \leq \frac{1}{2}\varepsilon_x + \frac{1}{2}\varepsilon(y) \leq \varepsilon_x$$

Dann wäre aber $y \in U$ im Widerspruch zur Voraussetzung, weil dann $y \in U_{\varepsilon_x}(x) \subset U$ gelten würde. Also ist $\tilde{U} \cap \tilde{V} = \emptyset$.

Jetzt können wir die Voraussetzungen auf \tilde{U} und \tilde{V} anwenden und erhalten $\tilde{U} \cap A = U = \emptyset$ oder $\tilde{V} \cap A = V = \emptyset$. \square

Bemerkung 1.5.2 Obwohl Zusammenhang ein rein topologischer Begriff ist, gilt die Rückrichtung des obigen Beweises nicht in topologischen Räumen, wohl aber in HAUSDORFF-Räumen. Seien z.B. $X := \{a, b, c\}$ für $a, b, c \in \mathbb{R}$ und $\mathcal{T} := \{\emptyset\} \cup \{Y \subset X : a \in X\}$. Dann ist \mathcal{T} eine Topologie auf X , jedoch enthält die Teilmenge $A := \{b, c\} \subset X$ die offenen disjunktten Teilmengen $\{b\}$ und $\{c\}$ aber die lassen sich nicht zu disjunktten in X offenen Teilmengen erweitern, weil jede solche den Punkt a enthält.

SATZ 1.5.2: ÜBER ZUSAMMENHANG UND ABSCHLUSS

Seien (X, \mathcal{T}) ein topologischer Raum sowie $A \subset X$ eine zusammenhängende Teilmenge und $A \subset B \subset \bar{A} \subset X$. Dann ist B und insbesondere \bar{A} zusammenhängend.

Beweis. Seien $U, V \subset X$ offen und disjunkt mit $B \subset U \cup V$. Wir müssen nach Satz 1.5.1 zeigen, dass $U \cap B = \emptyset$ oder $V \cap B = \emptyset$. Weil $A \subset B$ zusammenhängend ist, gilt $U \cap A = \emptyset$ oder $V \cap A = \emptyset$.

Sei o.B.d.A $U \cap A = \emptyset$. Wäre $U \cap B \neq \emptyset$, so gäbe es also ein $b \in U \cap B$ und, weil U offen ist, dazu ein $\varepsilon > 0$ mit $U_\varepsilon(b) \subset U$. Natürlich ist $b \in \bar{A}$. Also liegen nach Definition in $U_\varepsilon(b)$ auch Punkte von A . Die liegen dann aber in U . Widerspruch zur Annahme $U \cap A = \emptyset$! \square

SATZ 1.5.3: ZUSAMMENHANG IN \mathbb{R}

Die zusammenhängenden Teilmengen von \mathbb{R} sind genau die Intervalle.

Beweis. Sei J $\subset \mathbb{R}$ ein Intervall. Seien $U, V \subset \mathbb{R}$ offen mit $U \cap V = \emptyset$ und $J \subset U \cup V$.

Wir nehmen an, es gibt ein $p \in U \cap J, r \in V \cap J$ und o.B.d.A. $p < r$. Sei $q := \sup\{t \mid [p, t] \subset U\}$. Dann ist nach Voraussetzung $q \in J$, da $p \leq q \leq r < \infty$.

Dann ist $q \notin U$, denn andernfalls wäre wegen der Offenheit von U auch $[p, q + \varepsilon] \subset U$ für kleines $\varepsilon > 0$, was im Widerspruch zur Wahl von q steht.

Andererseits ist $q \notin V$, weil sonst wegen der Offenheit von V auch $q - \varepsilon \in V$ für kleines $\varepsilon > 0$ liegen würde, was im Widerspruch zur Wahl von q steht.

Damit ist $q \notin U \cup V$. Widerspruch zu $q \in J \subset U \cup V$.

[BILD!!]

Sei $J \subset \mathbb{R}$ zusammenhängend. Seien $p < q < r$ mit $p, r \in J$. Wäre $q \notin J$, so wäre $J = [(-\infty, q) \cup (q, \infty)] \cap J$, also $J \subset (-\infty, q)$ oder $J \subset (q, \infty)$ was im Widerspruch dazu steht, dass J zusammenhängend ist. \square

AUFGABEN

- ① [HW07 von Math104 von des USC Berkeley] Seien (X, d) ein metrischer Raum und $(S_i)_{i \in I}$ eine Familie zusammenhängender Teilmengen. Zeigen Sie, dass wenn $T := \bigcap_{i \in I} S_i \neq \emptyset$ gilt, die Vereinigung $S := \bigcup_{i \in I} S_i$ zusammenhängend ist.

②



LÖSUNGEN

- ① Wir nehmen an, S wäre nicht zusammenhängend. Dann existieren disjunkte nichtleere Mengen $A, B \subset X$, sodass $A \cup B = S$ gilt und A sowie B in S offen und abgeschlossen ist. Definieren nun $A_i := A \cap S_i$ und $B_i := B \cap S_i$ für alle $i \in I$. Dann gilt $S_i = A_i \cup B_i$, $A_i \cap B_i$ und

②



1.6 Stetige Abbildungen

- Nachdem der Konvergenzbegriff in metrischen Räumen erklärt ist, ist es leicht, auch die Stetigkeit von Abbildungen solcher Räume zu erklären.
- Wir machen uns mit der Bedeutung dieses Begriffes in verschiedenen einfachen Situationen vertraut und formulieren Rechenregeln für stetige Abbildungen.
- Etwas abstrakter ist die Charakterisierung der Stetigkeit mittels offener oder abgeschlossener Mengen.

DEFINITION 1.6.1 (STETIGKEIT)

Seien (X, d_X) und (Y, d_Y) metrische Räume, $G \subset X$ eine Teilmenge und $f : X \supset G \rightarrow Y$ eine Abbildung.

- ① f heißt stetig in $p \in G$, wenn $\lim_{x \rightarrow p} f(x) = f(p)$ ist, d.h. wenn für jede gegen p konvergente Folge $(x_k)_{k \in \mathbb{N}}$ in G auch $\lim_{k \rightarrow \infty} f(x_k) = f(p)$ ist.
- ② f heißt stetig in oder auf G , wenn es stetig in jedem Punkt $p \in G$ ist.

Bemerkung Offenbar ist f im Sinne dieser Definition stetig (in $p \in G$), genau dann, wenn es (in p) stetig ist als Abbildung des metrischen Raumes $(G, d_G) \rightarrow Y$.

KOROLLAR 1.6.2 (KOMPONENTENWEISE STETIGKEIT)

Ist (X, d_X) beliebig und $(Y, d_Y) = \mathbb{R}^m$, so ist $f = (f_1, \dots, f_m)$ genau dann stetig in p , wenn alle Komponentenfunktionen $f_i : X \rightarrow \mathbb{R}$ in p stetig sind.

Beweis. Das folgt unmittelbar aus der Definition und Satz 1.3.1. □

Bei einer Abbildung $f : X \rightarrow \mathbb{R}^m$ kann man also die Stetigkeit einfach an den (reellwertigen) Komponentenfunktionen untersuchen, \mathbb{R}^m oder \mathbb{R} auf der rechten Seite machen also "keinen großen Unterschied". Jetzt betrachten wir umgekehrt eine Funktion $f : \mathbb{R}^n \supset G \rightarrow Y$. Dann können wir $f(x) = f(x_1, \dots, x_n)$ als Funktion jeder einzelnen Variablen betrachten, indem wir uns vorstellen, dass die anderen festbleiben.

DEFINITION 1.6.3 (PARTIELLE STETIGKEIT)

Die Funktion $f : X \rightarrow \mathbb{R}^m$ heißt partiell stetig in p , wenn alle Funktionen

$$x_1 \mapsto f(x_1, p_2, \dots, p_n), \quad x_2 \mapsto f(p_1, x_2, p_3, \dots, p_n), \quad \dots, \quad x_n \mapsto f(p_1, p_2, \dots, x_n)$$

stetig sind.

SATZ 1.6.1: (PARTIELLE) STETIGKEIT

Jede stetige Funktion ist partiell stetig.

BEISPIEL 1.6.4 (BEISPIEL FÜR DIE OBIGE RÜCKRICHTUNG)

Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch $f(0, 0) := 0$ und

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

Für alle $x \neq 0 \neq y$ gilt $f(x, 0) = f(0, y) = 0$, also ist f in $(0, 0)$ partiell stetig.

Für $\lambda \in \mathbb{R}$ geht aber die Folge $(\frac{1}{k}, \frac{\lambda}{k})$ gegen $(0, 0)$, aber es ist

$$f\left(\frac{1}{k}, \frac{\lambda}{k}\right) = \frac{\lambda}{k^2\left(\frac{1}{k^2} + \frac{\lambda^2}{k^2}\right)} = \frac{\lambda}{1 + \lambda^2}.$$

Für $\lambda \neq 0$ und $k \rightarrow \infty$ geht das also nicht gegen $0 = f(0, 0)$.

Dieses Beispiel zeigt genauer, warum partielle Stetigkeit viel schwächer ist als "totale" Stetigkeit: Die Variable x muss sich der Stelle p auf beliebige Weise nähern dürfen. Bei der partiellen Stetigkeit schränkt man sich aber auf achsenparallele Annäherung ein.

In unserem Beispiel ist die Funktion auf allen Geraden durch den Nullpunkt jeweils konstant (Wert $\lambda/(1 + \lambda^2)$), nur im Nullpunkt hat sie definitionsgemäß den Wert 0. Der kommt heraus, wenn man auf der x -Achse ($\lambda = 0$) oder auf der y -Achse ($\lambda = \infty$) an den Nullpunkt heranläuft, aber eben nur dann. Selbst wenn $f(x) \rightarrow f(p)$ bei Annäherung auf allen Geraden durch p gilt, folgt daraus nicht die Stetigkeit in p . Ein Gegenbeispiel liefert die Funktion g mit $g(x, y) = 1$, falls $y = x^2 \neq 0$, und $g(x, y) = 0$ sonst. Wie sieht der Graph dieser Funktion aus (vgl Abb. ??)?

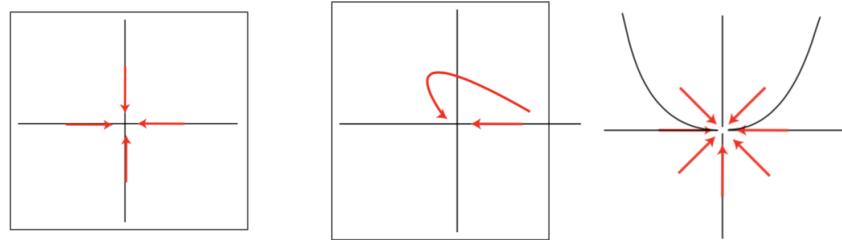


Abb. 20: Quelle: Ferus

Beispiel 1.6.5 Die Abbildungen

$$\alpha : \mathbb{R}^2 \rightarrow \mathbb{R}, (x_1, x_2) \mapsto x_1 + x_2$$

$$\mu : \mathbb{R}^2 \rightarrow \mathbb{R}, (x_1, x_2) \mapsto x_1 x_2$$

$$\eta : \mathbb{R}^2 \setminus \{(x_1, x_2) \mid x_2 \neq 0\} \rightarrow \mathbb{R}, (x_1, x_2) \mapsto \frac{x_1}{x_2}$$

sind stetig.

Wir zeigen das für α . Seien $p = (p_1, p_2) \in \mathbb{R}^2$ und $(x_k = (x_{k1}, x_{k2}))_{k \in \mathbb{N}}$ eine Folge mit $\lim x_k = p$. Dann ist

$$d(\alpha(x_k), \alpha(p)) = |(x_{k1} + x_{k2}) - (p_1 + p_2)| \leq |x_{k1} - p_1| + |x_{k2} - p_2|.$$

Aber nach Satz 1.3.1 folgt aus $\lim x_k = p$, dass $\lim x_{ki} = p_i$ für $i = 1, 2$. Daher geht die rechte Seite gegen Null und $\lim \alpha(x_k) = \alpha(p)$.

Beweis für μ und η selbst. Für den letzteren Fall benutzt man die Abschätzung

$$\left| \frac{x_1}{x_2} - \frac{p_1}{p_2} \right| = \left| \frac{x_1 p_2 - x_2 p_1}{x_2 p_2} \right| \leq \frac{|x_1 - p_1||p_2| + |p_1||x_2 - p_2|}{|x_2 p_2|}.$$

Beispiel 1.6.6 Dieselben Argumente wie im vorstehenden Beispiel zeigen die Stetigkeit der Determinante

$$\det : M(n \times n, \mathbb{R}) \rightarrow \mathbb{R}, \quad (x_{ij}) \mapsto \sum_{\sigma \in S_n} \text{sign}(\sigma) \prod_{i=1}^n x_{i,\sigma(i)},$$

wenn man den Raum $M(n \times n, \mathbb{R})$ der quadratischen n -reihigen Matrizen auf die offensichtliche Weise mit dem \mathbb{R}^{n^2} identifiziert und eine der l^p -Metriken verwendet. Auch die mit der Transponierten gebildete Abbildung

$$M(n \times n, \mathbb{R}) \rightarrow M(n \times n, \mathbb{R}), \quad A \mapsto AA^T$$

gebildete Abbildung ist stetig.

Korollar 1.6.7

Seien (X, d) metrischer Raum und $a \in X$. Dann ist $d(., a) : X \rightarrow \mathbb{R}, x \mapsto d(x, a)$ stetig.

Beweis. Wir zeigen $d(d(x, a), d(y, a)) = |d(x, a) - d(y, a)| \leq d(x, y) \rightarrow 0$ für alle $x, y \in X$.

Fall 1: $d(x, a) \geq d(y, a)$. Dann gilt $|d(x, a) - d(y, a)| = d(x, a) - d(y, a)$ und dann folgt die Behauptung durch Umstellen von der Dreiecksungleichung $d(x, a) \leq d(x, y) + d(y, a)$.

Fall 2: $d(y, a) \geq d(x, a)$. Dann gilt $|d(x, a) - d(y, a)| = d(y, a) - d(x, a)$ und dann folgt die Behauptung durch Umstellen von der Dreiecksungleichung $d(y, a) \leq d(y, a) + d(a, x)$. \square

Beispiel 1.6.8 Sei $(X, d) = (C^0([a, b], \mathbb{R}), d^{\text{sup}})$. Dann ist die Abbildung

$$\int_a^b : X \rightarrow \mathbb{R}, \quad f \mapsto \int_a^b f(x) dx$$

stetig, denn es gilt

$$\begin{aligned} \left| \int_a^b f(x) dx - \int_a^b g(x) dx \right| &\leq \int_a^b |f(x) - g(x)| dx \leq \int_a^b \sup_{a \leq x \leq b} |f(x) - g(x)| dx \\ &= |b - a| d^{\text{sup}}(f, g) \rightarrow 0. \end{aligned}$$

SATZ 1.6.2: KOMPOSITIONEN STETIGER ABBILDUNGEN SIND STEТИG.

Seien (X_i, d_i) , $i = 1, 2, 3$ metrische Räume und $f_i : X_i \supset G_i \rightarrow X_{i+1}$ für $i = 1, 2$ Abbildungen mit $f_1(G_1) \subset G_2$. Es sei f_1 stetig in $p_1 \in G_1$ und f_2 stetig in $p_2 := f_1(p_1)$. Dann ist $f_2 \circ f_1 : X_1 \supset G_1 \rightarrow X_3$ stetig in p_1 .

Beweis. Sei $p \in G_1$ und $(x_n)_{n \in \mathbb{N}}$ ein Folge in G_1 mit $x_n \rightarrow p$. Dann gilt $f_1(x_n) \xrightarrow{n \rightarrow \infty} f_1(p)$, weil f_1 stetig ist und somit $f_2(f_1(x_n)) = f_2(f_1(p))$, weil f_2 stetig ist. \square

Korollar 1.6.9

Ist (X, d) ein metrischer Raum, und sind $f, g : X \supset G \rightarrow \mathbb{R}$ stetig in $p \in G$, ist ferner $\lambda \in \mathbb{R}$, so sind auch die Abbildungen

$$f + g, f - g, fg, \lambda f : X \supset G \rightarrow \mathbb{R}$$

in p stetig. Ist $g(p) \neq 0$, so ist $\frac{f}{g} : X \supset \{x \in G \mid g(x) \neq 0\} \rightarrow \mathbb{R}$ in p stetig.

Insbesondere ist also der Vektorraum $C^0(X, \mathbb{R}) := \{f : X \rightarrow \mathbb{R} \mid f \text{ stetig}\}$ ein reeller Vektorraum.

Beweis. Die Abbildungen sind vom Typ $x \mapsto (f(x), g(x)) \xrightarrow{\alpha} f(x) + g(x)$, wobei die erste stetige Komponentenfunktionen besitzt. \square

SATZ 1.6.3: ε - δ -KRITERIUM FÜR STETIGKEIT

Seien $(X, d_X), (Y, d_Y)$ metrische Räume und $f : X \supset G \rightarrow Y$. Dann gilt ist f genau dann stetig in $p \in G$, wenn

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in G : [d_X(p, x) < \delta \implies d_Y(f(p), f(x)) < \varepsilon].$$

gilt. Die letzte Bedingung ist äquivalent zu $f(U_\delta(p) \cap G) \subset U_\varepsilon(f(p))$.

Beweis. Zu (\implies). Sei $\varepsilon > 0$. Gäbe es kein δ wie angegeben, so gäbe es insbesondere zu jedem $k \in \mathbb{N}$ ein $x_k \in G$ mit

$$d_X(p, x_k) < \delta_k := \frac{1}{k+1}, \text{ aber } d_Y(f(p), f(x_k)) \geq \varepsilon.$$

Dann wäre aber $\lim x_k = p$ und $\lim_{k \rightarrow \infty} f(x_k) \neq f(p)$. Widerspruch!

Zu (\impliedby). Sei $(x_k)_{k \in \mathbb{N}}$ eine gegen p konvergente Folge in G . Wir müssen zeigen, dass $\lim f(x_k) = f(p)$. Sei also $\varepsilon > 0$ beliebig. Sei $\delta > 0$ dazu gewählt wie im Satz. Dann existiert ein $k_0 \in \mathbb{N}$ mit $d_X(x_k, p) < \delta$ für alle $k \geq k_0$. Dann ist aber $d_Y(f(x_k), f(p)) < \varepsilon$ für alle $k \geq k_0$. \square

SATZ 1.6.4: STETIGE URBILDER OFFENER MENGEN SIND OFFEN

① Seien $(X, d_X), (Y, d_Y)$ metrische Räume und $f : X \rightarrow Y$.

Dann ist f stetig genau dann, wenn für jede offene Teilmenge $V \subset Y$ das Urbild $f^{-1}(V) \subset X$ offen in X ist.

② Die Aussage bleibt richtig, wenn man überall "offen" durch "abgeschlossen" ersetzt.

③ Ist $f : X \supset G \rightarrow Y$, so muss in (i) (X, d) durch (G, d_G) ersetzt werden.

Beweis. Zu (i \implies). Seien also f stetig und $V \subset Y$ offen. Wir wollen zeigen, dass $f^{-1}(V)$ offen ist. Sei dazu $x \in f^{-1}(V)$. Zu $f(x) \in V$ existiert dann ein $\varepsilon > 0$ mit

$$U_\varepsilon(f(x)) \subset V.$$

Dazu existiert dann ein $\delta > 0$ mit

$$f(U_\delta(x)) \subset U_\varepsilon(f(x)) \subset V.$$

Das bedeutet aber $U_\delta(x) \subset f^{-1}(V)$.

Zu (i \impliedby). Sei das Urbild jeder offenen Menge offen und sei $p \in X$.

Sei weiter $\varepsilon > 0$. Dann ist $U_\varepsilon(f(p))$ offen, also ist $f^{-1}(U_\varepsilon(f(p)))$ offen und damit eine Umgebung von p . Daher existiert ein $\delta > 0$ mit $U_\delta(p) \subset f^{-1}(U_\varepsilon(f(p)))$, d.h. mit

$$f(U_\delta(p)) \subset U_\varepsilon(f(p)).$$

Das war aber zu zeigen.

Zu (ii). Die Aussage über abgeschlossene Mengen folgt, weil $A \subset Y$ genau dann abgeschlossen ist, wenn $Y \setminus A$ offen ist, und weil andererseits $f^{-1}(Y \setminus A) = X \setminus f^{-1}(A)$ gilt.

Zu (iii). Die Offenheit in (G, d_G) ist klar nach (i). \square

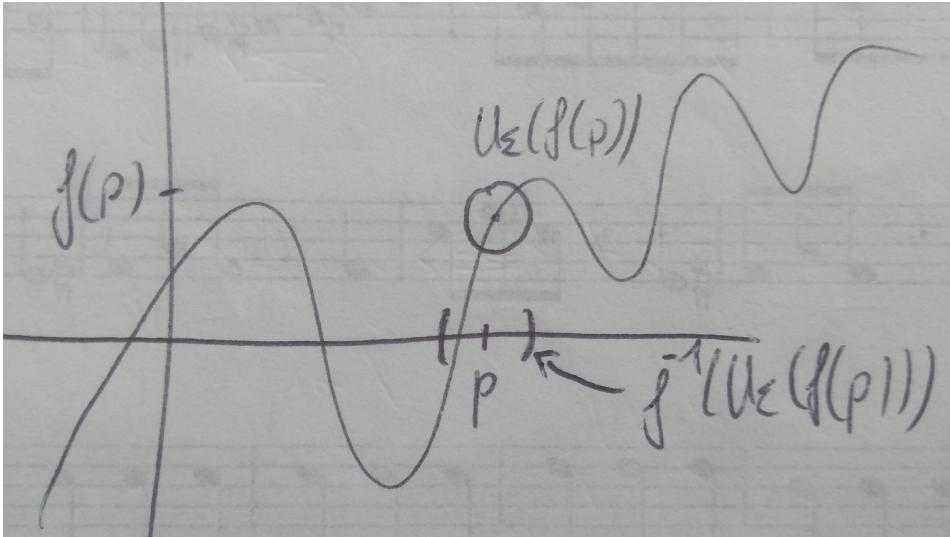


Abb. 21: Eine Verdeutlichung der Rückrichtung von (i).

Beispiel 1.6.10 (stetige Bilder offener Mengen sind i.A. nicht offen)

- Das Bild der offenen Menge $(0, 1)$ unter der konstanten Abbildung $f \equiv 1$ ist die abgeschlossene einelementige Menge $\{1\}$.
- Das Bild der offenen Menge $(-2\pi, 2\pi)$ unter der stetigen Abbildung $x \mapsto \sin(x)$ ist das kompakte Intervall $[-1, 1]$.

Beispiel 1.6.11 (Einheitssphäre & Matrizengruppen)

① Die Einheitssphäre

$$S^n := \left\{ x \in \mathbb{R}^{n+1} \mid \sum_{i=1}^{n+1} x_i^2 = 1 \right\}$$

ist das Urbild von $\{1\} \in \mathbb{R}$ unter der stetigen Abbildung $x \mapsto \sum x_i^2$, es gilt $d_2(\cdot, 0)^{-1} = S^n$. Also ist sie abgeschlossen und, weil beschränkt, auch kompakt.

allgemeine lineare ② Die allgemeine lineare Gruppe der invertierbaren Matrizen $\text{GL}(n, \mathbb{R}) := \{A \in \mathbb{R}^{n \times n} : A^{-1} \text{ existiert}\} = \det^{-1}(\mathbb{R} \setminus \{0\})$ ist offen in $M(n \times n, \mathbb{R})$.

spezielle lineare ③ Die spezielle lineare Gruppe $\text{SL}(n, \mathbb{R}) := \{A \in \text{GL}(n, \mathbb{R}) : \det(A) = 1\} = \det^{-1}(\{1\})$ ist abgeschlossen in $M(n \times n, \mathbb{R})$.

orthogonalen Matrizen ④ Die Gruppe der orthogonalen Matrizen $\mathcal{O}(n) := \{A \in \mathbb{R}^{n \times n} : AA^T = E_n\}$ ist abgeschlossen in $M(n \times n, \mathbb{R})$. Weil alle ihre Komponenten maximal vom Betrag 1 sind, ist sie auch beschränkt und damit kompakt in $M(n \times n, \mathbb{R}) = \mathbb{R}^{n^2}$.

Gruppen ⑤ Die angegebenen Teilmengen von $M(n \times n, \mathbb{R})$ sind tatsächlich Gruppen bezüglich der Matrixmultiplikation. Sie haben außerdem eine von $M(n \times n, \mathbb{R})$ geerbte Metrik, in der die Gruppenoperationen stetig sind.

AUFGABEN

- ① Betrachten Sie (\mathbb{R}^n, d_2) für $n \in \mathbb{N}$ eine konvexe Teilmenge $K \subseteq \mathbb{R}^n$, d.h. für alle $x, y \in K$ ist $tx + (1-t)y \in K$ für alle $t \in [0, 1]$.
- ① Zeigen Sie, dass K wegzusammenhängend ist.
 - ② Nehmen Sie an, dass $U, V \subseteq \mathbb{R}^n$ offen sind mit $K \subseteq U \cup V$ und $U \cap V = \emptyset$. Zeigen Sie, dass dann aus $K \cap U \neq \emptyset$ bereits $V \cap K = \emptyset$ folgt. Folgern Sie, dass K zusammenhängend sein muss.
 - ③ Zeigen Sie, dass für $n = 1$ und $x \in \mathbb{R}^n$ der Raum $X := \mathbb{R}^n \setminus \{x\}$ mit der induzierten Metrik nicht zusammenhängend ist.
- ② Sei $f : X \rightarrow Y$ eine Abbildung zwischen topologischen Räumen (X, \mathcal{T}) und (Y, \mathcal{S}) . Zeigen Sie, dass f immer stetig ist, wenn mindestens eine der folgenden Bedingungen gilt
- ① f ist konstant, d.h. $f(x) = y$ für alle $x \in X$ mit einem $y \in Y$.
 - ② die Topologie \mathcal{T} wird von der diskreten Metrik induziert,
 - ③ $\mathcal{S} = \emptyset, Y\}$.
- ③ In dieser Aufgabe sei \mathbb{R}^n immer mit der Standardmetrik d_2 versehen. Bestimmen Sie, in welchen Punkten die folgenden Abbildungen stetig sind:

$$g : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto \begin{cases} \frac{x(x^2 - y^2)}{x^2 + y^2} & \text{für } (x, y) \neq (0, 0) \\ 0, & \text{sonst.} \end{cases}$$

$$h : \mathbb{R}^2 \rightarrow \mathbb{R}^3, (x, y) \mapsto (xy, x^2 - y^2, 42e^x).$$

LÖSUNGEN

(1) (i) **Beweis.** Für alle $x, y \in K$ definieren wir $\gamma(t) := (1-t)x + ty \in K \forall t \in [0, 1]$. Dann gilt $\gamma(0) = x$ und $\gamma(1) = y$, und da die Komposition stetiger Abbildungen stetig ist, ist die Abbildung γ stetig. \square

(ii) **Beweis.** Sei $U \cap V \neq \emptyset$, dann existiert ein $x_0 \in U \cap K$. Angenommen, es existiert ein $y \in K \cap V \neq \emptyset$. Dann gilt $\gamma(t) := (1-t)x_0 + ty \in K$ für alle $t \in [0, 1]$.

Sei $q := \sup\{t \in [0, 1] : \gamma(t) \in U\}$. Es gilt $\gamma(0) = x_0 \in U$ aber $x_0 \notin V$, weil U und V disjunkt sind, also muss $q > 0$ gelten. Außerdem gilt $\gamma(1) = y \in V$ also muss $q < 1$ gelten.

Annahme 1: $\gamma(q) \in U$. Dann gilt, weil U offen und γ stetig ist, dass ein $\varepsilon > 0$ existiert, sodass $\gamma(q + \varepsilon) \in U$ gilt, was einen Widerspruch zur Definition von q darstellt.

Annahme 2: $\gamma(q) \in V$. Dann gilt, weil U offen und γ stetig ist, dass ein $\varepsilon > 0$ existiert, sodass $\gamma([q - \varepsilon, q]) \in V$ gilt, was einen Widerspruch zur Definition von q darstellt.

Also gilt $\gamma(q) \notin U \cup V$. Jedoch gilt $\gamma(q) \in K$, was einen Widerspruch zu $K \subset U \cup V$ darstellt. Also existiert kein $y \in K \cap V \neq \emptyset$ und somit gilt $U \cap K \neq \emptyset \implies V \cap K = \emptyset \implies K$ zusammenhängend. \square

Der vorangegangene Beweis orientierte sich an dem Beweis im Skript, der zeigte, dass Intervalle zusammenhängend sind. Der folgenden alternative Beweis ist eleganter und abstrakter.

Beweis. Sei $x_0 \in U \cap K$. Wir behaupten, dass kein $y \in V \cap K$ existiert, also K zusammenhängend ist.

Andernfalls existiert eine Funktion $\gamma : [0, 1] \rightarrow K$ mit $\gamma(0) = x_0$ und $\gamma(1) = y \in V$.

Weil die Funktion γ stetig ist (vgl. 1a), sind $\gamma^{-1}(U \cap K), \gamma^{-1}(V \cap K) \subseteq [0, 1]$ offen.

Außerdem ist $[0, 1] \subseteq \gamma^{-1}(U \cap K) \cup \gamma^{-1}(V) = \gamma^{-1}(U) \cup \gamma^{-1}(V)$ eine offene Überdeckung.

Nun gilt $\gamma^{-1}(U) \cap \gamma^{-1}(V) = \emptyset$ weil U und V disjunkt sind. Da $[0, 1]$ zusammenhängend ist, gilt $\gamma^{-1}(U) = \emptyset$ oder $\gamma^{-1}(V) = \emptyset$.

Wir wissen, dass $\gamma(0) = x_0 \in U$ und somit $\gamma^{-1}(U) \neq \emptyset$ gilt, also muss $\gamma^{-1}(V) = \emptyset$ gelten, was im Widerspruch zu der Annahme $y \in V$ ist. Somit gilt $K \cap V = \emptyset$. \square

(iii) **Beweis.** Seien $U := (-\infty, x)$ und $V := (x, \infty)$. Dann gilt $X = U \cup V$ und $U \cap V = \emptyset$. \square

Im Allgemeinen gilt

$$\mathbb{R}^n \setminus \{x\} \text{ ist } \begin{cases} \text{nicht zusammenhängend} & n = 1 \\ \text{zusammenhängend} & \text{sonst.} \end{cases}$$

(2) (i) **Beweis.** Sei $U \subset S$, dann ist U offen und es gilt $f^{-1}(U) = \begin{cases} X, & y \in U \\ \emptyset, & y \notin U \end{cases}$, und da X und \emptyset offen ist, gilt $f^{-1}(U) \in \mathcal{T}$, und somit ist f nach Satz 77 stetig. \square

- (ii) **Beweis.** \mathcal{T} induziert von Metrik d , falls $U \in \mathcal{T} \iff U$ offen im Sinne des metrischen Raums.

Sei d^{disk} die diskrete Metrik. Dann gilt $\mathcal{T} = \mathcal{P}(X)$, weil bzgl. d^{disk} alle Mengen offen sind.

Sei $U \in S$ beliebig. Dann gilt $f^{-1}(U) \subset X \implies f^{-1}(U) \subset \mathcal{T}$ und somit ist f stetig. \square

Wenn jedoch S von der diskreten Metrik induziert wird, gilt das nicht. Ggbsp.: (X, \mathcal{T}) mit $\mathcal{T} = \emptyset, X\}, X = \{1, 2\} = Y$ mit $S = \mathcal{P}(X)$ und $\text{id} : X \rightarrow Y, x \mapsto x$. Dann gilt $\text{id}^{-1}(\{1\}) \notin \mathcal{T}$. (WARUM?!)

- (iii) **Beweis.** Es gilt $f^{-1}(\emptyset) = \emptyset \in \mathcal{T}$ und $f^{-1}(Y) = X \in \mathcal{T}$. Somit ist f stetig. \square

- (3) (i) Da Stetigkeit mit Komposition verträglich ist, müssen wir g nur in $(0, 0)$ untersuchen.

Nun gilt für $x \geq y$ und $(x, y) \neq (0, 0)$

$$|g(x, y) - g(0, 0)| = \left| \frac{x(x^2 - y^2)}{x^2 + y^2} \right| = \left| \frac{x \left(1 - \underbrace{\frac{y^2}{x^2}}_{\leq 1} \right)}{1 + \underbrace{\frac{y^2}{x^2}}_{\geq 1}} \right| \leq \frac{|x|}{\underbrace{\left| 1 + \frac{y^2}{x^2} \right|}_{\geq 1}} \leq |x| \xrightarrow{(x,y) \rightarrow (0,0)} 0$$

Für $y > x$ teilt man durch y^2 und erhält das selbe Ergebnis. Alternativ geht man so vor:

$$|g(x, y) - g(0, 0)| = |x| \cdot \frac{|x^2 - y^2|}{x^2 + y^2} \stackrel{\Delta \neq}{=} |x| \cdot \frac{x^2 + y^2}{x^2 + y^2} = |x| \rightarrow 0 \quad (\text{s.o.})$$

- (ii) Aus der Vorlesung wissen wir, dass h stetig ist, da seine Komponentenfunktionen (vgl. (a)) stetig sind.

1.7 Fünf wichtige Sätze über stetige Abbildungen

Seien $(X, d_X), (Y, d_Y)$ metrische Räume.

SATZ 1.7.1: STETIGE BILDER ZUSAMMENHÄNGENDER/KOMPAKTER MENGEN SIND ZSMH./KOMPAKT

Seien $A \subset X$ zusammenhängend bzw. kompakt und $f : A \rightarrow Y$ stetig. Dann ist auch $f(A) \subset Y$ zusammenhängend bzw. kompakt.

Beweis. (Für Zusammenhang - Verallgemeinerung des ZWS) Seien $U, V \subset Y$ offen und disjunkt mit $f(A) \subset U \cup V$.

Nach Satz 1.6.4 sind $f^{-1}(U)$ und $f^{-1}(V)$ offen in (A, d_A) . Sie sind weiter disjunkt mit $A = f^{-1}(U) \cup f^{-1}(V)$. Weil A zusammenhängend ist, folgt $f^{-1}(U) = \emptyset$ oder $f^{-1}(V) = \emptyset$, und entsprechend ist $U \cap f(A) = \emptyset$ oder $V \cap f(A) = \emptyset$. \square

Beweis. (Für Kompaktheit) Sei $(U_i)_{i \in I}$ eine offene Überdeckung von $f(A)$. Dann ist $(f^{-1}(U_i))_{i \in I}$ eine offene Überdeckung von A . Also existiert eine endliche Teilmenge $J \subset I$, so dass $A \subset \bigcup_{i \in J} f^{-1}(U_i)$ gilt. Aber dann ist $f(A) \subset \bigcup_{i \in J} U_i$. \square

Beispiel 1.7.1 Wir untersuchen die Funktion $f : (0, \infty) \rightarrow \mathbb{R}$, $x \mapsto \frac{\ln x}{x}$ ⁶ um zu zeigen, dass stetige Abbildungen im Allgemeinen weder beschränkte Mengen in beschränkte Mengen noch abgeschlossene Mengen in abgeschlossene Mengen abbilden.

Durch Kurvenuntersuchen erkennen wir, dass (e, e^{-1}) ein globales Maximum von f ist. Das Intervall $(0, e)$ ist beschränkt in \mathbb{R} mit der Standardmetrik, aber $f((0, e)) = (-\infty, e^{-1})$ ist unbeschränkt. Analog ist $[e, \infty)$ ist abgeschlossen, aber $f([e, \infty)) = (0, e^{-1}]$ ist nicht abgeschlossen.

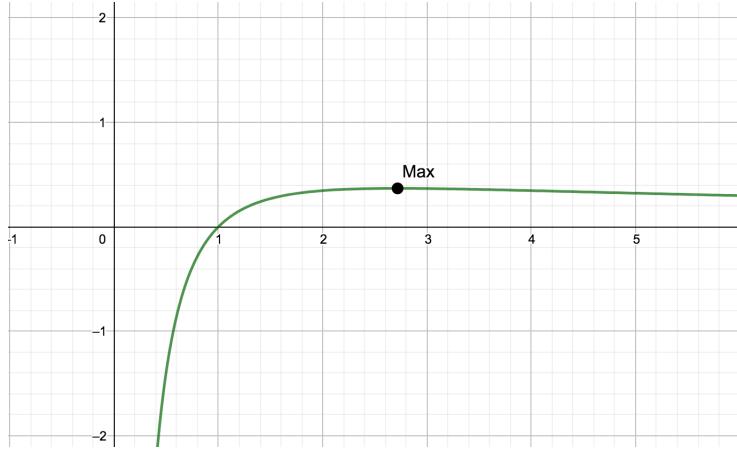


Abb. 22: Die Funktion $f : (0, \infty) \rightarrow \mathbb{R}$, $x \mapsto \frac{\ln x}{x}$ [Quelle: GeoGebra6]

SATZ 1.7.2: VOM MAXIMUM

Eine stetige reellwertige Funktion nimmt auf einem nichtleeren Kompaktum ihr Minimum und Maximum an.

⁶Geht äquivalent auch mit $g : (0, \infty) \rightarrow \mathbb{R}$, $x \mapsto \frac{1}{x}$.

Beweis. Nach Satz 1.7.1 ist $f(A) \subset \mathbb{R}$ kompakt. Sei $y^* := \sup_{a \in A} f(a) \in \mathbb{R}$. Nach Definition des Supremums existiert eine Folge $(y_n)_{n \in \mathbb{N}} \subset f(A)$ mit $y_n \rightarrow y^*$. Dann existiert aber auch eine beschränkte (weil A kompakt ist) Folge $(x_n)_{n \in \mathbb{N}} \subset A$ mit $y_k = f(x_k)$. Nach

Nach dem **Satz von BOLZANO-WEIERSTRASS** existiert eine konvergente Teilfolge $(x_{n_k})_{k \in \mathbb{N}} \subset (x_n)_{n \in \mathbb{N}}$ mit $x_{n_k} \rightarrow x^*$ für ein $x^* \in A$. Weil die Abbildung f stetig ist, gilt $f(x_{n_k}) \rightarrow f(x^*)$. Andererseits gilt $f(x_{n_k}) = y_k \rightarrow y^*$, also gilt $y^* = f(x^*) \in f(A)$. Somit ist das Supremum y^* in der Menge $f(A)$ enthalten, ist also das gesuchte Maximum.

Analog für das Minimum. □

Korollar 1.7.2

- ① Jede kompakte Menge in \mathbb{R} besitzt ein Maximum & Minimum. ($f = \text{id}?$)
- ② Sei (A, d_A) ein kompakter metrischer Raum. Dann ist jede positive stetige Funktion $f : A \rightarrow (0, \infty)$ "wegbeschränkt von Null", es existiert also ein $\varepsilon > 0$ mit $f(x) \geq \varepsilon$ für alle $x \in A$.

Bemerkung 1.7.3 (Zu ②) Die Funktion $f : (0, \infty) \rightarrow (0, \infty)$, $x \mapsto \frac{1}{x}$ ist zum Beispiel nicht "wegbeschränkt von Null", da $(0, \infty)$ nicht kompakt ist.

DEFINITION 1.7.4

Eine Abbildung $f : X \supset G \rightarrow Y$ heißt **gleichmäßig stetig** auf G , wenn gilt **gleichmäßig stetig**

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in G : f(U_\delta(x) \cap G) \subset U_\varepsilon(f(x)).$$

Die Abbildung f ist genau dann gleichmäßig stetig auf der Teilmenge $G \subset X$ in diesem Sinne, wenn es auf (G, d_G) gleichmäßig stetig ist.

SATZ 1.7.3: VON DER GLEICHMÄSSIGEN STETIGKEIT

Seien $f : X \supset A \rightarrow Y$ stetig und A kompakt. Dann ist f gleichmäßig stetig.

Beweis. Sei $\varepsilon > 0$. Aufgrund der Stetigkeit von f existiert zu jedem $x \in A$ ein $\delta_x > 0$ mit

$$f(U_{\delta_x}(x) \cap A) \subset U_{\frac{\varepsilon}{2}}(f(x)).$$

Dann ist $(U_{\delta_x}(x))_{x \in A}$ eine offene Überdeckung von A und es reichen endlich viele aus; es gilt $A \subset \bigcup_{j=1}^n U_{\delta_{x_j}}(x_j)$ für geeignete $n \in \mathbb{N}$ und $x_1, \dots, x_n \in A$.

Wir wählen nun eine **Lebesgue-Zahl** $\delta > 0$ dazu mit der Eigenschaft, dass jede **Lebesgue-Zahl** U_δ in einer der $U_{\delta_{x_j}}(x_j)$ liegt.

Dieses δ ist geeignet: ist $x \in A$ beliebig und y gewählt wie oben mit $U_\delta(x) \subset U_{\delta_y}(y)$, so folgt

$$f(U_\delta(x) \cap A) \subset f(U_{\delta_y}(y) \cap A) \subset U_{\frac{\varepsilon}{2}}(f(y))$$

Insbesondere ist $x \in U_\delta(x) \cap A$ und somit gilt nach der obigen Inklusion $f(x) \in U_{\frac{\varepsilon}{2}}(f(y))$ und deshalb

$$f(U_\delta(x) \cap A) \subset U_{\frac{\varepsilon}{2}}(f(y)) \subset U_\varepsilon(f(x)).$$

□

Die letzte Inklusion sieht man so: Wir wollen $d(f(y), a) < \frac{\varepsilon}{2} \implies d(f(x), a) < \varepsilon$ zeigen. Mit der Dreieckungleichung und der obigen Inklusion folgt

$$d(f(x), a) \stackrel{\Delta \neq}{\leqslant} d(f(x), f(y)) + d(f(y), a) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Lemma 1.7.5 (Existenz der LEBESGUEN Zahl)

Seien (X, d) ein metrischer Raum, $A \subset X$ kompakt und $(U_i)_{i \in I}$ eine offene Überdeckung von A . Dann existiert eine positive Zahl $\delta > 0$ sodass gilt

$$\forall a \in A \ \exists i \in I : U_\delta(a) \subset U_i.$$

Beweis. Wäre das nicht der Fall, würde ein $x \in A$ existieren, sodass für alle $i \in I$ $U_\delta(x) \not\subset U_i$ gelten würde. Wir wählen $(\delta_k := \frac{1}{k+1})_{k \in \mathbb{N}}$ und zu jedem δ_k ein $a_k \in A$. Wir wollen zeigen, dass für jedes $k \in \mathbb{N}$ die Kugel $U_{\delta_k}(x_k)$ in keiner der $U_{\delta_y}(y)$ enthalten ist. Weil A kompakt ist, existiert nach dem Satz von BOLZANO-Weierstrass eine konvergente Teilfolge (x_{n_k}) von (x_k) mit $x_{n_k} \rightarrow x^* \in A$. Für ein y gilt auch $x^* \in U_{\delta_y}(y)$. Weil $x_{n_k} \rightarrow x^*$ und $\delta_k \rightarrow 0$ geht, gilt $U_{\delta_{x_{n_k}}}(x_{n_k}) \subset U_{\delta_y}(y)$ für alle genügend großen k , was im Widerspruch zur Annahme steht. \square

Der in der Definition 1.3.5 eingeführte Begriff der gleichmäßigen Konvergenz einer Funktionenfolge hat mit der gleichmäßigen Stetigkeit nichts zu tun. Im Raum der beschränkten Funktionen bedeutete gleichmäßige Konvergenz einfach die Konvergenz bezüglich der Supremumsmetrik, und in dem Zusammenhang haben wir einen Spezialfall des folgenden Satzes schon kennengelernt, vgl. Satz 1.3.3.

SATZ 1.7.4: GLM. LIMITEN STETIGER FUNKTIONEN SIND STETIG

Sei die Folge $(f_i : G \rightarrow Y)_{i \in \mathbb{N}}$ auf $G \subset X$ gleichmäßig konvergent gegen $f : G \rightarrow Y$. Sind alle f_i stetig, so ist auch f stetig.

Beweis. Wir zeigen die Stetigkeit in $p \in G$. Sei $(x_j)_{j \in \mathbb{N}}$ eine Folge in G mit Grenzwert p . Wir müssen $\lim f(x_j) = f(p)$ zeigen.

Sei dazu $\varepsilon > 0$. Weil die Funktionenfolge gleichmäßig konvergiert existiert ein $k_0 \in \mathbb{N}$, so gilt

$$\forall k \geq k_0 \ \forall x \in G : d_Y(f_k(x), f(x)) < \varepsilon/3.$$

Weil f_{k_0} in p stetig ist, existiert ein j_0 mit $d_Y(f_{k_0}(x_j), f_{k_0}(p)) < \varepsilon/3$ für alle $j \geq j_0$. Dann ist aber für $j \geq j_0$

$$\begin{aligned} d_Y(f(x_j), f(p)) &\leq d_Y(f(x_j), f_{k_0}(x_j)) + d_Y(f_{k_0}(x_j), f_{k_0}(p)) + d_Y(f_{k_0}(p), f(p)) \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned} \quad \square$$

Korollar 1.7.6 (die konstante Majorante und Potenzreihen)

① Seien $(f_k : G \rightarrow \mathbb{R})_{k \in \mathbb{N}}$ eine Folge stetiger Funktionen und $S_n(x) := \sum_{k=1}^n f_k(x)$ definiert.

Ist $\sum_{k=1}^{\infty} \underbrace{\sup_{x \in G} |f_k(x)|}_{=: c_k} < \infty$, so ist $S(x) := \lim_{n \rightarrow \infty} S_n(x)$ stetig.

② Insbesondere konvergieren Potenzreihen $\sum_{k=0}^{\infty} a_k x^k$ mit Konvergenzradius $R > 0$ gleichmäßig gegen in jeder Kugel mit Radius $r < R$.

Beweis. **①** Zunächst ist $(S_n)_{n \in \mathbb{N}}$ eine CAUCHY-Folge in $(\mathcal{C}(G, \mathbb{R}), d^{\sup})$, denn für $n \geq m$ gilt

$$\begin{aligned} d^{\sup}(S_n(x), S_m(x)) &= \sup_{x \in G} \left| \sum_{k=m+1}^n f_k(x) \right| \leq \sup_{x \in G} \sum_{k=m+1}^n |f_k(x)| \\ &\leq \sum_{k=m+1}^n c_k \xrightarrow{n,m \rightarrow \infty} 0. \end{aligned} \quad (5)$$

Aufgrund der Vollständigkeit des zugrundeliegenden metrischen Raums existiert S als Grenzwert von S_n bezüglich d^{\sup} und wegen des vorangegangenen Satzes stetig.

② Es gilt

$$\sup_{\substack{x \in \mathbb{R}: \\ |x| < r}} |a_k x^k| \leq r^k |a_k|$$

und die Reihe $\sum_{k=0}^{\infty} r^k |a_k|$ konvergiert. \square

1.8 Normierte Vektorräume

- Wir lernen normierte Vektorräume kennen, die in der mehrdimensionalen Analysis als Definitions- und Zielbereiche der Funktionen dienen.
- Endlich-dimensionale normierte Vektorräume sind insbesondere **vollständige** metrische Räume, und auf ihnen ist jede lineare Abbildung stetig.
- 'Vektorraum' heißt hier immer **reeller** Vektorraum.

Differentialrechnung beschäftigt sich mit der linearen Approximation von Funktionen. In einem allgemeinen metrischen Raum macht das keinen Sinn, weil man keine **lineare Struktur** hat. Den richtigen Rahmen für Linearität bieten **Vektorräume**. Und wenn man außerdem über Konvergenz reden will, braucht man in den Vektorräumen eine Metrik, die sich mit der linearen Struktur verträgt. Das führt zur Klasse der **normierten Vektorräume**, mit denen wir uns jetzt befassen wollen.

Ausblick

Man kann nicht nur in normierten Vektorräumen Differentialrechnung betreiben, sondern auch in Räumen, die sich selber durch lineare Räume approximieren lassen: wie Flächen durch ihre Tangentialräume. Das führt zur Analysis auf sogenannten **differenzierbaren Mannigfaltigkeiten**.



DEFINITION 1.8.1 (NORMIERTER VEKTORRAUM)

Sei V ein Vektorraum.

- ① Eine **Norm** für V ist eine Abbildung $\|\cdot\| : V \rightarrow \mathbb{R}$, $v \mapsto \|v\|$, so dass für alle $u, v \in V$ und $\lambda \in \mathbb{R}$ gilt:
 - i (pos. Definitheit) $\|v\| \geq 0$ und ($\|v\| = 0 \iff v = 0$),
 - ii (abs. Homogenität) $\|\lambda v\| = |\lambda| \|v\|$,
 - iii (Dreiecksungleichung) $\|u + v\| \leq \|u\| + \|v\|$.
- ② Ein **normierter Vektorraum** $(V, \|\cdot\|)$ ist ein Vektorraum V zusammen mit einer Norm $\|\cdot\|$ auf V .

Bemerkung 1.8.2 (Normierte Vektorräume sind metrische Räume.) Durch die Definition $d(u, v) := \|u - v\|$ (Beweis einfach) wird daraus ein metrischer Raum, und alle Begriffe, die wir für metrische Räume erklärt haben, sind auch für normierte Vektorräume erklärt. Wenn wir in normierten Vektorräumen von Konvergenz, Stetigkeit, offenen Mengen etc. sprechen, beziehen wir uns immer auf diese Metrik.

Beispiel 1.8.3 (Standardnorm des \mathbb{R}^n) Auf dem \mathbb{R}^n gibt es unter anderem die **p -Norm**:

$$\|x\|_p := d_p(x, 0) = \begin{cases} \left(\sum_{k=1}^n |x_k|^p \right)^{\frac{1}{p}} & p \in [1, \infty) \\ \max_{i=1}^n |x_i| & p = \infty. \end{cases}$$

und sie induziert die Metrik $d_p(x, y) = \|x - y\|$.

Die $\|\cdot\|_2$ ist die **Standardnorm** auf \mathbb{R}^n .

Standardnorm

Beispiel 1.8.4 Der Vektorraum der beschränkten Funktionen $V = \mathcal{B}(X, \mathbb{R})$ gestattet die Norm $\|f\| = \|f\|_{sup} = \sup_{x \in X} |f(x)|$, die zur Supremumsmetrik führt.

Beispiel 1.8.5 Nicht jede Metrik kommt von einer Norm, schon weil metrische Räume im allgemeinen eben keine Vektorräume sind: Beliebige Teilmengen von normierten Vektorräumen sind als Teilmengen von metrischen Räumen wieder metrische Räume, im Allgemeinen aber nur normierten Vektorräume, wenn sie Unterräume (also linear abgeschlossen) sind.

Auf vielen Vektorräumen kann man keine Norm finden, die die diskreten Metrik induzieren.

Beweis. Sei (E, d) ein metrischer Raum, $E = \mathbb{R}$ ein Vektorraum, und d die diskrete Metrik. Dann gilt

$$\|x\| := d(x, 0) = \begin{cases} 1, & x \neq 0 \\ 0, & x = 0, \end{cases}$$

jedoch gilt (ii) nicht für $x \neq 0 \neq \lambda$, da in diesem Fall $d(\lambda \cdot x, 0) = 1 \neq \lambda = \lambda \cdot d(x, 0)$ gilt. \square

Lemma 1.8.6 (umgekehrte $\triangle \neq$)

In $(V, \|\cdot\|)$ gilt $|\|u\| - \|v\|| \leq \|u - v\|$ für alle $u, v \in V$.

Beweis. Es gilt nach der Dreiecksungleichung

$$\|u\| = \|(u - v) + v\| \leq \|u - v\| + \|v\|,$$

und daher $\|u\| - \|v\| \leq \|u - v\|$. Aus Symmetriegründen ist dann aber auch $\|v\| - \|u\| \leq \|u - v\|$, und daraus folgt die Behauptung. \square

Korollar 1.8.7 (Stetigkeit der Norm)

Die Funktion $\|\cdot\| : V \rightarrow \mathbb{R}$, $x \mapsto \|x\|$ ist LIPSCHITZ-stetig mit LIPSCHITZ-Konstante 1_V .

DEFINITION 1.8.8 (VEKTORRAUM DER LINEAREN ABBILDUNGEN)

Mit $L(V, W)$ bezeichnen wir den **Vektorraum der linearen Abbildungen** $F : V \rightarrow W$ vom Vektorraum V in den Vektorraum W .

SATZ 1.8.1: BESCHRÄNKTHEIT DER LINEARE ABBILDUNG

Seien $(V, \|\cdot\|_V)$ und $(W, \|\cdot\|_W)$ normierte Vektorräume und $F : V \rightarrow W$ linear. Dann ist F genau dann stetig, wenn es ein $C \in \mathbb{R}$ gibt, so dass für alle $v \in V$ gilt

$$\|F(v)\|_W \leq C\|v\|_V.$$

Beweis. Zu (\implies). Wenn F stetig ist, ist es insbesondere in 0 stetig. Also existiert zu $\varepsilon = 1$ ein $\delta > 0$ mit

$$F(U_\delta(0)) \subset U_\varepsilon(F(0)) = U_1(0).$$

Mit anderen Worten:

$$\|v\|_V < \delta \implies \|F(v)\|_W < 1.$$

Dann gilt aber für alle $v \neq 0$

$$1 > \left\| F\left(\underbrace{\frac{\delta}{2} \cdot \frac{v}{\|v\|_V}}_{\|\cdot\| < \delta}\right) \right\|_W \stackrel{\text{(ii)}}{=} \frac{\delta}{2\|v\|_V} \cdot \|F(v)\|_W.$$

Das impliziert auch für $v = 0$

$$\|F(v)\|_W \leq \frac{2}{\delta} \|v\|_V.$$

Also können wir $C = \frac{2}{\delta}$ wählen.

Zu (\Leftarrow). Existiert ein C wie im Satz, so ist für alle $v_1, v_2 \in V$

$$\|F(v_1) - F(v_2)\|_W = \|F(v_1 - v_2)\|_W \leq C\|v_1 - v_2\|_V.$$

Daraus folgt (mit $\delta = \frac{\varepsilon}{C}$) die (gleichmäßige) Stetigkeit von F . \square

Wir verzichten im weiteren auf den Index am Normsymbol.

DEFINITION 1.8.9 (VEKTORRAUM-DIMENSION, ISOMORPHISMUS)

Sie $n \in \mathbb{N}$. Ein (reeller) Vektorraum V heißt **n -dimensional**, wenn es einen Isomorphismus

$$\Phi : \mathbb{R}^n \rightarrow V$$

gibt. Dabei ist ein **Isomorphismus** eine **bijektive** Abbildung Φ , so dass Φ und Φ^{-1} linear sind.

SATZ 1.8.2

Seien V, W normierte Vektorräume und V endlich-dimensional. Dann ist jede lineare Abbildung $F : V \rightarrow W$ stetig.

Beweis. 1. Fall: $V = \mathbb{R}^n$. Zunächst gilt für $x \in \mathbb{R}^n$:

$$x = (x_1, \dots, x_n) = \sum_{j=1}^n x_j e_j,$$

wobei e_j den j -ten Standardbasisvektor der sogenannten Standardbasis des \mathbb{R}^n bezeichnet. Daher ist

$$\|F(x)\| = \left\| F\left(\sum_{j=1}^n x_j e_j\right) \right\| = \left\| \sum_{j=1}^n x_j F(e_j) \right\| \leq \sum_{j=1}^n |x_j| \|F(e_j)\| \leq M \sum_{j=1}^n |x_j|$$

mit $M := \max_{j=1}^n \|F(e_j)\|$.

Es gilt

$$|x_j| \leq \sqrt{\sum_{k=1}^n x_k^2} \Rightarrow \sum_{j=1}^n |x_j| \leq n \sqrt{\sum_{k=1}^n x_k^2},$$

also gilt $\|F(x)\| \leq Mn\|x\|$, und F ist nach Satz 1.8.1 stetig.

2. Fall: V beliebig, endlich-dimensional. Sei $\Phi : \mathbb{R}^n \rightarrow V$ ein Isomorphismus. Dann ist Φ nach dem 1. Fall stetig, und es gibt ein $C \in \mathbb{R}$ mit $\|\Phi(x)\| \leq C\|x\|$ für alle x . Wir zeigen, dass es auch ein $B > 0$ gibt, so dass

$$B\|x\| \leq \|\Phi(x)\| \leq C\|x\| \text{ für alle } x \in \mathbb{R}^n. \quad (6)$$

Die Funktion $\|\Phi\| : \mathbb{R}^n \rightarrow \mathbb{R}$ ist als Komposition stetiger Funktionen stetig und nimmt deshalb auf der kompakten Menge

$$S^{n-1} := \{x \in \mathbb{R}^n \mid \|x\| = 1\}$$

ihr Minimum B in einem Punkt $x^* \in S^{n-1}$ an. Weil $x^* \neq 0$ und Φ ein Isomorphismus ist, ist $\Phi(x^*) \neq 0$ und somit gilt $B := \|\Phi(x^*)\| > 0$.

Für alle $x \neq 0$ gilt

$$\|\Phi(x)\| = \left\| \Phi\left(\|x\| \frac{x}{\|x\|}\right) \right\| = \|x\| \left\| \underbrace{\Phi\left(\frac{x}{\|x\|}\right)}_{\in S^{n-1}} \right\| \geq B\|x\|.$$

und die Ungleichung $\|\Phi(x)\| \geq B\|x\|$ gilt offenbar auch für $x = 0$. Damit ist (6) bewiesen. Es folgt

$$\|\Phi^{-1}(v)\| \leq \frac{1}{B} \|v\| \text{ für alle } v \in V \quad \text{WARUMMM?}.$$

Schließlich ist nach dem 1. Fall die lineare Abbildung $F \circ \Phi : \mathbb{R}^n \rightarrow W$ stetig mit

$$\|F \circ \Phi(x)\| \leq A\|x\| \text{ für alle } x \in \mathbb{R}^n \text{ und ein } A \in (0, \infty).$$

Damit erhalten wir

$$\|F(v)\| = \|F \circ \Phi(\Phi^{-1}(v))\| \leq A\|\Phi^{-1}(v)\| \leq \frac{A}{B}\|v\|. \quad \square$$

Seien V und W normierte Vektorräume und $V \neq \{0\}$ endlich-dimensional. Dann definiert

$$\|F\| := \sup_{v \neq 0} \frac{\|F(v)\|}{\|v\|} \quad \text{für } F \in L(V, W)$$

eine Norm auf dem Vektorraum $L(V, W)$.

Operator-Korollar 1.8.10 (Die Operatornorm auf $L(V, W)$)

Beweis. Nach dem Satz existiert ein $C \in \mathbb{R}$, so dass für alle $v \neq 0$

$$\frac{\|F(v)\|}{\|v\|} \leq C.$$

Daher ist $\|F\| \in \mathbb{R}$. Die Norm-Eigenschaften sind leicht zu verifizieren. \square

Korollar 1.8.11

endlich-dimensionaler Sei V ein endlich-dimensionaler normierter Vektorraum mit zwei Normen $\|\cdot\|_1, \|\cdot\|_2$. Dann gilt

- (1) Es gibt $c, C > 0$, so dass für alle $v \in V$ gilt

$$c\|v\|_1 \leq \|v\|_2 \leq C\|v\|_1.$$

Man sagt: Je zwei Normen sind äquivalent.

- (2) Eine Menge $G \subset V$ ist genau dann bezüglich $\|\cdot\|_1$ offen, wenn sie bezüglich $\|\cdot\|_2$ offen ist. Daher sind auch Begriffe wie Konvergenz, Kompaktheit, Stetigkeit usw. unabhängig von der in V verwendeten Norm.
- (3) $\|\cdot\|_1$ -CAUCHY-Folgen sind auch $\|\cdot\|_2$ -CAUCHY-Folgen. Also ist auch der Begriff CAUCHY-Folge unabhängig von der in V verwendeten Norm.

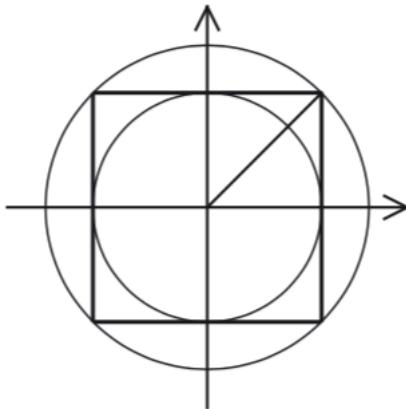


Abb. 23: Äquivalenz der euklidischen und der Supremumsnorm

Beweis. Zu (i). Weil die Identität $\text{id} : (V, \|\cdot\|_2) \rightarrow (V, \|\cdot\|_1)$ linear, also nach Satz 1.8.2 stetig ist, existiert ein nach Satz 1.8.1 ein $M > 0$ mit

$$\|v\|_1 \leq M\|v\|_2 \implies \frac{1}{M}\|v\|_1 \leq \|v\|_2$$

für alle v . Die Stetigkeit von id in der anderen Richtung liefert die zweite Ungleichung.

Zu (ii). Ist G offen bezüglich $\|\cdot\|_1$ und betrachtet man wieder id als stetige Abbildung von $(V, \|\cdot\|_2)$ nach $(V, \|\cdot\|_1)$, so ist auch

$$G = \text{id}^{-1}(G)$$

offen. Die umgekehrte Richtung folgt aus Symmetriegründen.

Zu (iii). Folgt leicht aus (i). \square

Alternativer Beweis. (von ①) Seien $\|\cdot\|_1$ und $\|\cdot\|_2$ zwei Normen auf V und o.B.d.A. $V = \mathbb{R}^n$. Die Einschränkung der Norm Funktion auf die abgeschlossene $\|\cdot\|_2$ -Einheitssphäre

$$\Psi : \overline{B}_{\|\cdot\|_2}(0, 1) \rightarrow \mathbb{R}, v \mapsto \|v\|_1$$

ist eine stetige Funktionen auf einem Kompaktum (nach HEINE-BOREL) und nimmt also ihr Maximum an. Vertauscht man die Rollen der beiden Norm, erhält man die Behauptung.

Alternative kann man so argumentieren: Die Äquivalenz von Normen ist eine Äquivalenzrelation, also genügt es zu zeigen, dass eine beliebige Norm äquivalent zu der $\|\cdot\|_1$ -Norm ist. Sei nun o.B.d.A. $v \in V$ so gewählt, dass $\|v\|_1 = 1$ ist. Da jede Norm $\|\cdot\|$ auf V stetig bezüglich der $\|\cdot\|_1$ -Norm ist, und die Einheitssphäre, auf der v liegt, kompakt ist, also wird eine Minimum und Maximum angenommen.

[Quelle] \square

Beispiel 1.8.12 Für die l^p -Normen auf \mathbb{R}^n aus Beispiel 1.8.3 gilt: Ist $1 \leq p \leq q \leq +\infty$ und $x \in \mathbb{R}^n$, so ist

$$\|x\|_q \leq \|x\|_p \leq n^{\frac{1}{p} - \frac{1}{q}} \|x\|_q, \quad (7)$$

wobei $\frac{1}{+\infty} := 0$.

Beweis. Sei zunächst $q < +\infty$. Die linke Abschätzung ist leicht: Man kann o.E. annehmen, dass

$$1 = (\|x\|_q)^q = \sum_i |x_i|^q.$$

Insbesondere sind dann alle $|x_i| \leq 1$ und daher $|x_i|^p \geq |x_i|^q$. Damit ist $\sum_i |x_i|^p \geq 1$ und

$$\|x\|_p = \left(\sum_i |x_i|^p \right)^{1/p} \geq 1 = \|x\|_q.$$

Die rechte Ungleichung beweisen wir später im Beispiel ??.

Man kann auch die Hölder-Ungleichung verwenden. Dann gilt

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^n |x_i|^1 \right)^{\frac{p}{q}} \left(\sum_{i=1}^n |1|^{\frac{q}{p}} \right)^{\frac{1-\frac{p}{q}}{p}} = \|x\|_q \cdot n^{\frac{1}{p} - \frac{1}{q}}.$$

Für den Fall $q = +\infty$ vergleiche (2). \square

Beispiel 1.8.13 (von mir) Auf endlichdimensionalen Vektorräumen sind nur alle Metriken äquivalent, die von einer Norm kommen, sonst liefert die diskrete Metrik, welche die diskrete Topologie induziert, ein Gegenbeispiel, weil sie z.B. nicht mit der von der euklidischen Metrik induzierten Topologie übereinstimmt.

Beispiel 1.8.14 (von mir) Die Rückrichtung von ① gilt nicht. Zum Beispiel sind $\|(a_n)_{n \in \mathbb{N}}\|_1 := \sum_{n \in \mathbb{N}} |a_n|$ und $\|(a_n)_{n \in \mathbb{N}}\|_2 := \sqrt{\sum_{n \in \mathbb{N}} a_n^2}$ zwei Normen auf ℓ_1 , welche nicht äquivalent sind.

Sei X ein beliebiger unendlichdimensionaler Vektorraum mit der Basis $\{e_i\}_{i \in I}$. Jeder Vektor $x \in X$ lässt sich schreiben als $x = \sum_{i \in I} \alpha_i e_i$, wobei fast alle α_i Null sind. Nun sind $\|x\|_1 := \sum_{i \in I} |c_i|$ und $\|x\|_\infty := \max_{i \in I} |c_i|$ zwei Normen auf X , welche nicht

äquivalent sind, da für jedes $n \in \mathbb{N}$ ein $x \in X$ existiert, sodass $\|x\|_1 > n\|x\|_\infty$ gilt, z.B. $\hat{x} := \sum_{i=1}^{n+1} e_i$ mit $\|\hat{x}\|_1 = N + 1$ aber $\|\hat{x}\|_\infty = 1$.

DEFINITION 1.8.15

BANACH-Raum Ein vollständiger normierter Vektorraum heißt ein **BANACH-Raum**.

SATZ 1.8.3

Jeder **endlich-dimensionale** normierte Vektorraum $(V, \|\cdot\|)$ ist ein **BANACH-Raum**.

Beweis. 1. Fall: $V = \mathbb{R}^n$ mit der Standardnorm. Das haben wir bereits im Beispiel 1.3.11 gezeigt.

2. Fall: V beliebiger endlich-dimensionaler \mathbb{R} -Vektorraum. Sie $(v_k)_{k \in \mathbb{N}}$ eine CAUCHY-Folge in V . Nach unserer Definition (oder nach Linearer Algebra) existiert einen Isomorphismus $\Phi : \mathbb{R}^n \rightarrow V$. Dann ist auch $\Phi^{-1} : V \rightarrow \mathbb{R}^n$ linear, also stetig, und es gibt ein $C \in \mathbb{R}$ mit

$$\|\Phi^{-1}(v_j) - \Phi^{-1}(v_k)\| = \|\Phi^{-1}(v_j - v_k)\| \leq C\|v_j - v_k\|$$

für alle $j, k \in \mathbb{N}$. Also ist auch $(x_k = \Phi^{-1}(v_k))_{k \in \mathbb{N}}$ eine CAUCHY-Folge in \mathbb{R}^n . Sie ist nach dem 1. Fall konvergent gegen ein $x^* \in \mathbb{R}^n$. Wegen der Stetigkeit von Φ ist deshalb

$$\lim_{k \rightarrow \infty} v_k = \lim_{k \rightarrow \infty} \Phi(x_k) = \Phi(x^*).$$

□

Beispiel 1.8.16 Sei V ein n -dimensionaler Vektorraum, so ist $L(V, V)$ zu $\mathbb{R}^{n \times n}$ isomorph, wie man durch geeignete Isomorphismen $L(V, V) \rightarrow \mathbb{R}^{n^2}$ und $\mathbb{R}^{n^2} \rightarrow \mathbb{R}^{n \times n}$.

Die Abbildung $GL(n, \mathbb{R}) \rightarrow GL(n, \mathbb{R})$, $A \mapsto A^{-1}$ ist stetig, weil z.B. der GAUSS-Algorithmus nur aus stetigen Operationen besteht

Wir halten noch einmal das Ergebnis aus dem Korollar 1.8.11 fest:

Vergleiche Beispiele 1.6.6 und 1.6.11. Sei $(V, \|\cdot\|)$ ein n -dimensionaler BANACH-Raum. Die Wahl einer Basis von V liefert nach linearer Algebra einen Isomorphismus

$$\Phi : L(V, V) \rightarrow M(n \times n, \mathbb{R})$$

zwischen dem Raum der linearen Abbildungen von V in sich und dem Raum der $(n \times n)$ -Matrizen. Wenn wir $L(V, V)$ mit der Operatornorm und $M(n \times n, \mathbb{R})$ zum Beispiel mit dem $\mathbb{R}^{(n^2)}$ identifizieren und mit der entsprechenden Norm ausstatten, ist Φ nach Satz 1.8.2 ein Homöomorphismus. Die Determinanten ist nach Beispiel 1.6.6 stetig auf $M(n \times n, \mathbb{R})$, und weil die Determinante der linearen Abbildung $F \in L(V, V)$ nach linearer Algebra gerade die Determinante der Matrix $\Phi(F)$ ist, ist auch die Determinantenfunktion auf $L(V, V)$ stetig. Damit ist das Urbild von $\mathbb{R} \setminus \{0\}$, also die invertierbaren Endomorphismen von V , eine offene Teilmenge $GL(V)$, die unter Φ der Menge der invertierbaren Matrizen $GL(n, \mathbb{R})$ entspricht. Für invertierbare Matrizen sind die Komponenten der Inversen durch gebrochen-rationale Funktionen der originalen Komponenten gegeben, also insbesondere stetig. Daher ist die Inversenbildung auf $GL(n, \mathbb{R})$ und wegen der Φ -Invarianz auch auf $GL(V)$ eine stetige Abbildung.

Ein (abstrakter) endlich-dimensionaler \mathbb{R} -Vektorraum hat unendlich viele Basen, aber keine von diesen ist besonders ausgezeichnet. Ebenso besitzt er unendlich viele Normen, aber keine von diesen ist besonders ausgezeichnet. Allerdings sind sie alle äquivalent: Die durch sie definierten Metriken liefern alle dieselben offenen Mengen, dieselben konvergenten Folgen, dieselben stetigen Abbildungen. Um über Offenheit, Konvergenz oder Stetigkeit in endlich-dimensionalen \mathbb{R} -Vektorräumen zu sprechen, kann man eine beliebige Norm wählen. Weil es aber egal ist, welche man wählt, kann man eben unabhängig von einer solchen Wahl über Offenheit, Konvergenz oder Stetigkeit reden.

Der \mathbb{R}^n besitzt eine Standardbasis und eine Standardnorm, die die Standardmetrik d^2 liefert. Natürlich kann man davon Gebrauch machen, oft muss man aber nicht ...

AUFGABEN

(1)

(2)

LÖSUNGEN

(1)

(2)

2 Grundlagen der mehrdimensionalen Differentiation

Wir werden die Differentialrechnung in endlich-dimensionalen BANACH-Räumen entwickeln. Nach dem vorangehenden Abschnitt sind diese isomorph zu einem \mathbb{R}^n , und man könnte sich auch auf die letzteren beschränken.

Der Vorteil wäre, dass man im \mathbb{R}^n eine ausgezeichnete Basis und damit ausgezeichnete Koordinaten hat. Dadurch wird die Theorie konkreter. Man könnte die Differentialrechnung auf dem Begriff der partiellen Ableitung, also der Ableitung nach einer einzelnen Variablen, aufbauen.

Der Nachteil wäre, dass man im \mathbb{R}^n eine ausgezeichnete Basis und damit ausgezeichnete Koordinaten hat. Diese verschleiern die Tatsache, dass die Konzepte der Differentialrechnung geometrischer Natur sind und mit speziellen Koordinaten nichts zu tun haben, vielleicht aber sehr viel mit anderen Strukturen, die auf dem \mathbb{R}^n auch noch so selbstverständlich vorkommen, dass wir sie gar nicht bemerken.

Zum Beispiel ist $V := \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}$ ein zweidimensionaler Untervektorraum des \mathbb{R}^3 . Hat man darauf eine differenzierbare Funktion $f : V \rightarrow \mathbb{R}$ gegeben, so ist es zunächst unklar, was ihre partiellen Ableitungen sein sollen. Erst wenn man in V eine Basis gewählt hat und damit eine Isomorphie von V auf \mathbb{R}^2 , macht der Begriff der partiellen Ableitungen von f einen Sinn. Allerdings für jede Basiswahl einen anderen. Und es gibt keine "kanonische" Weise, eine Basis zu wählen. Hingegen **Gradiente kann** man den viel wichtigeren Begriff des **Gradienten** ohne partielle Ableitungen definieren, braucht dafür aber ein Skalarprodukt. Und das Skalarprodukt des \mathbb{R}^3 liefert auf ganz kanonische Weise eines für den Untervektorraum V .

2.1 Die Ableitung

- Wir lernen die Ableitung als lineare Approximation einer Abbildung in der Nähe eines Punktes kennen.
- Wir berechne die Ableitung in einfachsten Fällen.

Seien V, W endlich-dimensionale BANACH-Räume⁷ und G eine offene Teilmenge von V .

DEFINITION 2.1.1

Sei $f : V \supset G \rightarrow W$ eine Abbildung der offenen Menge G .

- (1) f heißt differenzierbar in $p \in G$, wenn es eine lineare Abbildung $F : V \rightarrow W$ gibt, so dass für die durch

$$f(x) = f(p) + F(x - p) + R(x) \quad (8)$$

definierte "Restfunktion" $R : G \rightarrow W$ gilt

$$\lim_{x \rightarrow p} \frac{R(x)}{\|x - p\|} = 0. \quad (9)$$

F ist dann eindeutig bestimmt, vgl. Lemma 2.1.3, und wir nennen es die Ableitung oder das Differential von f in p und schreiben $F = D_p f$. Ableitung

- (2) f heißt differenzierbar (auf G), falls f in allen $p \in G$ differenzierbar differenzierbar

Bemerkung 2.1.2 Analytisch gesprochen ist $D_p f$ die lineare Approximation von f in der Nähe von p . Schreibt man x statt p und Δx für $x - p$, so erhält man

$$\Delta f := f(x + \Delta x) - f(x) \approx D_x f(\Delta x).$$

oder (König)

$$f(x + \Delta x) = f(x) + D_x f(\Delta x) + o(\|\Delta x\|) \quad \text{für } \Delta x \rightarrow 0$$

Lemma 2.1.3 (Eindeutigkeit des Differentials D_p)

Ist f in p differenzierbar und F wie in der Definition, so gilt für alle v in V :

$$F(v) = D_p f(v) = \lim_{t \rightarrow 0} \frac{f(p + tv) - f(p)}{t},$$

welches die Richtungsableitung von f in p in Richtung v darstellt.

Richtungsableitung

Beachten Sie: Weil der Definitionsbereich G von f offen ist, liegt für jedes $v \in V$ und hinreichend kleines $|t|$ der Punkt $p + tv$ in G . Deshalb ist der Limes sinnvoll. Die Definition der Differenzierbarkeit kann man auch für Abbildungen von nicht-offenen Teilmengen hinschreiben, aber die Ableitung ist dann im allgemeinen nicht mehr eindeutig.

Beweis. Ist F wie in der Definition, so folgt für alle $v \in V$ und $t \in \mathbb{R} \setminus \{0\}$

$$\frac{f(p + tv) - f(p)}{t} = \frac{f(p) + F(tv) + R(p + tv) - f(p)}{t} \stackrel{L}{=} F(v) + \frac{R(p + tv)}{t}.$$

⁷Im folgenden genügt es, wenn V endlich-dimensional ist. Aber da wir keine konkreten Anwendungen für unendlich-dimensionales W im Sinn haben, sei der Einfachheit halber auch W endlich-dimensional.

Aber es gilt

$$\frac{R(p+tv)}{t} = \underbrace{\frac{R(p+tv)}{\|p+tv-p\|}}_{\rightarrow 0} \underbrace{\frac{|t| \|v\|}{t}}_{=\pm \|v\|} \xrightarrow{t \rightarrow 0} 0. \quad \square$$

Differenzierbarkeit und das Differential hängen wegen Korollar 1.8.11 nicht ab von den gewählten Normen auf V und W . Wir werden deshalb die Norm oft auch nicht spezifizieren.

Beispiel 2.1.4 (Zusammenhang mit der Differenzierbarkeit im Fall $\mathbb{R} \rightarrow \mathbb{R}$)

Die einzigen linearen Abbildungen von \mathbb{R} nach \mathbb{R} sind die Abbildungen $x \mapsto ax$ mit einem festen $a \in \mathbb{R}$. Eine Abbildung $f : \mathbb{R} \supset G \rightarrow \mathbb{R}$ ist deshalb differenzierbar im Sinne der Analysis I genau dann, wenn sie auch nach der neuen Definition differenzierbar ist. Dann gilt für $p \in G$ und $v \in \mathbb{R}$

$$D_p f(v) = f'(p)v,$$

d.h.

$$\boxed{f'(p) = D_p f(1)} \text{ oder: } \boxed{\text{Neue Ableitung} = \text{Multiplikation mit der alten Ableitung.}} \quad (10)$$

Die lineare Abbildung, mit der eine (Ana I)-differenzierbare Funktion $f : (p - \varepsilon, p + \varepsilon) \rightarrow \mathbb{R}$ im Sinne von Definition 2.1.1 approximiert wird, ist die Abbildung $D_p f : x \mapsto f'(p) \cdot x$.

[BILD!]

Beispiel 2.1.5 Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto 1 + 3x + 4y + 5xy^2$.

① Dann ist f in $(0, 0)$ differenzierbar mit $D_{(0,0)} f(u, v) = 3u + 4v$. Es ist nämlich

$$f(x, y) = \underbrace{1}_{f((0,0))} + \underbrace{3(x-0) + 4(y-0)}_{F((x,y)-(0,0))} + \underbrace{5xy^2}_{R((x,y))},$$

und weil $5xy^2$ in $(x-0)$ und $(y-0)$ "kubisch ist", geht der Rest für $(x, y) \rightarrow (0, 0)$ gegen null:

$$\left| \frac{5xy^2}{\sqrt{x^2 + y^2}} \right| = 5 \left| \frac{x}{\sqrt{x^2 + y^2}} y^2 \right| \leq 5y^2.$$

② Durch Nachrechnen können Sie bestätigen, dass

$$f(x, y) = 32 + \underbrace{23(x-1) + 24(y-2)}_{F((x,y)-(1,2))} + \underbrace{20(x-1)(y-2) + 5(y-2)^2 + 5(x-1)(y-2)^2}_{=: R(x,y)}.$$

Es gilt

$$\begin{aligned} \frac{|R(x)|}{\|(x, y) - (1, 2)\|} &= \frac{|20(x-1)(y-2) + 5(y-2)^2 + 5(x-1)(y-2)^2|}{\sqrt{(x-1)^2 + (y-2)^2}} \\ &\leq 20|y-2| + 5|y-2| + 5|y-2||x-1| \xrightarrow{(x,y) \rightarrow (1,2)} 0 \end{aligned}$$

Damit ist f auch in $(x, y) = (1, 2)$ differenzierbar und

$$D_{(1,2)} f(u, v) = 23u + 24v.$$

(Die Umrechnung von f auf den Punkt $(x, y) = (1, 2)$ geschieht erst für x und dann für y mit der TAYLOR-Formel aus Analysis I. Vgl. auch Satz 2.5.2).

[BILDER AUS GEOGEBRA!]

Bemerkung (Geometrische Interpretation) Die geometrische Interpretation ist am einfachsten im Fall $f : \mathbb{R}^2 \supset G \rightarrow \mathbb{R}$. Dann ist der Graph von $x \mapsto f(p) + D_p f(x - p)$ eine Ebene im \mathbb{R}^3 , die **Tangentialebene** an den Graphen von f im Punkt $(p, f(p))$ **Tangentialebene**. Das gilt natürlich auch für Funktionen $\mathbb{R} \rightarrow \mathbb{R}$ und kommt aus dem Satz von TAYLOR.

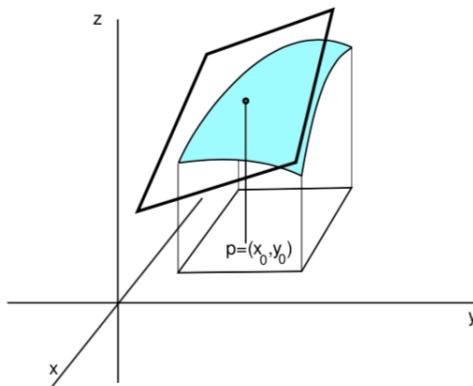


Abb. 24: Geometrische Interpretation [Quelle: Ferus]

Berechnung von Ableitungen

Beispiel 2.1.6 (Konstante Abbildungen) Sei $f : V \supset G \rightarrow W$ konstant. Dann ist $D_p f \equiv 0$ für alle $p \in G$, da $f(x) = f(p) + 0 + 0$ für alle $x, p \in G$ gilt.

Beispiel 2.1.7 (Lineare Abbildungen sind ihre eigenen Ableitungen) Sei $f : V \rightarrow W$ linear. Dann ist

$$f(x) = f(p) + f(x - p) + 0.$$

Also ist f in jedem Punkt $p \in V$ differenzierbar und $D_p f = f$. Zum Beispiel ist die Additionsabbildung

$$\alpha : V \times V \rightarrow V, (x, y) \mapsto \alpha(x, y) = x + y$$

vom Vektorraum $V \times V$ der Paare in den Vektorraum V linear. Also ist α differenzierbar, und für alle (x, y) und (u, v) in $V \times V$ ist

$$D_{(x,y)} \alpha(u, v) = u + v.$$

Beispiel 2.1.8 (Skalarmultiplikation) Die nichtlineare sondern bilineare Abbildung der Skalarmultiplikation

$$\mu : \mathbb{R} \times V \rightarrow V, (\lambda, x) \mapsto \lambda x$$

ist differenzierbar in jedem $(\lambda_0, x_0) \in \mathbb{R} \times V$. Es ist nämlich

$$\mu(\lambda, x) = \lambda x = \underbrace{\lambda_0 x_0}_{\mu(\lambda_0, x_0)} + \underbrace{(\lambda - \lambda_0)x_0}_{=: F(\lambda - \lambda_0, x - x_0)} + \underbrace{\lambda_0(x - x_0)}_{=: R(\lambda, x)} + \underbrace{(\lambda - \lambda_0)(x - x_0)}_{=: R(\lambda, x)}.$$

Diese Gleichung rechnet man leicht nach. Es bleibt zu zeigen, dass

$$\lim_{(\lambda, x) \rightarrow (\lambda_0, x_0)} \frac{R(\lambda, x)}{\|(\lambda, x) - (\lambda_0, x_0)\|} = 0.$$

Dazu braucht man eine Norm auf $\mathbb{R} \times V$. Wir nehmen an, dass auf V eine Norm $\|\cdot\|$ gegeben ist, und definieren

$$\|(\lambda, x)\| := |\lambda| + \|x\|.$$

Dass das wirklich eine Norm definiert ist klar, da $\|\cdot\|$ eine Norm auf V und $|\cdot|$ eine Norm auf \mathbb{R} ist. Damit gilt dann:

$$\frac{|R(\lambda, x)|}{\|(\lambda, x) - (\lambda_0, x_0)\|} = \frac{|\lambda - \lambda_0| \|x - x_0\|}{|\lambda - \lambda_0| + \|x - x_0\|} \leq \|x - x_0\| \xrightarrow{(\lambda, x) \rightarrow (\lambda_0, x_0)} 0$$

Daraus folgt die Behauptung. Wir halten fest:

$$D_{(\lambda_0, x_0)} \mu(\lambda, x) = \lambda_0 x + \lambda x_0.$$

Das ist eine Art Produktregel, auf die wir noch zurückkommen.

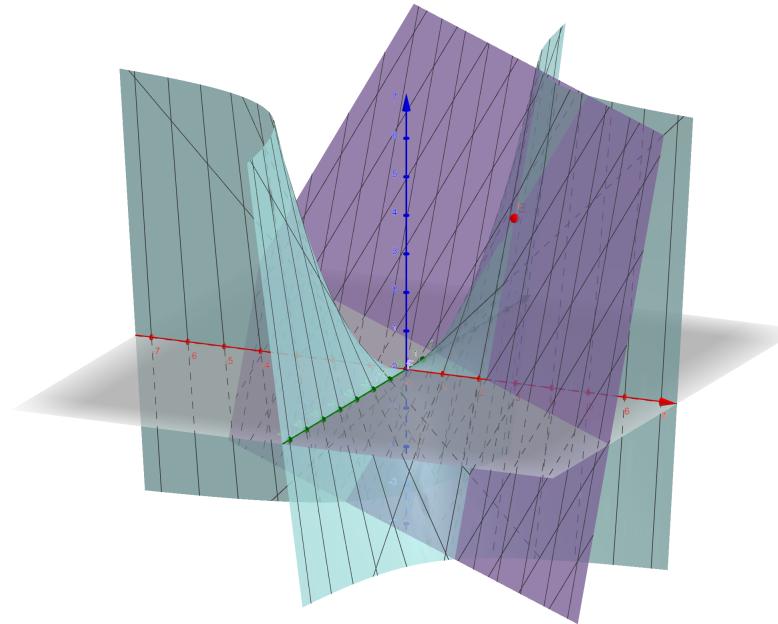


Abb. 25: Der Graph der Funktion $f(x, y) := xy$ und die Tangentialebene an dem Punkt $E = (2.2, 1.7)$, welche durch $f(a, b) + (x - a)b + (y - b)a$ gegeben ist [Quelle: GeoGebra6]

Lemma 2.1.9 (Komponentenweise Differentiation)

Sei

$$f = (f_1, \dots, f_m) : V \supset G \rightarrow \mathbb{R}^m$$

mit Komponentenfunktionen $f_i : G \rightarrow \mathbb{R}$. Dann ist f genau dann in p differenzierbar, wenn alle f_i in p differenzierbar sind. In diesem Fall gilt

$$D_p f(v) = (D_p f_1(v), \dots, D_p f_m(v)) \text{ für alle } v \in V. \quad (11)$$

Beweis. In Komponenten sieht die Gleichung (8) so aus:

$$f_i(x) = f_i(p) + F_i(x - p) + R_i(x).$$

Nun ist F linear genau dann, wenn alle Komponenten F_i linear sind. Und weil Konvergenz im \mathbb{R}^n einfach komponentenweise Konvergenz ist, folgt die Behauptung durch Betrachtung der Komponenten $\frac{R_i(x)}{\|x - p\|}$ von $\frac{R(x)}{\|x - p\|}$. \square

Dieses Beispiel gestattet eine Verallgemeinerung auf folgende Situation

Bemerkung Seien V und W_1, \dots, W_m endlich-dimensionale BANACH-Räume, $G \subset V$ offen und seien

$$f_i : V \supset G \rightarrow W_i$$

für $i \in \{1, \dots, m\}$ Abbildungen. Wir definieren

$$\begin{aligned} f : V \supset G &\rightarrow W_1 \times \dots \times W_m \\ x &\mapsto (f_1(x), \dots, f_m(x)). \end{aligned}$$

Dann ist f genau dann differenzierbar in p , wenn alle f_i in p differenzierbar sind, und es gilt wieder die Gleichung (11).

DEFINITION 2.1.10 (FUNKTIONALMATRIX)

Sei $f : \mathbb{R}^n \supset G \rightarrow \mathbb{R}^m$ eine differenzierbar Abbildung, G eine offene Teilmenge und in $p \in G$. Dann ist $D_p f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine lineare Abbildung, und eine solche wird nach Linearer Algebra dargestellt durch eine Matrix, die wir mit $f'(p)$ bezeichnen und die **Jacobimatrix** oder **Funktionalmatrix** von f in p nennen. Die Spalten sind gerade die Bilder der Basisvektoren e_1, \dots, e_n :

$$f'(p) = (D_p f(e_1) \dots D_p f(e_n)) = (D_p f_i(e_j)) = \begin{pmatrix} D_p f_1(e_1) & \dots & D_p f_1(e_n) \\ \vdots & & \vdots \\ D_p f_m(e_1) & \dots & D_p f_m(e_n) \end{pmatrix}.$$

Die Formel im Lemma 2.1.3 liefert eine Möglichkeit, die $D_p f_i(e_j)$ zu berechnen. Wir kommen im Abschnitt 2.3 darauf zurück.

Beispiel 2.1.11 (Kurven & Schraubenlinie) Eine Abbildung $f : \mathbb{R} \supset (a, b) \rightarrow W$ nennt man eine Kurve in W . Ist f in $t \in (a, b)$ differenzierbar, so ist für alle $\lambda \in \mathbb{R}$

$$D_t f(\lambda) = \lambda D_t f(1),$$

Also ist $D_t f : \mathbb{R} \rightarrow W$ durch den **Tangentialvektor** $\dot{f}(t) := D_t f(1)$ eindeutig bestimmt.

[BILD!]

Ist $W = \mathbb{R}^m$ und $f = (f_1, \dots, f_m)$, so ist

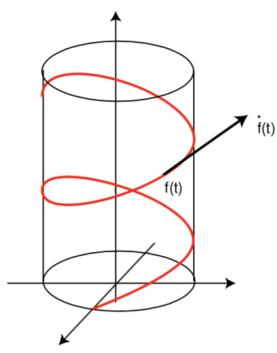
$$\dot{f}(t) = (\dot{f}_1(t), \dots, \dot{f}_m(t)).$$

Dabei ist nach Lemma 2.1.3

$$\dot{f}_i(t) = D_{t,i} f_i(1) = \lim_{\tau \rightarrow 0} \frac{f_i(\tau + t) - f_i(t)}{\tau}.$$

$\dot{f}_i(t)$ ist also die gewöhnliche Ableitung der Analysis I. Insbesondere können wir die Definition von \dot{f} auch auf den Fall kompakter Intervalle $[a, b]$ ausdehnen.

Konkret: Die Kurve $f : \mathbb{R} \rightarrow \mathbb{R}^3$, $t \mapsto (\cos t, \sin t, t)$ ist eine Spiralkurve. Sie hat den Geschwindigkeitsvektor $\dot{f}(t) = (-\sin t, \cos t, 1)$ mit $\|\dot{f}(t)\| = \sqrt{2}$. Und es gilt zum Beispiel $D_2 f(-5) = (5 \sin 2, -5 \cos 2, -5)$.



2.2 Rechenregeln für differenzierbare Abbildungen

- Differenzierbare Abbildungen sind stetig.
- Kettenregel: $D_p(f \circ g)(u) = D_qf(D_pg(u))$.
- Produktregel: $D_{(p_1, \dots, p_k)}\mu(v_1, \dots, v_k) = \sum_{i=1}^k \mu(p_1, \dots, p_{i-1}, v_i, p_{i+1}, \dots, p_k)$.
- $D_p\|x\| = \frac{\langle x, p \rangle}{\|x\|}$.
- $D_A \det(B) = \det(A) \cdot \text{Spur}(B \cdot A^{-1})$
- $D_A \text{inv}(B) = -A^{-1} \cdot B \cdot A^{-1}$.

SATZ 2.2.1: DIFFERENZIERBARE ABBILDUNGEN SIND STETIG.

Seien V, W endlich-dimensionale BANACH-Räume. Ist $f : V \supset G \rightarrow W$ in $p \in G$ differenzierbar, so ist es dort auch stetig.

Beweis. Die Behauptung folgt aus

$$f(x) = f(p) + D_p f(x - p) + R(x),$$

weil für $x \rightarrow p$ das Restglied gegen 0 geht, und weil die lineare Abbildung $D_p f$ auf einem endlich-dimensionalen BANACH-Raum stetig ist. Dann gilt

$$f(x) - f(p) = D_p f(x - p) + R(x) \xrightarrow{x \rightarrow p} D_p f(0) + 0 = 0. \quad \square$$

SATZ 2.2.2: KETTENREGEL

Seien U, V, W endlich-dimensionale BANACH-Räume, $G \subset U$ und $H \subset V$ offen, $g : G \rightarrow V$ und $f : H \rightarrow W$ Abbildungen mit $g(G) \subset H$. Sei g differenzierbar in $p \in G$ und f differenzierbar in $q = g(p) \in H$. Dann ist die Abbildung $f \circ g : G \rightarrow W$ differenzierbar in p , und es gilt:

$$D_p(f \circ g) = D_{g(p)}f \circ D_pg \quad \text{bzw.} \quad D_p(f \circ g)(u) = D_qf(D_pg(u)) \quad \forall u \in U.$$

Beweis. Die definierenden Gleichungen

$$\begin{aligned} g(x) &= g(p) + D_pg(x - p) + R(x) \\ f(y) &= f(q) + D_qf(y - q) + S(y) \end{aligned}$$

implizieren ($y = g(x)$, $q = g(p)$) aufgrund der Linearität der Abbildung $D_{(\cdot)}$

$$\begin{aligned} f(g(x)) &= f(g(p)) + D_qf(g(x) - g(p)) + S(g(x)) \\ &= f(g(p)) + D_qf(D_pg(x - p)) + \underbrace{D_qf(R(x)) + S(g(x))}_{=: T(x)}. \end{aligned}$$

Mit Hilfe der im Korollar 1.8.10 definierte Operatornorm gilt

$$\begin{aligned} \frac{T(x)}{\|x - p\|} &\stackrel{\triangle}{=} \frac{\|D_qf(R(x))\|}{\|x - p\|} + \frac{\|S(g(x))\|}{\|g(x) - g(p)\|} \cdot \frac{\|g(x) - g(p)\|}{\|x - p\|} \\ &\leq \underbrace{\|D_qf\|_{\text{op}}}_{<\infty} \underbrace{\frac{\|R(x)\|}{\|x - p\|}}_{\xrightarrow{x \rightarrow p} 0} + \underbrace{\frac{\|S(g(x))\|}{\|g(x) - g(p)\|}}_{\xrightarrow{x \rightarrow p} 0, \text{ da } g \text{ stetig}} \cdot \frac{\|D_pg(x - p) + R(x)\|}{\|x - p\|}. \end{aligned}$$

Außerdem gilt analog

$$\frac{\|D_p g(x-p) + R(x)\|}{\|x-p\|} \leq \|D_p g\|_{\text{op}} + \underbrace{\frac{R(x)}{\|x-p\|}}_{\rightarrow 0}.$$

Da die Vektorräume endlichdimensional sind, ist die Operatornormen der Abbildungen zwischen ihnen immer beschränkt. \square

Der folgende Satz verallgemeinert das Beispiel 2.1.8 auf multilinear Abbildungen.

SATZ 2.2.3: PRODUKTREGEL

Seien V_1, \dots, V_k, W endlich-dimensionale BANACH-Räume und

$$\mu : V_1 \times \dots \times V_k \rightarrow W$$

k -lineare eine **k -lineare** Abbildung, d.h. μ ist in jedem seiner k Argumente linear.

Dann ist μ differenzierbar und es gilt

$$D_{(p_1, \dots, p_k)} \mu(v_1, \dots, v_k) = \sum_{i=1}^k \mu(p_1, \dots, p_{i-1}, v_i, p_{i+1}, \dots, p_k).$$

Bemerkung: Der erste Summand ist zu interpretieren als $\mu(v_1, p_2, \dots, p_k)$, der letzte entsprechend.

Beweis. ① Stetigkeit von μ . Wir zeigen zunächst, dass für alle $x_i \in V_i$ ein C existiert mit

$$\|\mu(x^{(1)}, \dots, x^{(k)})\| \leq C \|x^{(1)}\| \cdot \dots \cdot \|x^{(k)}\|. \quad (12)$$

Ist (e_1, \dots, e_n) eine Basis von V , so sind die **Koordinatenabbildungen** $\Phi : x = \sum_{j=1}^{\dim(V)} x_j e_j \mapsto x_i$ linear, also stetig. Anwendung auf $V = V_i$ ergeben (Satz 1.8.1) Konstanten $C_{1,j_1}, \dots, C_{k,j_k}$ mit $|\Phi(x)| = |x_j^{(i)}| \leq C_{i,j} \|x^{(i)}\|$ für alle $x^{(i)} \in V_i$ und alle $j \in \{1, \dots, \dim(V_i)\}$, wobei $x_j^{(i)}$ die j -te Koordinate von x bezüglich einer Basis $(e_j^{(i)})_{j \in \{1, \dots, \dim(V_i)\}}$ ist.

Wir wählen nun Basen e_{i1}, \dots, e_{in_i} für V_i und schreiben $x_i = \sum_j x_{ij} e_{ij} \in V_i$.

Aus der Multilinearität von μ folgt

$$\begin{aligned} \|\mu(x^{(1)}, \dots, x^{(k)})\| &= \left\| \mu \left(\sum_{j_1} e_{j_1}^{(1)} x_{j_1}^{(1)}, \dots, \sum_{j_k} e_{j_k}^{(k)} x_{j_k}^{(k)} \right) \right\| \\ &= \left\| \sum_{j_1, \dots, j_k} x_{j_1}^{(1)} \cdot \dots \cdot x_{j_k}^{(k)} \cdot \mu \left(e_{j_1}^{(1)}, \dots, e_{j_k}^{(k)} \right) \right\| \\ &\leq \sum_{j_1, \dots, j_k} |x_{j_1}^{(1)}| \cdot \dots \cdot |x_{j_k}^{(k)}| \cdot \left\| \mu \left(e_{j_1}^{(1)}, \dots, e_{j_k}^{(k)} \right) \right\| \\ &\leq \|x^{(1)}\| \cdot \dots \cdot \|x^{(k)}\| \underbrace{\sum_{j_1, \dots, j_k} C_{1,j_1} \cdot \dots \cdot C_{k,j_k}}_{=:C} \cdot \left\| \mu \left(e_{j_1}^{(1)}, \dots, e_{j_k}^{(k)} \right) \right\|. \end{aligned}$$

② Nun zum eigentlichen Beweis. Dazu müssen wir den Restterm

$$R(x_1, \dots, x_k) = \mu(x_1, \dots, x_k) - \mu(p_1, \dots, p_k) - \sum_{i=1}^k \mu(p_1, \dots, p_{i-1}, x_i - p_i, p_{i+1}, \dots, p_k)$$

berechnen. Dann müssen wir eine Norm $\|\cdot\|$ auf $V_1 \times \dots \times V_k$ wählen und zeigen, dass

$$\lim_{(x_1, \dots, x_k) \rightarrow (p_1, \dots, p_k)} \frac{R(x_1, \dots, x_k)}{\|(x_1, \dots, x_k) - (p_1, \dots, p_k)\|} = 0.$$

Eigentlich müssen wir den Restterm natürlich gar nicht berechnen, sondern wir müssen ihn in einer Form schreiben, die es ermöglicht, den Grenzwert zu berechnen. Dazu führen wir folgende Schreibweise ein:

$$\sum_{1 \leq j_1 < \dots < j_m \leq k} \mu_{j_1 \dots j_m}(p, x) := \sum_{1 \leq j_1 < \dots < j_m \leq k} \mu(p_1, \dots, x_{j_1} - p_{j_1}, \dots, x_{j_m} - p_{j_m}, \dots, p_k), \quad (13)$$

wobei über alle Produkte summiert wird, die aus $\mu(p_1, \dots, p_k)$ entstehen, indem man die p_{j_i} ersetzt durch $x_{j_i} - p_{j_i}$. Der Restterm ist dann also

$$R(x_1, \dots, x_k) = \mu(x_1, \dots, x_k) - \mu(p_1, \dots, p_k) - \sum_{1 \leq j_1 \leq k} \mu_{j_1}(p, x).$$

Wir zeigen gleich in einem Lemma, dass dann

$$R(x_1, \dots, x_k) = \sum_{m=2}^k \sum_{1 \leq j_1 < \dots < j_m \leq k} \mu_{j_1 \dots j_m}(p, x). \quad (14)$$

Jeder Summand der rechten Seite enthält also mindestens zwei ‘‘Faktoren’’ der Form $(x_j - p_j)$ und geht deshalb für $(x_1, \dots, x_k) \rightarrow (p_1, \dots, p_k)$ mindestens quadratisch gegen 0.

Genauer: Ist $\|\cdot\|_i$ eine Norm auf V_i , so definiert

$$\|(x_1, \dots, x_k)\| := \|x_1\|_1 + \dots + \|x_k\|_k$$

eine Norm auf $V_1 \times \dots \times V_k$, und weil nach (12)

$$\begin{aligned} & \frac{\|\mu(p_1, \dots, x_{j_1} - p_{j_1}, \dots, x_{j_2} - p_{j_2}, \dots, p_k)\|}{\|(x_1, \dots, x_k) - (p_1, \dots, p_k)\|} \\ & \leq C \|p_1\| \dots \underbrace{\|p_1\| + \dots + \|x_{j_1} - p_{j_1}\| + \dots + \|p_k\|}_{\leq 1} \dots \underbrace{\|x_{j_2} - p_{j_2}\|}_{\rightarrow 0} \dots \|p_k\|, \end{aligned}$$

geht das Restglied gegen 0. □

Die Restgliedformel (14) folgt aus

Lemma 2.2.1

Für jede k -lineare Abbildung $\mu : V_1 \times \dots \times V_k \rightarrow W$ und alle (x_1, \dots, x_k) und $(p_1, \dots, p_k) \in V_1 \times \dots \times V_k$ gilt unter der Verwendung von (13)

$$\mu(x_1, \dots, x_k) = \mu(p_1, \dots, p_k) + \sum_{m=1}^k \sum_{1 \leq j_1 < \dots < j_m \leq k} \mu_{j_1 \dots j_m}(p, x)$$

Die p -Terme auf der rechten Seite heben sich also weg.

Beweis. Wir zeigen das durch vollständige Induktion über k .

$k = 1$. Die Formel

$$\mu(x_1) = \mu(p_1) + \mu(x_1 - p_1)$$

folgt aus der 1-Linearität.

$k \rightarrow k+1$. Sei also $\mu : V_1 \times \dots \times V_{k+1} \rightarrow W$ eine k -lineare Abbildung. Dann ist

$$\begin{aligned} (*) &:= \mu(p_1, \dots, p_{k+1}) + \sum_{m=1}^{k+1} \sum_{1 \leq j_1 < \dots < j_m \leq k+1} \mu_{j_1 \dots j_m}(p, x) \\ &= \mu(p_1, \dots, p_{k+1}) + \sum_{m=1}^{k+1} \sum_{1 \leq j_1 < \dots < j_m \leq k} \mu_{j_1 \dots j_m}(p, x) \\ &\quad + \sum_{m=1}^{k+1} \sum_{1 \leq j_1 < \dots < j_m = k+1} \mu_{j_1 \dots j_m}(p, x). \end{aligned}$$

Im mittleren Term kann $m = k+1$ nicht vorkommen. Deshalb können wir fortfahren

$$\begin{aligned} (*) &= \mu(p_1, \dots, p_{k+1}) + \sum_{m=1}^k \sum_{1 \leq j_1 < \dots < j_m \leq k} \mu_{j_1 \dots j_m}(p, x) \\ &\quad + \mu(p_1, \dots, p_k, x_{k+1} - p_{k+1}) + \sum_{m=2}^{k+1} \sum_{1 \leq j_1 < \dots < j_m = k+1} \mu_{j_1 \dots j_m}(p, x). \end{aligned}$$

Wir definieren nun zwei k -lineare Abbildungen auf $V_1 \times \dots \times V_k$ durch

$$\begin{aligned} \mu^0(x_1, \dots, x_k) &:= \mu(x_1, \dots, x_k, p_{k+1}), \\ \mu^1(x_1, \dots, x_k) &:= \mu(x_1, \dots, x_k, x_{k+1}). \end{aligned}$$

Wir erhalten

$$\begin{aligned} (*) &= \mu^0(p_1, \dots, p_k) + \sum_{m=1}^k \sum_{1 \leq j_1 < \dots < j_m \leq k} \mu_{j_1 \dots j_m}^0(p, x) \\ &\quad + \mu^1(p_1, \dots, p_k) - \mu^0(p_1, \dots, p_k) \\ &\quad + \sum_{m=2}^{k+1} \sum_{1 \leq j_1 < \dots < j_{m-1} \leq k} \mu_{j_1 \dots j_m}^1(p, x) - \sum_{m=2}^{k+1} \sum_{1 \leq j_1 < \dots < j_{m-1} \leq k} \mu_{j_1 \dots j_m}^0(p, x) \\ &= \sum_{m=1}^k \sum_{1 \leq j_1 < \dots < j_m \leq k} \mu_{j_1 \dots j_m}^0(x, p) + \mu^1(p_1, \dots, p_k) \\ &\quad + \sum_{m=1}^k \sum_{1 \leq j_1 < \dots < j_m \leq k} \mu_{j_1 \dots j_m}^1(p, x) - \sum_{m=1}^k \sum_{1 \leq j_1 < \dots < j_m \leq k} \mu_{j_1 \dots j_m}^0(p, x) \\ &= \mu^1(p_1, \dots, p_k) + \sum_{m=1}^k \sum_{1 \leq j_1 < \dots < j_m \leq k} \mu_{j_1 \dots j_m}^1(p, x) \\ &\stackrel{\text{Ind.Vor.}}{=} \mu^1(x_1, \dots, x_k) = \mu(x_1, \dots, x_{k+1}). \end{aligned} \quad \square$$

Die Produktregel aus der Analysis I ist eine Kombination aus der vorstehenden Produktregel mit der Kettenregel. Das erklären wir genauer:

Beispiel 2.2.2 (Alte Produktregel) Seien $J \subset \mathbb{R}$ ein offenes Intervall, $p \in J$ und seien $f, g : J \rightarrow \mathbb{R}$ differenzierbare Funktionen. Sei weiter $\mu : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ die Multiplikationsabbildung $(x, y) \mapsto xy$ und sei

$$(f, g) : J \rightarrow \mathbb{R} \times \mathbb{R}, \quad t \mapsto (f(t), g(t)).$$

Wir betrachten die Komposition

$$h := \mu \circ (f, g) : t \mapsto f(t)g(t).$$

Dann gilt

$$\begin{aligned} h'(p) &\stackrel{(10)}{=} D_p h(1) \\ &= D_{(f(p), g(p))} \mu \circ D_p(f, g)(1) && \text{(Kettenregel)} \\ &\stackrel{(11)}{=} D_{(f(p), g(p))} \mu \circ (D_p f(1), D_p g(1)) \\ &\stackrel{(10)}{=} D_{(f(p), g(p))} \mu \circ (f'(p), g'(p)) \\ &= \mu(f'(p), g(p)) + \mu(f(p), g'(p)) && \text{(Produktregel)} \\ &= f'(p)g(p) + f(p)g'(p). \end{aligned}$$

Beispiel 2.2.3 Hier sind wichtige multilinear Produkte. Überlegen Sie, was in jedem einzelnen Fall die Produktregel besagt.

- ① Das normale Produkt reeller Zahlen hatten wir gerade

$$\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, (x, y) \mapsto xy.$$

- ② Matrix mal Vektor: Sei $M(m \times n, \mathbb{R})$ der Vektorraum der reellen $(m \times n)$ -Matrizen.

$$M(m \times n, \mathbb{R}) \times \mathbb{R}^n \rightarrow \mathbb{R}^m, (A, x) \mapsto Ax.$$

- ③ Allgemeiner

$$L(V, W) \times V \rightarrow W, (f, v) \mapsto f(v)$$

- ④ Das kanonische Skalarprodukt im \mathbb{R}^n

$$\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}, (x, y) \mapsto \langle x, y \rangle = \sum_{i=1}^n x_i y_i.$$

- ⑤ Allgemeiner jedes Skalarprodukt

$$V \times V \rightarrow \mathbb{R}, (x, y) \mapsto \langle x, y \rangle$$

auf einem Euklidischen Vektorraum V , oder noch allgemeiner jede Bilinearform.

- ⑥ Das Kreuzprodukt

$$\mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3, (x, y) \mapsto x \times y$$

auf dem \mathbb{R}^3 .

- ⑦ Die Determinante als Funktion der Spalten:

$$\mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}, (x_1, \dots, x_n) \mapsto \det(x_1, \dots, x_n).$$

Dieses führen wir in einem Beispiel weiter unten aus.

Beispiel 2.2.4 Wir betrachten einen Euklidischen Vektorraum $(V, \langle \cdot, \cdot \rangle)$ und dazu: die Produktabbildung

$$\mu : V \times V \rightarrow \mathbb{R}, (x, y) \mapsto \langle x, y \rangle,$$

die lineare ‘‘Diagonal-Abbildung’’

$$\delta : V \rightarrow V \times V, \quad x \mapsto (x, x),$$

und die ‘‘Norm-Abbildung’’

$$r : V \rightarrow \mathbb{R}, \quad x \mapsto \|x\| = \sqrt{\langle x, x \rangle}$$

Dann ist

$$r = \sqrt{} \circ \mu \circ \delta \quad \text{und} \quad D_p r = D_{\mu(\delta(p))} \sqrt{} \circ D_{\delta(p)} \mu \circ (D_p \delta).$$

Seien $x, p \in V$ und $p \neq 0$. Wir wollen zeigen, dass r in p differenzierbar ist und das Differential ausrechnen.

- δ ist als lineare Abbildung differenzierbar und

$$D_p \delta(x) = \delta(x) = (x, x).$$

- μ ist nach der Produktregel differenzierbar und

$$D_{(p,p)} \mu(x, x) = \mu(x, p) + \mu(p, x) = 2\mu(x, p).$$

- Die Wurzel $\sqrt{}$ ist nach Analysis I differenzierbar und

$$(\sqrt{})'(\tilde{t}) = \frac{1}{2\sqrt{\tilde{t}}}.$$

Das bedeutet

$$D_{\tilde{t}} \sqrt{}(t) = \frac{1}{2\sqrt{\tilde{t}}} t.$$

Nimmt man die vorstehenden Ergebnisse zusammen, so sieht man, dass r in p nach der Kettenregel differenzierbar ist und

$$D_p r(x) = \frac{1}{2\sqrt{\mu(x, p)}} 2\mu(x, p) = \frac{\langle x, p \rangle}{r(p)}. \quad (15)$$

Beispiel 2.2.5 Für Vektoren $x_1, \dots, x_n \in \mathbb{R}^n$ schreiben wir

$$X := (x_1, \dots, x_n)$$

für die Matrix mit den Spalten x_j . Ist (e_1, \dots, e_n) die kanonische Basis des \mathbb{R}^n , so ist also

$$E := (e_1, \dots, e_n)$$

die n -reihige Einheitsmatrix. So identifizieren wir den Raum $\mathbb{R}^n \times \dots \times \mathbb{R}^n$ mit dem Raum $M(n \times n, \mathbb{R})$ der quadratischen n -reihigen Matrizen. Die Determinante – nehmen Sie das als Definition, wenn Sie in der Linearen Algebra noch nicht so weit sind – ist mit dieser Identifikation eine n -lineare Abbildung

$$\det : \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad (x_1, \dots, x_n) \mapsto \det(x_1, \dots, x_n)$$

mit folgenden zusätzlichen Eigenschaften:

$$\det(x_1, \dots, x_n) = 0, \text{ falls zwei der } x_j \text{ gleich sind,} \quad \det(e_1, \dots, e_n) = 1. \quad (16)$$

Damit gilt für $A = (a_1, \dots, a_n) \in M(n \times n, \mathbb{R})$ und entsprechendes B nach der Produktregel

$$\begin{aligned} D_A \det(B) &= \sum_{j=1}^n \det(a_1, \dots, a_{j-1}, b_j, a_{j+1}, \dots, a_n) \\ &= \sum_{j=1}^n \det(a_1, \dots, a_{j-1}, \sum_{k=1}^n b_{kj} e_k, a_{j+1}, \dots, a_n) \\ &= \sum_{k=1}^n \sum_{j=1}^n b_{kj} \underbrace{\det(a_1, \dots, a_{j-1}, e_k, a_{j+1}, \dots, a_n)}_{=: a_{jk}^v} \\ &= \sum_{k=1}^n \sum_{j=1}^n b_{kj} a_{jk}^v. \end{aligned}$$

Definiert man also die **Adjunge** $\text{adj}(A)$ der Matrix A durch

Adjunge

$$\text{adj}(A) := (a_{jk}^v)_{j,k=1,\dots,n} = (\det(a_1, \dots, a_{j-1}, e_k, a_{j+1}, \dots, a_n))_{j,k=1,\dots,n},$$

so ist $D_A \det(B)$ gerade die Summe der Diagonalelemente der Matrix $B \text{adj}(A)$, die sogenannte **Spur** dieser Matrix:

Spur

$$D_A \det(B) = \text{Spur}(B \text{adj}(A)).$$

Wir merken noch an:

(1) Ist $A = E$, so ist nach (??)

$$e_{jk}^v = \det(e_1, \dots, e_{j-1}, e_k, e_{j+1}, \dots, e_n) = \delta_{jk},$$

also $\text{adj}(E) = E$ und

$$D_E \det(B) = \text{Spur}(B). \quad (17)$$

(2) Allgemein gilt nach (??)

$$\begin{aligned} \sum_{k=1}^n a_{ik}^v a_{kj} &= \det(a_1, \dots, a_{i-1}, \sum_{k=1}^n a_{kj} e_k, a_{i+1}, \dots, a_n) \\ &= \det(a_1, \dots, a_{i-1}, a_j, a_{i+1}, \dots, a_n) = \delta_{ij} \det(A), \end{aligned}$$

also $\text{adj}(A)A = \det(A)E$. Ist $\det A \neq 0$, so ist A also invertierbar und

$$\text{adj}(A) = \det(A)A^{-1}.$$

In diesem Fall ist

$$D_A \det(B) = \det(A) \text{Spur}(BA^{-1}). \quad (18)$$

Beispiel 2.2.6 Sei V ein endlich-dimensionaler BANACH-Raum. Wir wollen zeigen

- $\text{GL}(V) := \{A \in L(V, V) \mid A \text{ invertierbar}\}$ ist offen in $L(V, V)$,
- die Abbildung

$$\text{inv} : \text{GL}(V) \rightarrow L(V, V), \quad A \mapsto A^{-1}$$

ist differenzierbar und

- ihre Ableitung ist

$$D_A \text{inv}(B) = -A^{-1}BA^{-1}. \quad (19)$$

Beachten Sie: Für $V = \mathbb{R}$ ist das Ergebnis die aus der Schule bekannte Formel für die Ableitung von $\frac{1}{x}$, es gilt $D_x(x^{-1})(\lambda) = -x^{-1}\lambda x^{-1} = \lambda(-x^{-2})$.

Beweis. Wir benutzen auf $L(V, V)$ die Operatornorm und die Ungleichung

$$\|AB\| \leq \|A\| \|B\| \quad (20)$$

für die Norm der Komposition von A und B (Beweis in HA).

Sei $A \in \text{GL}(V)$ und $B \in L(V, V)$ mit

$$\|A - B\| < \frac{1}{\|A^{-1}\|}. \quad (21)$$

Aus (20) folgt dann also

$$\|A^{-1}(A - B)\| < 1. \quad (22)$$

Nun benutzen wir

$$B = A - (A - B) = A(E - A^{-1}(A - B)),$$

wobei E die identische Abbildung von V ist, und denken an die geometrische Reihe.

Wir definieren

$$S_n := \left(\sum_{k=0}^n (A^{-1}(A - B))^k \right) A^{-1}.$$

Aus

$$\|S_{n+m} - S_n\| = \left\| \left(\sum_{k=n+1}^{n+m} (A^{-1}(A - B))^k \right) A^{-1} \right\| \leq \left(\sum_{k=n+1}^{n+m} \|A^{-1}(A - B)\|^k \right) \|A^{-1}\|$$

folgt mit (22), dass die S_n eine CAUCHY-Folge bilden. Weil $L(V, V)$ endlich-dimensional, also ein BANACH-Raum und somit vollständig ist, existiert $S := \lim S_n$. Aus

$$\begin{aligned} S_n A &= E + A^{-1}(A - B) + \dots + (A^{-1}(A - B))^n \\ S_n A (A^{-1}(A - B)) &= \quad A^{-1}(A - B) + \dots + (A^{-1}(A - B))^n + (A^{-1}(A - B))^{n+1} \end{aligned}$$

folgt durch Subtraktion:

$$S_n B = E - (A^{-1}(A - B))^{n+1}.$$

Der letzte Term geht für $n \rightarrow \infty$ gegen Null, also ist $SB = E$, d.h. $S = B^{-1}$, also ist B invertierbar und es gilt $B \in U_{\frac{1}{\|A^{-1}\|}}(A) \implies B \in \text{GL}(n)$.

Verschärft man (21) zu

$$\|A - B\| \leq \frac{1}{2} \cdot \frac{1}{\|A^{-1}\|}, \quad (23)$$

(anstatt $\frac{1}{2}$ kann man jedes $0 < \lambda < 1$ wählen) so folgt mit der Dreiecksungleichung, der Eigenschaften der geometrische Reihe und (20)

$$\begin{aligned} \|S_n\| &\leq \left(\sum_{k=0}^n \|A^{-1}\|^k \underbrace{\|(A - B)\|^k}_{\leq \frac{1}{2^k} \|A^{-1}\|^k} \right) \|A^{-1}\| \leq \frac{1}{1 - \frac{1}{2}} \|A^{-1}\| = 2\|A^{-1}\|. \end{aligned}$$

Aus (23) folgt also $\|B^{-1}\| \leq 2\|A^{-1}\|$ und damit

$$\begin{aligned}\|B^{-1} - A^{-1}\| &= \|B^{-1}AA^{-1} - B^{-1}BA^{-1}\| \\ &= \|B^{-1}(A - B)A^{-1}\| \leq \|B^{-1}\|\|A - B\|\|A^{-1}\| \\ &\leq 2\|A^{-1}\|^2\|A - B\|.\end{aligned}$$

Das impliziert die Stetigkeit der Abbildung inv .

Aus

$$\text{inv}(B) = \text{inv}(A) + \underbrace{-A^{-1}(B - A)A^{-1}}_{D_A \text{ inv}(B)} + R(B) = \text{inv}(A) + A^{-1}(A - B)A^{-1} + R(B)$$

folgt

$$\begin{aligned}R(B) &= \text{inv}(B) - \text{inv}(A) + A^{-1}(B - A)A^{-1} \\ &= B^{-1} - A^{-1} + A^{-1}(B - A)A^{-1} \\ &= -A^{-1}(B - A)B^{-1} + A^{-1}(B - A)A^{-1} \\ &= A^{-1}(B - A)(A^{-1} - B^{-1}).\end{aligned}$$

Dann ist nach (20)

$$\frac{\|R(B)\|}{\|B - A\|} \leq \|A^{-1}\|\|A^{-1} - B^{-1}\|.$$

Wegen der Stetigkeit von inv geht das für $B \rightarrow A$ gegen 0 und inv ist differenzierbar mit der angegebenen Ableitung. \square

2.3 Richtungsableitungen, partielle Ableitungen

- Nun endlich die Differentiation für bescheidenere Ansprüche!
- Richtungs- und insbesondere partielle Ableitungen kann man mit Methoden der Analysis I ausrechnen.
- Aber der Zusammenhang mit der (totalen) Differenzierbarkeit ist nicht ganz trivial.
- Insbesondere gilt für $V = \mathbb{R}^n$: partielle Ableitungen existieren und sind stetig $\Rightarrow f$ differenzierbar \Rightarrow partielle Ableitungen existieren, die Rückrichtungen im Allgemeinen jeweils nicht.



Seien V, W endlich-dimensionale BANACH-Räume, $G \subset V$ offen, $p \in G$ und $f : V \supset G \rightarrow W$ eine Abbildung.

DEFINITION 2.3.1 (PARTIELLE- UND RICHTUNGSABLEITUNG)

Richtungsableitung ① Für $v \in V$ definiere die **Richtungsableitung** von f in p in Richtung v durch

$$\partial_v f(p) := \lim_{t \rightarrow 0} \frac{f(p + tv) - f(p)}{t},$$

falls dieser Limes existiert.

② Ist $V = \mathbb{R}^n$ mit der kanonischen Basis (e_1, \dots, e_n) , so nennt man die Richtungsableitungen in Richtungen e_i in $p \in G$

$$\partial_{e_i} f(p) := \frac{\partial f}{\partial x_i}(p) := \partial_i f(p) := \lim_{t \rightarrow 0} \frac{f(p_1, \dots, p_i + t, \dots, p_n) - f(p_1, \dots, p_n)}{t}$$

(statt x_i auch andere Variablennamen) die i -te **partiellen Ableitungen** von f in p .

Das ist (für $W = \mathbb{R}$) die Ableitung von f nach der i -ten Variablen im Sinne der Analysis I.

③ Ist $V = \mathbb{R}^n$ und $W = \mathbb{R}$, so heißt der Vektor

$$\text{grad } f(p) := \nabla f(p) := (\partial_{e_i} f(p))_{i \in \{1, \dots, n\}}$$

Gradient der **Gradient** von f in p .

Beispiel 2.3.2 Ist f in p differenzierbar, so existieren dort alle Richtungsableitungen und nach Lemma 2.1.3 gilt $\boxed{\partial_v f(p) = D_p f(v)}$. Speziell im Fall $V = \mathbb{R}^n$ gilt also $\boxed{\partial_i f(p) = D_p f(e_i)}$.

Beispiel 2.3.3 (Funktionalmatrix zu Fuß) Ist weiter $f : \mathbb{R}^n \supset G \rightarrow \mathbb{R}^m$, differenzierbar, so ist nach Beispiel 2.1.10 die Darstellungsmatrix von $D_p f$, also die Funktionalmatrix, mit Methoden der Analysis I zu berechnen:

$$\boxed{f'(p) = (\partial_j f_i(p))_{\substack{i=1, \dots, m \\ j=1, \dots, n}} = \left(\frac{\partial f_i}{\partial x_j}(p) \right)_{\substack{i=1, \dots, m \\ j=1, \dots, n}}}.$$

Beispiel 2.3.4 (Vgl. Beispiel 1.6.4) Existieren in p alle Richtungsableitungen, so muss f in p nicht differenzierbar, ja nicht einmal stetig sein, wie man an folgender Funktionen sehen kann:

$$f(x, y) := \begin{cases} 1 & \text{falls } y = x^2 > 0 \\ 0 & \text{sonst.} \end{cases}.$$

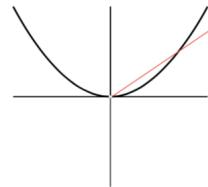


Abb. 26: Die im obigen Beispiel angeführte Funktion [Quelle: Ferus]

SATZ 2.3.1: DIFFERENZIERBARKEIT UND PARTIELLE DIFFERENZIERBARKEIT

Existieren auf ganz G alle Richtungsableitungen (oder im Fall $V = \mathbb{R}^n$ auch nur alle partiellen Ableitungen) und sind diese stetig, so ist f in G differenzierbar.

Beweis. Wir führen den Beweis nur für $V = \mathbb{R}^n, W = \mathbb{R}$. Mittels [komponentenweiser Differentiation bzw. partieller Differentiation](#) folgt daraus der Satz für $V = \mathbb{R}^n, W = \mathbb{R}^m$. Sind schließlich $\Phi : \mathbb{R}^n \rightarrow V$ und $\Psi : \mathbb{R}^m \rightarrow W$ Isomorphismen und setzt man $\tilde{f} := \Psi \circ f \circ \Phi^{-1}$, so ist f genau dann differenzierbar bzw. partiell differenzierbar, wenn das entsprechende für \tilde{f} gilt. Daraus folgt der Satz dann für beliebige V, W .

Seien also $V = \mathbb{R}^n, W = \mathbb{R}$. Für $p \in G$ definiere

$$F_p(x_1, \dots, x_n) := \sum_{j=1}^n x_j \partial_j f(p).$$

Dann ist $F_p : \mathbb{R}^n \rightarrow \mathbb{R}$ linear und der offensichtliche Kandidat für die Ableitung an der Stelle p . Wir betrachten eine offene ε -Kugel $U = U_\varepsilon(p)$, die ganz in der offenen Menge G liegt, und beschränken uns im folgenden auf $x \in U$. Beachten Sie, dass dann auch die Punkte $(p_1, \dots, p_j, x_{j+1}, \dots, x_n)$ in U und damit im Definitionsbereich von f liegen. Für $x \in U$ gilt daher

$$\begin{aligned} f(x) - f(p) &= f(x_1, \dots, x_n) - f(p_1, \dots, p_n) \\ &= f(x_1, \dots, x_n) - f(p_1, x_2, \dots, x_n) \\ &\quad + f(p_1, x_2, \dots, x_n) - f(p_1, p_2, x_3, \dots, x_n) \\ &\quad \vdots \\ &\quad + f(p_1, \dots, p_{n-1}, x_n) - f(p_1, \dots, p_n). \end{aligned}$$

Wir wenden auf jede Zeile den Mittelwertsatz an.

$$\begin{aligned} f(x) - f(p) &= \partial_1 f(\xi_1, x_2, \dots, x_n)(x_1 - p_1) \\ &\quad + \partial_2 f(p_1, \xi_2, \dots, x_n)(x_2 - p_2) \\ &\quad \vdots \\ &\quad + \partial_n f(p_1, \dots, p_{n-1}, \xi_n)(x_n - p_n) \end{aligned}$$

mit einem geeigneten ξ_i zwischen x_i und p_i . Für den Grenzübergang $x \rightarrow p$ gilt

$$(p_1, \dots, p_{i-1}, \xi_i, x_{i+1}, \dots, x_n) \rightarrow p \quad \forall i \in \{1, \dots, n\}.$$

Daraus folgt

$$\begin{aligned} \frac{\|f(x) - f(p) - F_p(x - p)\|}{\|x - p\|} &= \frac{\|\sum_{i=1}^n (x_i - p_i)(\partial_i f(p_1, \dots, p_{i-1}, \xi_i, x_{i+1}, \dots, x_n) - \partial_i f(p))\|}{\|x - p\|} \\ &\leq \underbrace{\sum_{i=1}^n \frac{|x_i - p_i|}{\|x - p\|}}_{\leq 1} \underbrace{\overbrace{\partial_i f(p_1, \dots, p_{i-1}, \xi_i, x_{i+1}, \dots, x_n)}^{\xrightarrow{x \rightarrow p} p \text{ weil } \partial_i f \text{ stetig.}} - \partial_i f(p)}_{\xrightarrow{x \rightarrow p} 0} \xrightarrow{x \rightarrow p} 0. \end{aligned}$$

Beispiel 2.3.5 Die Abbildung $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $(x, y) \mapsto (\sin x \cos y, \sin x \sin y, \cos y)$ hat die folgende Matrix partieller Ableitungen:

$$(\partial_j f_i(x, y)) = \begin{pmatrix} \cos x \cos y & -\sin x \sin y \\ \cos x \sin y & \sin x \cos y \\ 0 & -\sin y \end{pmatrix}.$$

Die partiellen Ableitungen sind offensichtlich stetig. Daher ist die Funktion differenzierbar und das Differential $D_{(x,y)} f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ wird durch die obige Matrix $f'(x, y)$ repräsentiert.

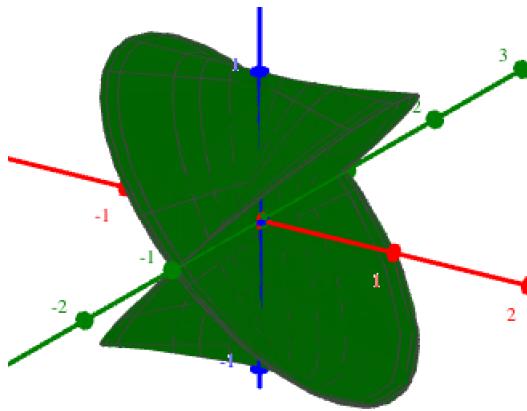


Abb. 27: [Quelle: GeoGebra5]

Bemerkung (Kugelkoordinaten) Ersetzt man im letzten Argument von $f \cos(y)$ durch $\cos(x)$, so erhält man die Einheitskugel S^2 , und kann auch $\|(\sin x \cos y, \sin x \sin y, \cos y)\| = 1$ mit der euklidischen Norm leicht nachrechnen. Schränkt man diese Funktion jetzt auf $[-\pi, \pi] \times [0, \pi]$ (??) ein, so eine **Parametrisierung** auf die Oberfläche von S^2 , d.h. das neue f liefert eine **Parametrisierung** dieser zweidimensionalen Untermannigfaltigkeit des \mathbb{R}^3 .

Beispiel 2.3.6 (Das "totale Differential") Die Koordinaten-Abbildungen

$$x_i : \mathbb{R}^n \rightarrow \mathbb{R}, (v_1, \dots, v_n) \mapsto v_i$$

sind linear. Deshalb gilt $D_p x_i = x_i$ für alle p in \mathbb{R}^n .

Jedes $v \in \mathbb{R}^n$ lässt sich schreiben als

$$v = \sum x_i(v) e_i = \sum D_p x_i(v) e_i.$$

Ist f in $p \in G \subset \mathbb{R}^n$ differenzierbar, so folgt

$$D_p f(v) = D_p f(\sum D_p x_i(v) e_i) \stackrel{\text{L}}{=} \sum_i D_p x_i(v) D_p f(e_i) = \sum_i \partial_i f(p) D_p x_i(v).$$

Das schreibt man auch so:

$$Df = \sum \partial_i f D x_i \quad (24)$$

Man nennt diesen Ausdruck das **totale Differential** von f im Gegensatz zu **teiliale Differential** einzelnen partiellen Differentialen $\frac{\partial f}{\partial x_i}$. Bei Lichte besehen ist das totale Differential an der Stelle p aber einfach nur die Ableitung $D_p f$.

Beispiel 2.3.7 (Kettenregel in partiellen Ableitungen) Für differenzierbare Abbildungen zwischen den Standardräumen sieht die Kettenregel in partiellen Ableitungen folgendermaßen aus:

Aus $D_p(f \circ g) = D_{g(p)} f \circ D_p g$ folgt nach linearer Algebra (Übung!) (und $V = \mathbb{R}^n$, $W = \mathbb{R}^k$ und $U = \mathbb{R}^m$)

$$(f \circ g)'(p) = f'(g(p))g'(p),$$

wobei rechts das Produkt der Matrizen steht. Also, wenn wir die Variablen im Definitionsraum von f mit y_j und im Definitionsraum von g mit x_k bezeichnen,

$$\left(\frac{\partial(f \circ g)_i}{\partial x_k}(p) \right)_{i,k} = \left(\frac{\partial f_i}{\partial y_j}(g(p)) \right)_{i,j} \cdot \left(\frac{\partial g_j}{\partial x_k}(p) \right)_{j,k}$$

oder $(\forall i, k)$

$$\frac{\partial(f \circ g)_i}{\partial x_k} = \sum_j \frac{\partial f_i}{\partial y_j}(g(p)) \frac{\partial g_j}{\partial x_k}(p).$$

Abgekürzte Notation:

$$\frac{\partial f_i}{\partial x_k} = \sum_j \frac{\partial f_i}{\partial y_j} \frac{\partial y_j}{\partial x_k}.$$

2.4 Höhere Ableitungen und der Satz von SCHWARZ

- Die Ableitung einer Funktion von mehreren Variablen ist nicht eine Zahl, sondern eine Lineare Abbildung. Entsprechend werden erst recht die höheren Ableitungen solcher Funktionen kompliziertere Gebilde, nämlich multilinear Abbildungen. Wir lernen, wie man sie "trotzdem" effektiv berechnen kann.
- Wir zeigen: k -te partielle Ableitungen existieren und sind stetig $\Rightarrow f$ k -mal differenzierbar \Rightarrow die k -ten partiellen Ableitungen existieren.
- Wir lernen den Satz von Schwarz (f k -mal differenzierbar \Rightarrow Reihenfolge der Ableitungen vertauschbar), der manche Rechnung vereinfacht, aber auch wichtige Anwendungen auf Differentialgleichungen hat.

Bemerkung 2.4.1 Wir erinnern daran, dass $L(V, W)$ den Vektorraum aller linearen Abbildungen von V nach W bezeichnet. Sind V und W endlich-dimensional, so ist auch $L(V, W)$ endlich-dimensional und es gilt $\dim L(V, W) = (\dim V)(\dim W)$, vgl. Lineare Algebra.

DEFINITION 2.4.2 (ZWEITE ABLEITUNG)

Sei $f : V \supset G \rightarrow W$ auf G differenzierbar. Dann existiert

$$Df : G \rightarrow L(V, W), p \mapsto D_p f \in L(V, W) \quad \text{mit } D_p f(v) \in W \quad \forall v \in V.$$

Ist diese Abbildung differenzierbar in $p \in G$, so heißt f in p zweimal differenzierbar, wenn

$$D_p^2 f := D_p(Df) : V \rightarrow L(V, W),$$

die zweite Ableitung von f in p , existiert.

Bemerkung Es gilt

$$\begin{aligned} D^2 f &\in L(V, L(V, W)) \quad \text{aber} \quad p \mapsto D_p^2 f : V \rightarrow L(V, W) \quad \text{sowie} \\ D_p^2 f(v_1) &\in L(V, W) \quad \text{und} \quad D_p^2 f(v_1)(v_2) \in W. \end{aligned}$$

Beispiel 2.4.3 (Zweite Ableitung der Normabbildung mit der Kettenregel)

Sei $(V, \langle \cdot, \cdot \rangle)$ ein Euklidischer Vektorraum und sei $r : V \setminus \{0\} \rightarrow \mathbb{R}$ gegeben durch $r(x) = \sqrt{\langle x, x \rangle}$. Wir haben im Beispiel 2.2.4 ausgerechnet, dass

$$D_p r(v) = \frac{1}{r(p)} \langle p, v \rangle,$$

also

$$Dr : x \mapsto D_x r = \frac{1}{r(x)} \langle x, . \rangle \in L(V, \mathbb{R}).$$

Das ist das Produkt der Abbildung $\frac{1}{r} : V \setminus \{0\} \rightarrow \mathbb{R}$ mit der Abbildung

$$g : V \rightarrow L(V, \mathbb{R}), x \mapsto \langle x, . \rangle$$

1. Faktor: Weil r differenzierbar ist, ist nach der Kettenregel auch $\frac{1}{r}$ differenzierbar, und es ist

$$D_p \frac{1}{r}(v) = -\frac{1}{r^2(p)} D_p r(v) = -\frac{1}{r^3(p)} \langle p, v \rangle.$$

2. Faktor: Die Abbildung $g : V \rightarrow L(V, \mathbb{R})$ ist linear, also auch differenzierbar und

$$D_p g(v) = g(v) = \langle v, \cdot \rangle.$$

Produktregel: Daher ist Dr nach der Produktregel differenzierbar, und es gilt

$$\begin{aligned} D_p(Dr)(v) &= -\frac{1}{r^3(p)} \langle p, v \rangle \langle p, \cdot \rangle + \frac{1}{r(p)} \langle v, \cdot \rangle \\ &= \frac{1}{r^3(p)} (r^2(p) \langle v, \cdot \rangle - \langle p, v \rangle \langle p, \cdot \rangle) \\ &= \frac{1}{r^3(p)} (\langle p, p \rangle \langle v, \cdot \rangle - \langle p, v \rangle \langle p, \cdot \rangle), \end{aligned}$$

also

$$D_p^2 r(v)(w) = \frac{1}{r^3(p)} (\langle p, p \rangle \langle v, w \rangle - \langle p, v \rangle \langle p, w \rangle). \quad (25)$$

Es ist klar, wie man höhere als 2. Ableitungen definiert. Dabei entsteht allerdings ein kleines Problem: Wir erhalten $D_p^3 f \in L(V, L(V, L(V, W)))$, und den Zielraum der 7. Ableitung mag man nicht mehr hinschreiben.

Bezeichnen wir mit $L^k(V^k, W)$ den Vektorraum der k -linearen Abbildungen von V^k nach W , so haben wir allgemeiner einen kanonischen Isomorphismus

$$j_k : L(V, L^{k-1}(V, W)) \rightarrow L^k(V^k, W)$$

mit

$$j_k(\Phi)(v_1, \dots, v_k) := \Phi(v_1)(v_2, \dots, v_k)$$

für $(v_1, \dots, v_k \in V)$ (Beweisen Sie das! (s. HA) Es folgt, dass $\dim L^k(V, W) = (\dim V)^k (\dim W) < \infty$, wenn V und W endlich-dimensional sind.)

Damit definieren wir induktiv die k -Ableitung $D_p^k f$ einer Funktion f an der Stelle p wie folgt:

DEFINITION 2.4.4 (HÖHERE ABLEITUNGEN)

Ist $f : V \supset G \rightarrow W$ bereits $(k-1)$ mal differenzierbar und ist die $(k-1)$ -te Ableitung von f in p

$$D^{k-1} f : G \rightarrow L^{k-1}(V, W), x \mapsto D_x^{k-1} f$$

in $p \in G$ differenzierbar, so heißt f in p k -mal differenzierbar und die k -te Ableitung von f in p angewendet auf $v_1, \dots, v_k \in V$ ist gegeben durch

$$D_p^k f(v_1, \dots, v_k) := j_k(D_p(D^{k-1} f))(v_1, \dots, v_k) = D_p(D^{k-1} f)(v_1)(v_2, \dots, v_k) \in W.$$

Die k -te Ableitung einer k -mal differenzierbaren Abbildung $f : V \supset G \rightarrow W$ an einer Stelle p ist also eine k -lineare Abbildung

$$D_p^k f \in L^k(V, W).$$

Lemma 2.4.5

Ist f in p k -mal differenzierbar so gilt für $v_1, \dots, v_k \in V$

$$D_p^k f(v_1, \dots, v_k) = \partial_{v_1} \dots \partial_{v_k} f(p).$$

Insbesondere existiert die rechte Seite.

Beweis. Durch Induktion über k .

$k = 1$. $D_p f(v) = \partial_v f(p)$ wissen wir schon.

$(k - 1) \rightarrow k$. Die Abbildung

$$\begin{array}{ccc} g : L^{k-1}(V, W) & \rightarrow & W \\ \mu & \mapsto & \mu(v_2, \dots, v_k). \end{array}$$

ist linear. Nach Kettenregel und Voraussetzung ist daher $g \circ D^{k-1} f$ differenzierbar mit

$$\begin{aligned} D_p(g \circ D^{k-1} f)(v_1) &= g(D_p(D^{k-1} f)(v_1)) = D_p(D^{k-1} f)(v_1)(v_2, \dots, v_k) \\ &= D_p^k f(v_1, \dots, v_k). \end{aligned}$$

Andererseits ist nach Induktionsvoraussetzung

$$g \circ D^{k-1} f = D^{k-1} f(v_2, \dots, v_k) = \partial_{v_2} \dots \partial_{v_k} f$$

und deshalb nach dem Fall $k = 1$

$$D_p(g \circ D^{k-1} f)(v_1) = \partial_{v_1}(g \circ D^{k-1} f)(p) = \partial_{v_1} \dots \partial_{v_k} f(p).$$

□

Bemerkung Dieses Lemma impliziert insbesondere folgende Vereinfachung: Um $D_p^2 f(v_1, v_2)$ zu berechnen, muss ich nicht die schwerer vorzustellende Abbildung $Df : G \rightarrow L(V, W)$ differenzieren, sondern ich kann $Df(v_2) : G \rightarrow W$ in Richtung v_1 differenzieren: Ich darf **vor** der zweiten Ableitung den Vektor v_2 einsetzen.

$$D_p^2 f(v_1, v_2) = \partial_{v_1} \partial_{v_2} f(p) = D_p(Df(v_2))(v_1).$$

Dabei muss man auf die Reihenfolge der Vektoren achten – bis wir gleich gezeigt haben, dass sie keine Rolle spielt!

Bemerkung (Erinnerung) Ist $f : \mathbb{R}^n \supset G \rightarrow W$ in $p \in G$ k -mal differenzierbar, so gilt

$$D_v f(p) = \langle v, \text{Grad}(f) \rangle = \sum_{i=1}^n \underbrace{v_i}_{\in \mathbb{R}} \partial_i f(p).$$

Durch Iteration erhält man das folgende

Korollar 2.4.6 (Höhere Ableitungen auf dem \mathbb{R}^n , Anwendung des vorherigen Lemmas)

Ist $f : \mathbb{R}^n \supset G \rightarrow W$ in $p \in G$ k -mal differenzierbar, und hat man k Vektoren

$v_j = (v_{1j}, \dots, v_{nj}) \in \mathbb{R}^n$ mit $j \in \{1, \dots, k\}$ gegeben, so gilt

$$D_p^k f(v_1, \dots, v_k) = \sum_{i_1, \dots, i_k=1}^n \partial_{i_1} \dots \partial_{i_k} f(p) v_{i_1 1} \dots v_{i_k k}. \quad (26)$$

Also lässt sich die k -te Ableitung von f mittels k -facher partieller Ableitungen ausrechnen.

Beweis. (König) Seien $p \in G$ und $v_1, \dots, v_n \in V$. Dann gilt

$$\underbrace{D_p^k f}_{\in L^k(V^k, W)}(v_1, \dots, v_n) = D_p(D^{k-1} f)(v_1)(v_2, \dots, v_k),$$

wobei $D^{k-1}f : V \supset G \rightarrow L(V, L^{k-1}(V, W))$, $p \mapsto D_p^{k-1}f$ und $L^{k-1}(V, W)$ die Menge der $k-1$ linearen Abbildungen von $V^{k-1} \rightarrow W$ ist.

Ist $(e_i)_{i \in I}$ ein Basis von V , so gilt $v_i = \sum_j v_{j,i} e_i$ und somit aufgrund der Linearität von ∂_ξ in ξ

$$D_p^k f(v_1, \dots, v_n) = \partial_{v_1} \dots \partial_{v_k} f(p) = \sum_{j_1, \dots, j_k} \partial_{e_{j_1}} \dots \partial_{e_{j_k}} f(p) \cdot v_{j_1,1} \dots v_{j_k,k}. \quad \square$$

Beispiel 2.4.7 Für den Fall $k = 2$ gilt

$$\begin{aligned} D_p^2 f(v_1, v_2) &= (v_{1,1}, \dots, v_{n,1}) (\partial_i \partial_j f(p))_{i,j \in \{1, \dots, n\}} \begin{pmatrix} v_{1,2} \\ \vdots \\ v_{n,2} \end{pmatrix} = v_1^T \left(\sum_{j=1}^n \partial_i \partial_j v_{j,2} \right)_{i \in \{1, \dots, n\}} \\ &= \sum_{i=1}^n v_{i,1} \sum_{j=1}^n \partial_i \partial_j f(p) v_{j,2} = \sum_{i,j=1}^n \partial_i \partial_j f(p) v_{i,1} v_{j,2}. \end{aligned}$$

Beispiel 2.4.8 Konkret betrachten wir die Normfunktion $r(x) = \|x\| = \sqrt{\sum_{i=1}^n x_i^2}$ auf $\mathbb{R}^n \setminus \{0\}$. Nach den Ableitungsregeln aus Analysis I gilt

$$\begin{aligned} \partial_j r(x) &= \frac{1}{2} (\|x\|)^{-1} \cdot 2x_j = \frac{x_j}{r(x)}, \\ \partial_i \partial_j r(x) &= \frac{r(x) \cdot \delta_{ij} - x_j \frac{x_i}{r(x)}}{r(x)^2} = \frac{1}{r(x)^3} (r(x)^2 \delta_{ij} - x_i x_j) \end{aligned}$$

und damit

$$\begin{aligned} D_p^2 r(v_1, v_2) &= \frac{1}{\|p\|^3} \sum_{i=1}^n v_{i,1} v_{j,2} (r^2(p) \delta_{i,j} - p_i p_j) \\ &= \frac{1}{\|p\|^3} \left[r^2(p) \left(\sum_{i=1}^n v_{i,1} v_{i,2} \right) - \left(\sum_{i=1}^n v_{i,1} p_i \right) \left(\sum_{i=1}^n v_{j,2} p_j \right) \right] \\ &= \frac{1}{\|p\|^3} [\|p\|^2 \langle v_1, v_2 \rangle - \langle p, v_1 \rangle \langle p, v_2 \rangle]. \end{aligned}$$

Vergleichen Sie das mit (25).

Für spätere Verwendung zeigen wir hier noch das folgende

Lemma 2.4.9 (Ableiten ist "assoziativ")

Für k -mal differenzierbares $f : V \supset G \rightarrow W$, $p \in G$ und $v_1, \dots, v_k \in V$ gilt

$$D_p^k f(v_1, \dots, v_k) = D_p^{k-1}(Df)(v_1, \dots, v_{k-1})(v_k).$$

Beweis. (weggelassen) Ich beweise das für $V = \mathbb{R}^n$ mit den Koordinatenprojektionen $x_j : \mathbb{R}^n \rightarrow \mathbb{R}$. Der allgemeine Fall geht nach Wahl einer Basis genauso, nur treten an die Stelle der x_j dann die dualen Basisvektoren. Es ist

$$\begin{aligned} D_p^{k-1}(Df)(v_1, \dots, v_{k-1})(v_k) &= D_p^{k-1} \left(\sum_{j=1}^n (\partial_j f) x_j \right) (v_1, \dots, v_{k-1})(v_k) \\ &= \sum_{i_1, \dots, i_{k-1}, j=1}^n \partial_{i_1} \dots \partial_{i_{k-1}} \partial_j f(p) v_{i_1,1} \dots v_{i_{k-1}, k-1} x_j(v_k) \\ &= \sum_{i_1, \dots, i_{k-1}, j=1}^n \partial_{i_1} \dots \partial_{i_{k-1}} \partial_j f(p) v_{i_1,1} \dots v_{i_{k-1}, k-1} v_{j,k}. \end{aligned}$$

Vergleich mit (26) liefert die Behauptung. \square

Für die Frage, ob f zweimal differenzierbar ist, existiert ebenfalls ein gutes Kriterium mittels partieller Ableitungen: Sei f differenzierbar (z.B. weil es überall stetige partielle Ableitungen besitzt). Nach (24) ist Df wegen der Konstanz der Abbildungen $p \mapsto D_p x_i = x_i$ genau dann differenzierbar, wenn die partiellen Ableitungen $\partial_i f$ alle differenzierbar sind. Das lässt sich wieder mittels partieller Ableitungen testen, und man erhält: Existieren alle zweiten partiellen Ableitungen von f auf G und sind sie dort stetig, so ist f zweimal differenzierbar.

Entsprechendes gilt für höhere Ableitungen.

DEFINITION 2.4.10 (C^k -FUNKTIONEN)

Ist $f : V \supset G \rightarrow W$ k -mal differenzierbar und die Abbildung

$$D^k f : G \rightarrow L^k(V, W), \quad p \mapsto D_p^k f$$

k -mal stetig, so heißt f k -mal stetig differenzierbar.

differenzierbar

Schließlich soll $f \in C^\infty$ bedeuten, dass f beliebig oft differenzierbar ist – die Stetigkeit der Ableitungen folgt dann natürlich von selbst.

Bemerkung (klar.) $f : \mathbb{R}^n \rightarrow W$ k -mal stetig partiell differenzierbar $\iff f$ k -mal stetig differenzierbar ($\iff f \in C^k$) $\implies f$ k -mal differenzierbar $\implies f$ k -mal partiell differenzierbar.

Lemma 2.4.11

Sei $f : V \supset G \rightarrow W$ differenzierbar und in $p \in G$ zweimal differenzierbar. Dann gilt für alle $u, v \in V$

$$D_p^2 f(u, v) = \lim_{t \rightarrow 0} \frac{f(p + tu + tv) - f(p + tu) - f(p + tv) + f(p)}{t^2}.$$

Bemerkung 2.4.12 (Eindimensionaler Vergleich, von mir) Sei I ein reelles Intervall und $f : I \rightarrow \mathbb{R}$ zweimal differenzierbar. Für $x \in I$ gilt dann $f''(x) = \lim_{t \rightarrow 0} \frac{f(x + h) - 2f(x) + f(x - h)}{t^2}$.

Bemerkung (Anwendungsbeispiel: Diskretisierung von part. DGlen) In der Numerik benutzt man oft Approximationen für $D_p^2 f(u, v)$ der Form

$$f(p + u + v) - f(p + u) - f(p + v) + f(p) \approx D_{p+u} f(v) - D_p f(v) \approx D_p^2 f(u, v).$$

für u, v "hinreichend klein".

In Dimension eins gilt analog für eine Funktion $g : \mathbb{R} \rightarrow \mathbb{R}$

$$g''(t_0) \approx \frac{1}{h} \left[\underbrace{\frac{g(t_0 + h) - g(t_0)}{h}}_{\approx g'(t_0 + \frac{h}{2})} + \underbrace{\frac{g(t_0 - h) - g(t_0)}{h}}_{\approx g'(t_0 - \frac{h}{2})} \right]$$

"Rechtsableitung" "Linksableitung"

[BILD!!]

Weil im Lemma der Zähler rechts in u und v symmetrisch ist, folgt daraus der wichtige

SATZ 2.4.1: H.A. SCHWARZ

Sei $f : V \supset G \rightarrow W$ in G zweimal differenzierbar. Dann gilt für alle $u, v \in V$ und $p \in G$

$$D_p^2 f(u, v) = D_p^2 f(v, u).$$

Beweis. (des Lemmas 2.4.11) 1. Schritt. Es genügt der Beweis für den Fall $W = \mathbb{R}$, weil $f = \sum_i f_i e_i$, wo die e_i eine Basis von W und die f_i reellwertige Funktionen sind.

2. Schritt. In Anlehnung an die heuristische Betrachtung oben definieren wir eine Funktion

$$F(t) := f(p + u + tv) - f(p + tv).$$

Dabei seien $\|u\|$ und $\|v\|$ hinreichend klein, so dass $p + u + tv$ und $p + tv$ für $0 \leq t \leq 1$ in G liegen. Dann ist nach dem Mittelwertsatz für ein $\tau \in]0, 1[$

$$\begin{aligned} f(p + u + v) - f(p + u) - f(p + v) + f(p) &= F(1) - F(0) = F'(\tau) \\ &= D_{p+u+\tau v} f(v) - D_{p+\tau v} f(v) \quad (\text{Kettenregel}) \\ &= (D_{p+u+\tau v} f - D_p f)(v) - (D_{p+\tau v} f - D_p f)(v) \\ &=: (\star) \end{aligned}$$

Wir wenden jetzt auf die Funktion Df die Definition der Differenzierbarkeit an der Stelle p an, und erhalten für $x = p + u + \tau v$ bzw. $x = p + \tau v$

$$\begin{aligned} (\star) &= D_p^2 f(u + \tau v)(v) + R(p + u + \tau v)(v) \\ &\quad - D_p^2 f(\tau v)(v) - R(p + \tau v)(v) \\ &= D_p^2 f(u, v) + (R(p + u + \tau v) - R(p + \tau v))(v). \end{aligned}$$

Zu gegebenem $\varepsilon > 0$ existiert ein $\delta > 0$, so dass

$$\|R(x)\| \leq \varepsilon \|x - p\|, \text{ falls } \|x - p\| < \delta. \quad (27)$$

Beachten Sie, dass $R(x) \in L(V, W)$. Für die Norm von $R(x)$ verwenden wir daher wie üblich die Operatornorm auf $L(V, W)$.

Für $\|u\| + \|v\| < \delta$ ist dann

$$\begin{aligned} \|R(p + u + \tau v)\| &\leq \varepsilon \|u + \tau v\| \leq \varepsilon (\|u\| + \|\tau v\|), \\ \|R(p + \tau v)\| &\leq \varepsilon \|\tau v\| \leq \varepsilon (\|u\| + \|\tau v\|), \end{aligned}$$

also

$$\|(\star) - D_p^2 f(u, v)\| = \|(R(p + u + \tau v) - R(p + \tau v))(v)\| \leq 2\varepsilon (\|u\| + \|\tau v\|) \|v\|.$$

Nun seien $u, v \in V$ beliebig und $t_0 > 0$ bei vorgegebenem $\varepsilon > 0$ so klein gewählt, dass $\|t_0 u\| + \|t_0 v\| < \delta$ ist, vgl. (27). Dann folgt für alle t mit $|t| \leq t_0$

$$\begin{aligned} &\left\| \frac{f(p + tu + tv) - f(p + tu) - f(p + tv) + f(p)}{t^2} - D_p^2 f(u, v) \right\| \\ &= \left\| \frac{f(p + tu + tv) - f(p + tu) - f(p + tv) + f(p) - D_p^2 f(tu, tv)}{t^2} \right\| \\ &\leq 2\varepsilon \frac{(\|tu\| + \|tv\|) \|tv\|}{t^2} = 2\varepsilon (\|u\| + \|v\|) \|v\| \end{aligned}$$

Daraus folgt die Behauptung. \square

Korollar 2.4.13 (zum Satz von Schwarz)

Ist $f : \mathbb{R}^n \supset G \rightarrow W$ in G zweimal differenzierbar (oder zweimal partiell differenzierbar mit stetigen zweiten partiellen Ableitungen), so gilt für alle i, j

$$\partial_i \partial_j f = \partial_j \partial_i f.$$

Beispiel 2.4.14 (Wichtige Anwendung: Integrabilitätskriterium) Seien $f_1, \dots, f_n : \mathbb{R}^n \supset G \rightarrow \mathbb{R}$. Die elementarste Frage der Theorie der partiellen Differentialgleichungen ist die, ob es eine Funktion $y : G \rightarrow \mathbb{R}$ gibt, so dass für alle i gilt:

$$\partial_i y = f_i. \quad (28)$$

Haben die f_i stetige partielle Ableitungen, so hat ein solches y , falls es existiert, **notwendige** partielle Ableitungen bis zur Ordnung 2. Also ist eine **notwendige** Bedingung für die Lösbarkeit von (28), dass

$$\partial_j f_i = \partial_j \partial_i y = \partial_i \partial_j y = \partial_i f_j,$$

d.h.

$$\partial_j f_i = \partial_i f_j \quad \text{für alle } i, j.$$

Aus dem Satz von Schwarz in Verbindung mit dem Lemma 2.4.5 ergibt sich unmittelbar die folgende Verallgemeinerung:

Korollar 2.4.15 (zum Satz von Schwarz)

Ist $f : V \supset G \rightarrow W$ k -mal differenzierbar in G , so gilt für jede Permutation (i_1, \dots, i_k) von $(1, \dots, k)$ und für alle $p \in G$ und $v_1, \dots, v_k \in V$

$$D_p^k f(v_1, \dots, v_k) = D_p^k f(v_{i_1}, \dots, v_{i_k})$$

bzw.

$$\partial_{v_1} \dots \partial_{v_k} f(p) = \partial_{v_{i_1}} \dots \partial_{v_{i_k}} f(p).$$

Aus dem Vergleich des Satzes von Schwarz mit Korollar 2.4.13 ergeben sich zwei Fragen:

- ① existiert Funktionen, die zweimal differenzierbar, aber nicht zweimal stetig differenzierbar sind? Dann ist die Version des Satzes 2.4.1 stärker als der Korollar.

Antwort: Ja. Die Funktion

$$g : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto x^4 \sin \frac{1}{x}$$

ist auf \mathbb{R} zweimal differenzierbar, aber die 2. Ableitung ist in 0 nicht stetig. Entsprechendes gilt dann auch für die durch $f(x, y) := g(x)$ definierte Funktion auf \mathbb{R}^2 .

- ② existiert Funktionen mit (nicht stetigen) partiellen Ableitungen 2. Ordnung, für die der Satz von Schwarz nicht gilt?

Antwort: Ja. Die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$f(x, y) := \begin{cases} 0 & \text{für } x = y = 0 \\ xy \frac{x^2 - y^2}{x^2 + y^2} & \text{sonst} \end{cases}$$

besitzt überall stetige 1. partielle Ableitungen und (in 0 unstetige) 2. partielle Ableitungen. Für sie gilt

$$\partial_1 \partial_2 f(0, 0) = 1 \neq -1 = \partial_2 \partial_1 f(0, 0).$$

Beweise als Übung.

2.5 Schrankensatz und der Satz von TAYLOR

- Schrankensatz und Satz von Taylor kann man auch für Funktionen mehrerer Variablen formulieren.

In höherdimensionalen Räumen wird die naive Verallgemeinerung des Mittelwertsatzes

$$f(b) - f(a) \stackrel{?}{=} D_\xi f(b-a) \text{ für geeignetes } \xi \text{ zwischen (?) } a \text{ und } b$$

falsch. Zum Beispiel gilt für die Funktion $f : [0, 2\pi] \rightarrow \mathbb{R}^3$ mit $f(t) = (\cos t, \sin t, t)$, dass

$$f(2\pi) - f(0) = (0, 0, 2\pi).$$

Aber für alle $\xi \in]0, 2\pi[$ ist $D_\xi f(2\pi - 0) = 2\pi(-\sin t, \cos t, 1) \neq (0, 0, 2\pi)$.

Der Schrankensatz, den wir jetzt beweisen, liefert einen Ersatz für den Mittelwertsatz.

SATZ 2.5.1: SCHRANKENSATZ

Seien V, W endlich-dimensionale Banachräume, G eine offene Teilmenge und $f : V \supset G \rightarrow W$ eine differenzierbare Abbildung sowie $a, b \in G$, sodass die Strecke

$$\overline{ab} := \{a + t(b-a) \mid 0 \leq t \leq 1\}$$

in G enthalten ist. Dann gilt

$$\|f(b) - f(a)\| \leq \sup_{x \in \overline{ab}} \|D_x f\| \|b - a\|.$$

Beweis. Sei $K := \sup_{x \in \overline{ab}} \|D_x f\|$ und sei $\varepsilon > 0$. Sei o.E. $K < \infty$, sonst ist nichts zu zeigen. Wir definieren

$$A := \{t \in [0, 1] : \|f(a + t(b-a)) - f(a)\| \leq t(K + \varepsilon)\|b - a\|\}.$$

Weil beide Seiten der Ungleichung in der Definition von A in $t \in [0, 1]$ stetig sind, ist A abgeschlossen. Wegen $0 \in A$ ist $A \neq \emptyset$. Insbesondere gilt

$$\sup A =: s \in A \subset [0, 1].$$

Wir zeigen $s = 1$, d.h.

$$\|f(b) - f(a)\| \leq (K + \varepsilon)\|b - a\|.$$

Weil das für alle $\varepsilon > 0$ gilt, folgt daraus die Behauptung.

Annahme: $s < 1$. Die Funktion f ist in $p = a + s(b-a)$ differenzierbar, und wir haben

$$f(a + t(b-a)) = f(a + s(b-a)) + D_p f((t-s)(b-a)) + R(a + t(b-a)) \quad (29)$$

mit $\lim_{x \rightarrow p} \frac{R(x)}{\|x-p\|} = 0$, also $\lim_{t \rightarrow s} \frac{R(a+t(b-a))}{|t-s|\|b-a\|} = 0$.

(Beachten Sie $\|(a + t(b-a)) - (a + s(b-a))\| = \|(t-s)(b-a)\| = |t-s|\|b-a\|$.)

Daher existiert $\delta > 0$, so dass für $s < t < s + \delta$ gilt

$$\|R(a + t(b - a))\| \leq \varepsilon(t - s) \cdot \|b - a\|.$$

Aus Umstellen von (29) und der Dreiecksungleichung folgt für $s < t < s + \delta$

$$\begin{aligned} \|f(a + t(b - a)) - f(a + s(b - a))\| &\leq \|D_p f\| (t - s) \cdot \|b - a\| + \varepsilon(t - s) \cdot \|b - a\| \\ &\leq (t - s)(K + \varepsilon)\|b - a\|. \end{aligned}$$

und weiter

$$\begin{aligned} \|f(a + t(b - a)) - f(a)\| &\leq \|f(a + t(b - a)) - f(a + s(b - a))\| + \|f(a + s(b - a)) - f(a)\| \\ &\leq (t - s)(K + \varepsilon)\|b - a\| + s(K + \varepsilon)\|b - a\| \\ &= t(K + \varepsilon)\|b - a\|. \end{aligned}$$

Das ist ein Widerspruch zur Definition von s , da die zu zeigende Ungleichung auch für $t > s$ gilt $s \neq \sup(A)$. Somit ist $s = 1$. \square

Korollar 2.5.1 (zum Schrankensatz)

Wenn f in der Situation des obigen Satzes in den Endpunkten von \overline{ab} nicht differenzierbar, aber stetig ist, gilt dieselbe Behauptung, wobei das Supremum über alle Punkte von $\overline{ab} \setminus \{a, b\}$ zu nehmen ist.

Beweis. Ist f nur im "Inneren" der Strecke \overline{ab} differenzierbar, so gilt nach dem zu Beweisenen für alle $0 < t_1 < t_2 < 1$

$$\begin{aligned} \|f(a + t_2(b - a)) - f(a + t_1(b - a))\| &\leq \sup_{0 < t < 1} \|D_{a+t(b-a)} f\| \|(a + t_2(b - a)) - (a + t_1(b - a))\| \\ &= \sup_{0 < t < 1} \|D_{a+t(b-a)} f\| (t_2 - t_1)\|b - a\|. \end{aligned}$$

Durch Grenzübergang $t_1 \searrow 0$ und $t_2 \nearrow 1$ folgt mit der Stetigkeit von f die Behauptung. \square

Beispiel 2.5.2 (Berechnung der Operatornorm) Der Schrankensatz macht es wünschenswert, die Operatornorm explizit zu berechnen. Das ist ein Problem der linearen Algebra. Wir geben die Resultate für zwei einfache Fälle.

- ① $F : \mathbb{R}^n \rightarrow \mathbb{R}$ sei bezüglich der Standardbasen gegeben durch die $m \times 1$ -Matrix $A = (a_1, \dots, a_n)$. Dann ist die Operatornorm bezüglich der Euklidischen Norm von \mathbb{R}^n und dem Betrag $|\cdot|$ auf \mathbb{R} gegeben durch

$$\|A\| = \sqrt{\sum_{i=1}^n a_i^2}.$$

Im Falle des Schrankensatzes, ist $F = Df$ gegeben durch die Matrix

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

und

$$\|Df\| = \sqrt{\sum_{i=1}^n (\partial_i f)^2}.$$

- ② $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ sei bezüglich der Standardbasen gegeben durch die $m \times n$ -Matrix A . Die Normen auf \mathbb{R}^n und \mathbb{R}^m seien die üblichen Euklidischen Normen. Mit A^T bezeichnen wir die transponierte Matrix. Dann ist $A^T A$ eine symmetrische (=selbstadjungierte) $n \times n$ -Matrix und $\|A\|$ ist die Wurzel aus dem Maximum der Eigenwerte von $A^T A$.

Korollar 2.5.3 (Konstanzkriterium I)

Sei $G \subset V$ offen und zusammenhängend und sei $f : G \rightarrow \mathbb{R}$ differenzierbar mit $D_p f = 0$ für alle $p \in G$. Dann ist f konstant.

Beweis. Sei $p \in G$ und

$$U := \{x \in G \mid f(x) = f(p)\}.$$

Ist $a \in U$, so existiert $r > 0$ mit $U_r(a) \subset G$. Aus dem Schrankensatz folgt dann für alle $b \in U_r(a)$

$$\|f(b) - f(p)\| = \|f(b) - f(a)\| \leq 0 \cdot \|b - a\|,$$

also $b \in U$. Deshalb ist U offen. Andererseits ist

$$G \setminus U = f^{-1}(\mathbb{R} \setminus \{f(p)\})$$

ebenfalls offen. Weil G zusammenhängend ist, folgt dass entweder $U = \emptyset$ oder $G \setminus U = \emptyset$. Wegen $p \in U$ folgt $U = G$. \square

Korollar 2.5.4

Seien $f, g \in V \supset U \rightarrow W$ differenzierbar, U offen und zusammenhängend sowie $D_x f = D_x g$ für alle $x \in U$, so gilt $f = g + \text{Konstante}$.

Beweis. Betrachte $h(x) := f(x) - g(x)$. Es gilt $D_x(h) = D_x(f) - D_x(g) = 0$, somit ist h nach 1. konstant. \square

Beispiel 2.5.5 Wenn wir den Zusammenhang von U weglassen, gilt das Korollar nicht mehr, betrachte dafür z.B.

$$f : (0, 1) \cup (2, 3) \rightarrow \mathbb{R}, \quad x \mapsto \begin{cases} 1 & x \in (0, 1) \\ 2 & x \in (2, 3) \end{cases},$$

welche nicht konstant ist, für die aber $D_x f \equiv 0$ für alle x im Definitionsbereich gilt.

Für spätere Verwendung “verfeinern” wir das Konstanzkriterium noch etwas:

Korollar 2.5.6 (Konstanzkriterium II)

Seien $G \subset V_1 \times V_2$ offen und $f : G \rightarrow W$ differenzierbar. Es gelte für alle $(p_1, p_2) \in G$, dass

$$0 = D_{(p_1, p_2)} f(0, \cdot) : V_2 \rightarrow W.$$

Offenbar existiert zu jedem $(p_1, p_2) \in G$ ein $\varepsilon > 0$ mit

$$U_\varepsilon(p_1) \times U_\varepsilon(p_2) \subset G.$$

Dafür gilt dann

$$f(q_1, q_2) = f(q_1, r_2) \text{ für alle } q_1 \in U_\varepsilon(p_1) \text{ und } q_2, r_2 \in U_\varepsilon(p_2).$$

Mit andern Worten: Ist die V_2 -Ableitung von f Null, so hängt f lokal nicht von der V_2 -Variablen ab.

Beweis. Betrachte die Funktion

$$g : V_2 \supset U_\varepsilon(p_2) \rightarrow W, \quad t \mapsto f(q_1, t).$$

Für diese gilt nach der Kettenregel angewendet auf $t \mapsto (q_1, t) \mapsto f(q_1, t)$

$$D_t g(v_2) = D_{(q_1, t)} f(0, v_2) = 0,$$

und aus dem Schrankensatz angewendet auf g folgt $g(q_2) = g(r_2)$ und damit die Behauptung. \square

SATZ 2.5.2: TAYLOR-FORMEL

Sei $f : V \supset G \rightarrow W$ n -mal differenzierbar, (G eine offene Teilmenge und $p \in G$). Dann lässt sich f als

$$f(x) = \left(\sum_{k=0}^n \frac{1}{k!} D_p^k f(\underbrace{x-p, \dots, x-p}_{k\text{-mal}}) \right) + R(x) \quad (30)$$

darstellen, wobei für die dadurch definierte Restfunktion R gilt

$$\lim_{x \rightarrow p} \frac{R(x)}{\|x - p\|^n} = 0.$$

Zusatz: Ist f sogar $(n+1)$ -mal differenzierbar und reellwertig(!) und ist $\overline{px} \subset G$, so existiert $q \in \overline{px}$, so dass

$$R(x) = \frac{1}{(n+1)!} D_q^{n+1} f(\underbrace{x-p, \dots, x-p}_{(n+1)\text{-mal}}).$$

Beweis. Der Zusatz folgt direkt aus dem 1-dimensionalen Fall: Wir setzen $v := x - p$ und

$$g : [0, 1] \rightarrow \mathbb{R}, t \mapsto g(t) := f(p + tv).$$

Dann existiert $\tau \in [0, 1]$, so dass

$$\begin{aligned} f(x) &= g(1) \\ &= \left(\sum_{k=0}^n \frac{1}{k!} g^{(k)}(0) \right) + \frac{1}{(n+1)!} g^{(n+1)}(\tau) \\ &= \left(\sum_{k=0}^n \frac{1}{k!} \underbrace{\partial_v \dots \partial_v}_{k\text{-mal}} f(p) \right) + \frac{1}{(n+1)!} \underbrace{\partial_v \dots \partial_v}_{(n+1)\text{-mal}} f(p + \tau v) \\ &= \sum_{k=0}^n \frac{1}{k!} D_p^k f(\underbrace{x-p, \dots, x-p}_{k\text{-mal}}) + \frac{1}{(n+1)!} D_q^{n+1} f(\underbrace{x-p, \dots, x-p}_{(n+1)\text{-mal}}) \end{aligned}$$

mit $q = p + \tau v$.

Nun zum Beweis der ersten TAYLOR-Formel. Die kann man nicht einfach auf den eindimensionalen Fall zurückführen. Die Mehrdimensionalität von W ist dabei sekundär. Aber ein Ansatz wie oben führt nur zu Informationen über $\lim_{t \rightarrow 0} \frac{R(p+tv)}{\|tv\|^n}$, und das ist eine deutlich eingeschränkte Aussage.

Beweis durch vollständige Induktion über n .

$n = 1$. Das ist einfach die Definition der Differenzierbarkeit.

$(n-1) \rightarrow n$. Die Induktionsvoraussetzung angewendet auf die $(n-1)$ -mal differenzierbare Funktion Df liefert

$$D_x f = \left(\sum_{k=0}^{n-1} \frac{1}{k!} D_p^k (Df)(x-p, \dots, x-p) \right) + \tilde{R}(x) \quad (31)$$

mit

$$\lim_{x-p \rightarrow 0} \frac{\tilde{R}(x)}{\|x-p\|^{n-1}} = 0. \quad (32)$$

Nun berechnen wir die Ableitung von

$$R(x) = f(x) - \sum_{k=0}^n \frac{1}{k!} D_p^k f(x-p, \dots, x-p).$$

Der erste Term unter dem Summenzeichen ist $f(p)$, fällt bei der Differentiation also weg. Unter Benutzung der Produktregel, des Satzes von Schwarz und des Lemmas 2.4.9

$$\begin{aligned} D_x R(v) &= D_x f(v) - \sum_{k=1}^n \frac{1}{k!} k D_p^k f(x-p, \dots, x-p, v) \\ &= D_x f(v) - \sum_{k=1}^n \frac{1}{(k-1)!} D_p^{k-1} (Df)(x-p, \dots, x-p)(v). \end{aligned}$$

Also ist

$$D_x R = D_x f - \sum_{k=0}^{n-1} \frac{1}{k!} D_p^k (Df)(x-p, \dots, x-p) = \tilde{R}(x).$$

Zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$ mit $U_\delta(p) \subset G$, so dass für alle $x \in U_\delta(p)$

$$\|\tilde{R}(x)\| \leq \varepsilon \|x - p\|^{n-1} \quad (33)$$

Wegen $R(p) = 0$ folgt aus dem Schrankensatz dann für alle $x \in U_\delta(p)$

$$\|R(x)\| = \|R(x) - R(p)\| \leq \sup_{y \in U_\delta(p)} \|D_y R\| \|x - p\| \leq \varepsilon \|x - p\|^n,$$

also

$$\lim_{x \rightarrow p} \frac{R(x)}{\|x - p\|^n} = 0. \quad \square$$

Bemerkung 2.5.7 (Zum Zusatz) f muss reellwertig (d.h. $f(V) \subset \mathbb{R}$) sein, denn wenn man eine Aussage für in den \mathbb{R}^n haben will könnte man versuchen genau so vorzugehen und kriegt dann ein ξ für jede Komponente aber diese müssen nicht übereinstimmen.

2.6 Lokale Extrema mit und ohne Nebenbedingungen

- Wir wenden die TAYLOR-Formel auf Extremalprobleme an.

DEFINITION 2.6.1

Sei $l \in L^k(V, \mathbb{R})$ eine k -lineare Abbildung. Wenn der Zielraum \mathbb{R} ist, nennt man solche Abbildungen auch **k -Linearformen**, insbesondere für $k = 2$ **Bilinearformen**

Bilinearformen

men.

l heißt

- | | |
|---|--|
| <ol style="list-style-type: none"> ① positiv definit, wenn $l(v, \dots, v) > 0$ für alle $v \neq 0$. ② positiv-semidefinit, wenn $l(v, \dots, v) \geq 0$ für alle v. ③ negativ definit, wenn $l(v, \dots, v) < 0$ für alle $v \neq 0$. ④ negativ-semidefinit, wenn $l(v, \dots, v) \leq 0$ für alle v. ⑤ indefinit, wenn $v \mapsto l(v, \dots, v)$ das Vorzeichen wechselt. | <small>positiv definit</small>
<small>positiv-semidefinit</small>
<small>negativ definit</small>
<small>negativ-semidefinit</small>
<small>indefinit</small> |
|---|--|

Lemma 2.6.2

Sei $l \in L^k(V, \mathbb{R})$ eine **symmetrische** k -Linearform, d.h. es gelte

$$l(v_1, \dots, v_k) = l(v_{i_1}, \dots, v_{i_k})$$

für jede Permutation (i_1, \dots, i_k) von $(1, \dots, k)$. Dann gilt:

- ① Ist $l(v, \dots, v) = 0$ für alle $v \in V$, so ist $l = 0$.
- ② Ist k ungerade, so ist $l = 0$ oder l indefinit.

Beweis. Zu (i). Vollständige Induktion über k .

$k = 1$. Trivial.

$(k - 1) \rightarrow k$. Dann gilt

$$\begin{aligned} 0 &= l(tv + w, \dots, tv + w) \\ &= \sum_{i=0}^k \binom{k}{i} l(\underbrace{tv, \dots, tv}_{i-\text{mal}}, w, \dots, w) \\ &= \sum_{i=0}^k \binom{k}{i} t^i l(\underbrace{v, \dots, v}_{i-\text{mal}}, w, \dots, w). \end{aligned}$$

Ein Polynom verschwindet aber nur dann identisch, wenn alle Koeffizienten=0 sind. Daher ist insbesondere $l(v, \dots, v, w) = 0$ für alle v, w . Bei festem w ist $l(., \dots, ., w)$ symmetrisch und $(k - 1)$ -linear. Deshalb ist nach Induktionsvoraussetzung $l(v_1, \dots, v_{k-1}, w) = 0$ für alle v_i, w .

Zu (ii). existiert ein v mit $l(v, \dots, v) \neq 0$, etwa > 0 , so ist $l(-v, \dots, -v) < 0$ und l indefinit. Andernfalls ist $l = 0$ nach (i). □

DEFINITION 2.6.3

Die Funktion $f : V \supset G \rightarrow \mathbb{R}$ hat in p ein **strenges lokales Maximum**, wenn es ein $\varepsilon > 0$ gibt, so dass

$$\forall x \in G (0 < \|x - p\| < \varepsilon \implies f(x) < f(p)).$$

Analog erklärt man strenge lokale Minima.

SATZ 2.6.1: LOKALE EXTREMA

Sei $f : V \supset G \rightarrow \mathbb{R}$ k -mal differenzierbar, $k \geq 1$, $p \in G$ und

$$\begin{aligned} D_p f = 0, \dots, D_p^{k-1} f = 0, \\ D_p^k f \neq 0. \end{aligned}$$

Dann gilt:

- ① Ist $D_p^k f$ negativ definit, so hat f in p ein strenges lokales Maximum.
- ② Ist $D_p^k f$ positiv definit, so hat f in p ein strenges lokales Minimum.
- ③ Ist $D_p^k f$ indefinit, insbesondere k ungerade, so hat f in p kein lokales Extremum, sondern einen sogenannten Sattelpunkt.

Im semidefiniten Fall wird keine Aussage gemacht.

(iii) Aus (iii) folgt insbesondere die wichtige notwendige Bedingung:

Hat f in p ein lokales Extremum, so ist $D_p f = 0$.

Beweis. (Beweis des Satzes) Die Idee des Beweises ergibt sich aus der TAYLOR-Formel. Es gilt

$$f(x) - f(p) = \frac{1}{k!} D_p^k f(x - p, \dots, x - p) + R(x), \quad (34)$$

mit

$$\lim_{x \rightarrow p} \frac{R(x)}{\|x - p\|^k} = 0. \quad (35)$$

Also haben $f(x) - f(p)$ und $D_p^k f(x - p, \dots, x - p)$ dasselbe Vorzeichen, vorausgesetzt, man kann das Restglied vernachlässigen. Letzteres zu zeigen, ist das technische Problem des Beweises.

Die Einheitssphäre $S := \{v \in V \mid \|v\| = 1\}$ ist abgeschlossen und beschränkt, nach dem Satz von Heine-Borel (der in endlich-dimensionalen Banachräumen ebenso gilt wie im Standard- \mathbb{R}^n . Warum?) also kompakt. Daher existieren

$$m := \min_{v \in S} D_p^k f(v, \dots, v) \text{ und } M := \max_{v \in S} D_p^k f(v, \dots, v).$$

Für beliebiges $v \in V$ folgt daraus

$$m\|v\|^k \leq D_p^k f(v, \dots, v) \leq M\|v\|^k.$$

Ist $D_p^k f$ positiv definit, negativ definit oder indefinit, so ist $\varepsilon := \frac{1}{2k!} \min(|m|, |M|) > 0$. Wegen (35) existiert ein $\delta > 0$ mit $U_\delta(p) \subset G$ und

$$|R(x)| \leq \varepsilon \|x - p\|^k \text{ für alle } x \in U_\delta(p).$$

Zu (i). Ist $D_p^k f$ negativ definit, also $M < 0$, so folgt für $x \in U_\delta(p)$

$$\frac{1}{k!} D_p^k f(x - p, \dots, x - p) + R(x) \leq \frac{M}{k!} \|x - p\|^k - \frac{M}{2k!} \|x - p\|^k = \frac{M}{2k!} \|x - p\|^k \leq 0$$

mit Gleichheit nur für $x = p$. Aus (34) folgt die Behauptung (i). (i)

Zu (ii). Analog.

Zu (iii). Im indefiniten Fall ist $m < 0 < M$, und es gibt $v_1, v_2 \in S$ mit

$$D_p^k f(v_1, \dots, v_1) = m, \quad D_p^k f(v_2, \dots, v_2) = M.$$

Dann ist $x_j := p + \frac{\delta}{2}v_j \in U_\delta(p)$ und es gilt

$$\begin{aligned} \frac{1}{k!} D_p^k f(x_1 - p, \dots, x_1 - p) + R(x_1) &\leq \frac{m}{k!} \|x_1 - p\|^k - \frac{m}{2k!} \|x_1 - p\|^k = \frac{m}{2k!} \|x_1 - p\|^k < 0, \\ \frac{1}{k!} D_p^k f(x_2 - p, \dots, x_2 - p) + R(x_2) &\geq \frac{M}{k!} \|x_2 - p\|^k - \frac{M}{2k!} \|x_2 - p\|^k = \frac{M}{2k!} \|x_2 - p\|^k > 0. \end{aligned}$$

Aus (34) folgt die Behauptung (iii). (iii) □

Bemerkung. Die Frage, wann eine symmetrische k -Linearform zum Beispiel positiv definit ist, ist eine Frage an die (multi)lineare Algebra. Ein häufiger Spezialfall ist $k = 2$. Wir wollen überdies annehmen, dass $V = \mathbb{R}^n$ ist. Dann ist

$$D_p^2 f(u, v) = \sum_{i,j=1}^n \partial_i \partial_j f(p) u_i v_j.$$

Die (symmetrische) Matrix

$$H = \begin{pmatrix} \partial_1 \partial_1 f(p) & \dots & \partial_1 \partial_n f(p) \\ \vdots & & \vdots \\ \partial_n \partial_1 f(p) & \dots & \partial_n \partial_n f(p) \end{pmatrix}$$

der zweiten partiellen Ableitungen heißt auch die **Hessesche Matrix** von f . Für **Heisesche Matrix** gilt also

$$D_p^2 f(u, v) = \langle Hu, v \rangle$$

mit dem kanonischen Skalarprodukt $\langle u, v \rangle = \sum_{i=1}^n u_i v_i$. Es ist also eine interessante Frage, wann die durch eine symmetrische Matrix A gegebene Bilinearform $\langle Au, v \rangle$ positiv definit ist. In der Linearen Algebra lernt man (z.B. im Zusammenhang mit der Hauptachsentransformation), dass dies genau dann gilt, wenn alle Eigenwerte von A positiv sind. Dann nennt man auch A positiv definit. In der linearen Algebra lernt man auch, wie man die Eigenwerte bestimmt, und hat damit eine Methode, um im Fall $k = 2$ positive Definitheit nachzuprüfen.

Ein anderes Kriterium ist das folgende:

Lemma 2.6.4 (Hauptminorenkriterium)

Eine symmetrische $(n \times n)$ -Matrix

$$A = (a_{ij})_{i,j=1,\dots,n}$$

ist genau dann positiv definit, wenn alle **Hauptminoren** positiv sind. **Hauptminoren** sind **Hauptminoren** oder **Hauptabschnittsdeterminanten** die Determinanten **Hauptabschnittsdeterminanten** Matrizen

$$A_k := (a_{ij})_{i,j=1,\dots,k}$$

negativ definit ist genau dann **negativ definit**, wenn die Hauptminoren wechselndes Vorzeichen beginnend mit $a_{11} < 0$ haben.

Man findet dieses Kriterium oft in der Literatur zitiert (als Kriterium von Sylvester oder Hurwitz), aber selten bewiesen. Wir geben deshalb einen Beweis im Anhang.

Im Falle $n = 2$ ist die Hessematrix gegeben durch

$$\begin{pmatrix} \partial_x^2 f & \partial_y \partial_x f \\ \partial_x \partial_y f & \partial_y^2 f \end{pmatrix}$$

und wir erhalten folgendes Kriterium für lokale Extrema:

SATZ 2.6.2

Sei $f : \mathbb{R}^2 \supset G \rightarrow \mathbb{R}$ zweimal differenzierbar auf der offenen Menge G und sei $p \in G$. Dann gilt:

- (1) Hat f in p ein lokales Extremum, so ist $D_p f = 0$.
- (2) Ist $D_p f = 0$ und gilt

$$\partial_x^2 f(p) \partial_y^2 f(p) - (\partial_x \partial_y f(p))^2 > 0,$$

so hat f in p ein strenges lokales Extremum, und zwar

- ein Maximum, falls $\partial_x^2 f(p) < 0$,
- ein Minimum, falls $\partial_x^2 f(p) > 0$.

- (3) Ist

$$\partial_x^2 f(p) \partial_y^2 f(p) - (\partial_x \partial_y f(p))^2 < 0,$$

so hat f in p kein lokales Extremum. (Sattelpunkt)

Wir geben dafür noch einen direkten Beweis ohne weiteren Bezug auf die lineare Algebra:

Beweis. Wir bezeichnen die Hessematrix kurz mit

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

Dann ist

$$D_p^2 f \left(\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right) = ax^2 + 2bxy + cy^2 =: \varphi(x, y).$$

Wählt man $y = 0$, so sieht man, dass $a > 0$ bzw. $a < 0$ notwendig für die positive bzw. negative Definitheit ist. Die ist in diesem Fall dann aber äquivalent dazu, dass

$$0 < x^2 + 2\frac{b}{a}xy + \frac{c}{a}y^2 = \left(x + \frac{b}{a}y \right)^2 + \frac{c}{a}y^2 - \frac{b^2}{a^2}y^2 = \left(x + \frac{b}{a}y \right)^2 + \frac{ac - b^2}{a^2}y^2.$$

Wählt man nun $y \neq 0$ und $x = -\frac{b}{a}y$, so ergibt sich $ac - b^2$ als weitere notwendige Bedingung. Diese ist aber auch hinreichend: Die rechte Seite ist dann ≥ 0 , und verschwindet nur für $y = 0$ und $x = 0$. \square

2.7 Differentialoperatoren der klassischen Vektoranalysis

- Wir interpretieren Abbildungen - wie in der Physik - als Vektor- oder Skalarfelder.
- Gradient, Divergenz, Rotation und Laplaceableitung sind Felder, die mit Hilfe von Differentiationsprozessen aus anderen Feldern entstehen. Diese Operationen haben sich in der Physik als wichtig erwiesen.
- Wir lernen elementare Definitionen dieser Operationen im \mathbb{R}^n , bemühen uns aber auch um Definitionen in abstrakten Vektorräumen um zu klären, welche zusätzlichen Strukturen ggf. noch erforderlich sind.

DEFINITION 2.7.1

Sei $G \subset V$ eine offene Teilmenge des endlich-dimensionalen Banachraums V . Für diese Vorlesung vereinbaren wir folgende Sprechweisen:

- Ein **Vektorfeld** auf G ist eine Abbildung $X : G \rightarrow V$. **Vektorfeld**
- Ein **skalares Feld** auf G ist eine reellwertige Funktion $f : G \rightarrow \mathbb{R}$. **skalares Feld**

Wir wollen im folgenden eine kurze Einführung der klassischen Differentialoperatoren geben. Wir geben jeweils zwei Definitionen, eine elementare im \mathbb{R}^n und eine abstrakte, die etwas mehr lineare Algebra voraussetzt und die aufzeigt, welche „Hintergrundstrukturen“ in die Definition einfließen.

Gradient

- **Naiv.** Mit $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$ bezeichnen wir das kanonische Skalarprodukt auf \mathbb{R}^n . Der **Gradient** eines differenzierbaren skalaren Feldes $f : \mathbb{R}^n \supset G \rightarrow \mathbb{R}$ ist das folgende Vektorfeld:

$$\text{grad } f : G \rightarrow \mathbb{R}^n, p \mapsto \text{grad}_p f := (\partial_1 f(p), \dots, \partial_n f(p)).$$

Fundamentale Eigenschaften:

- ① $\langle \text{grad}_p f, v \rangle = D_p f(v)$ für alle $v \in V$.
- ② Der Gradient steht senkrecht auf den Niveaus von f . Genauer gilt für eine differenzierbare Kurve $c :]a, b[\rightarrow G$

$$f \circ c \text{ konstant} \iff \langle \text{grad}_{c(t)} f, \dot{c}(t) \rangle = 0 \text{ für alle } t. \quad (36)$$

Das folgt aus der Kettenregel, weil $\langle \text{grad}_{c(t)} f, \dot{c}(t) \rangle = D_{c(t)} f(\dot{c}(t)) = \frac{d}{dt} (f \circ c)$.

- ③ Der Gradient ist ein linearer Differentialoperator:

Für $\alpha, \beta \in \mathbb{R}$ und $f, g : G \rightarrow \mathbb{R}$ ist

$$\text{grad}(\alpha f + \beta g) = \alpha \text{grad } f + \beta \text{grad } g.$$

- ④ Der Gradient gibt die Richtung und Größe des stärksten Wachstums der Funktion f an:

Ist $\|v\| = 1$ und φ der Winkel zwischen dem Gradienten und der Richtung v , so ist

$$\partial_v f(p) = \|\operatorname{grad}_p f\| \cos \varphi.$$

- **Für Fortgeschrittene.** Ist $l : V \times V \rightarrow \mathbb{R}$ eine (nicht notwendig symmetrische) Bilinearform, so liefert

$$j_l : V \rightarrow V^* = L(V, \mathbb{R}), v \mapsto l(v, \cdot)$$

eine lineare Abbildung von V in V^* . Ist diese Abbildung ein Isomorphismus, so heißt l *nicht-degeneriert*. Ist l nicht-degeneriert, so kann man den l -Gradienten eines differenzierbaren skalaren Feldes $f : G \rightarrow \mathbb{R}$ definieren durch

$$\operatorname{grad}_p^l f := j_l^{-1}(D_p f),$$

d.h. durch die Gleichung

$$l(\operatorname{grad}_p^l f, v) = D_p f(v) \quad \text{für alle } v \in V.$$

Er ist ebenfalls ein linearer Differentialoperator und die obigen Eigenschaften (i), (ii) gelten mit l statt $\langle \cdot, \cdot \rangle$.

Beispiel 2.7.2 Euklidischer Gradient Seien $V = \mathbb{R}^n$ und $l(x, y) = \langle x, y \rangle = \sum x_i y_i$ das übliche Skalarprodukt. Das liefert den „naiven“ Gradienten wie oben. Allgemeiner Euklidis gibt es in jedem Euklidischen Vektorraum einen kanonischen Gradienten.

Vektor Beispiel 2.7.3 (Vierergradient) Sei $V = \mathbb{R}^4$ und

$$L(x, y) = x_1 y_1 + x_2 y_2 + x_3 y_3 - x_4 y_4$$

das sogenannte **LORENTZ-Produkt**. Der zugehörigen Gradient, der sogenannte **Vierergradient** ist gegeben durch

$$\operatorname{grad}^L f = (\partial_1 f, \partial_2 f, \partial_3 f, -\partial_4 f).$$

Er spielt – wie das LORENTZ-Produkt – eine große Rolle in der Relativitätstheorie.

Beispiel 2.7.4 (Symplektischer Gradient) Sei $V = \mathbb{R}^{2n}$ und

$$\sigma(x, y) = x_{n+1} y_1 + \dots + x_{2n} y_n - x_1 y_{n+1} - \dots - x_n y_{2n}$$

das sogenannte **symplektische Skalarprodukt**. Der entsprechende **symplektische Gradient** ist gegeben durch

$$\operatorname{grad}^\sigma f = (-\partial_{n+1} f, \dots, -\partial_{2n} f, \partial_1 f, \dots, \partial_n f).$$

Er spielt eine wichtige Rolle in der Hamilton-Jacobi-Theorie der klassischen Mechanik, vgl. Beispiel ??.

Divergenz

- **Naiv.** Sei $V = \mathbb{R}^n$ und $X = (X_1, \dots, X_n) : G \rightarrow \mathbb{R}^n$ ein differenzierbares Divergenz Vektorfeld. Dann ist die **Divergenz** von X das folgende skalare Feld:

$$\operatorname{div} X : G \rightarrow \mathbb{R}, p \mapsto \operatorname{div}_p X := \sum_{i=1}^n \partial_i X_i(p).$$

Beachte: $\operatorname{div} X$ ist gerade die Summe der Diagonalelemente der Jacobimatrix $(\partial_j X_i)$ von X .

- **Für Fortgeschrittene.** Ist V ein beliebiger endlich-dimensionaler Vektorraum und $X : G \rightarrow V$ ein differenzierbares Vektorfeld, so ist für p in G das Differential $D_p X$ ein Endomorphismus von V . Man definiert

$$\operatorname{div} X = \operatorname{Spur}(D_p X).$$

Der Satz von Gauß (Spezialfall des in der Analysis III zu beweisenden Stokesschen Integralsatzes) gibt eine Interpretation des Divergenz als "Quellstärke" des Feldes X . Das hat damit zu tun, dass die Spur die Ableitung der Determinante ist und die Determinante Volumina misst.

Rotation

- **Naiv.** Sei $V = \mathbb{R}^3$ und $X : G \rightarrow \mathbb{R}^3$ ein differenzierbares Vektorfeld. Die *Rotation* von X ist das folgende Vektorfeld:

$$\operatorname{rot} X : G \rightarrow \mathbb{R}^3$$

mit

$$\operatorname{rot}_p X := \begin{pmatrix} \partial_2 X_3(p) - \partial_3 X_2(p) \\ \partial_3 X_1(p) - \partial_1 X_3(p) \\ \partial_1 X_2(p) - \partial_2 X_1(p) \end{pmatrix}.$$

- **Für Fortgeschrittene.** Für zwei Vektoren $a, b \in \mathbb{R}^3$ ist das Vektorprodukt $a \times b$ charakterisiert durch die Bedingungen

(1) $a \times b = 0$, falls a, b linear abhängig,

und andernfalls

(2) $\|a \times b\| = \|a\| \|b\| \sin \angle(a, b)$,

(3) $\langle a \times b, a \rangle = \langle a \times b, b \rangle = 0$ und $(a, b, a \times b)$ ist eine positiv orientierte Basis des \mathbb{R}^3 .

Durch diese Bedingungen lässt sich ein Vektorprodukt in jedem orientierten 3-dimensionalen Euklidischen Vektorraum definieren.

Wir erklären nun zwei Methoden, um Achsrotationen in einem orientierten 3-dimensionalen Euklidischen Vektorraum V zu beschreiben. Die Achse sei gegeben durch einen Einheitsvektor u . Das Geschwindigkeitsfeld der Drehung in einem Punkt $x \in V$ muss dann senkrecht zu x und u stehen und mit dem Abstand von der Achse linear anwachsen.

Das wird geleistet

(1) durch ein Feld

$$x \mapsto \omega u \times x,$$

wo $\omega \in \mathbb{R}$ die sogenannte Winkelgeschwindigkeit bezeichnet, oder

(2) durch ein Feld

$$x \mapsto Ax,$$

wobei A ein schiefadjungierter (=schiefsymmetrischer) Endomorphismus von V mit $\operatorname{Kern}(A) = \mathbb{R} u$ ist. (Jeder schiefadjungierte Endomorphismus $\neq 0$ eines dreidimensionalen Raumes hat einen 1-dimensionalen Kern. Warum?)

Der Zusammenhang zwischen diesen beiden Methoden ist einfach: Für $a \in V$ ist $A : x \mapsto a \times x$ schiefadjungiert, weil

$$\langle a \times x, y \rangle = -\langle x, a \times y \rangle,$$

und die Abbildung

$$a \mapsto a \times \dots$$

liefert eine Isomorphismus(!) von V auf den Vektorraum der schiefadjungierten Endomorphismen von V .

Das Differential $D_p X$ eines differenzierbaren Vektorfeldes an der Stelle p ist im allgemeinen weder schiefsymmetrisch noch symmetrisch, aber man kann es in einen schiefsymmetrischen Anteil (= Rotationsanteil) und in einen symmetrischen Anteil zerlegen:

$$D_p X = \frac{1}{2}(D_p X - D_p X^*) + \frac{1}{2}(D_p X + D_p X^*).$$

(Der * bezeichnet die Adjungierte oder transponierte Matrix.) Dann gilt (Nachrechnen!)

$$\text{rot}_p X \times \dots = D_p X - D_p X^*.$$

In diesem Sinne ist die Rotation $\text{rot } X$ der doppelte Rotationsanteil von DX .

Es gibt eine Verallgemeinerung der Rotation auf Vektorräume beliebiger Dimension, aber nicht mehr für Vektorfelder, sondern für kompliziertere Objekte, die sogenannten Differentielformen vgl. (Analysis III).

SATZ 2.7.1

Für zweimal differenzierbare Felder gilt

$$\begin{aligned} \text{rot grad } f &= 0, \\ \text{div rot } X &= 0. \end{aligned}$$

Das gibt also notwendige Bedingungen dafür, dass sich ein differenzierbares Vektorfeld als Gradient (eines Potentials) oder Rotation (eines Vektorpotentials) schreiben lässt: die Rotation bzw. Divergenz muß verschwinden. Lokal, nicht aber global, sind diese Bedingungen auch hinreichend, vgl. Analysis III.

Beweis. Stures Nachrechnen unter Benutzung des Satzes von Schwarz über die Vertauschbarkeit der zweiten partiellen Ableitungen. \square

Laplaceoperator

Für zweimal differenzierbare skalare Felder auf $G \subset \mathbb{R}^n$ (oder in einem Euklidischen LAPLACEoperatorVektorraum) ist der LAPLACEoperator definiert durch

$$\Delta f = \text{div grad } f.$$

In Koordinaten bedeutet das

$$\Delta_p f = \sum_{i=1}^n \partial_i^2 f(p).$$

Funktionen mit $\Delta f = 0$ heißen **harmonische Funktionen**.

Der Laplaceoperator spielt eine fundamentale Rolle für die Beschreibung sehr vieler physikalischer Phänomene (Wärmeleitungsgleichung, Wellengleichung, Schrödingergleichung). Zum Beispiel ist die Amplitude f einer Welle in einem homogenen 3-dimensionalen Medium eine Funktion der Raumkoordinaten x_i und der Zeit t und genügt der Gleichung

$$\Delta_p f = \sum_{i=1}^3 \frac{\partial^2 f}{\partial x_i^2} = \frac{1}{c^2} \frac{\partial^2}{\partial t^2}.$$

Normiert man die Ausbreitungsgeschwindigkeit auf $c = 1$ und verwendet im \mathbb{R}^4 den Vierergradienten, so schreibt man den entsprechenden Laplaceoperator auch als $\Delta f := \operatorname{div} \operatorname{grad} f$, und die Wellengleichung wird einfach

$$\Delta f = 0.$$

Aufgaben

- ① [Aus den Übungsaufgaben vom Forster, 5D] Seien $U \subset \mathbb{R}^3$ eine offene Teilmenge und $v = (v_1, v_2, v_3) : U \rightarrow \mathbb{R}^3$ ein zweimal differenzierbares Vektorfeld. Verifizieren Sie, dass $\text{rot}(\text{rot}(v)) = \nabla(\text{div}(v)) - (\Delta v_1, \Delta v_2, \Delta v_3)$.
- ② [Forster, 5E] Seien $U \subset \mathbb{R}^n$ eine offene Teilmenge und $f, g : U \rightarrow \mathbb{R}^n$ zwei mal stetig partiell differenzierbare Funktionen. Zeigen Sie

$$\Delta(fg) = f\Delta g + 2\langle \nabla f, \nabla g \rangle + g\Delta f.$$



Lösungen

(1)

(2) 

3 Die großen Sätze der mehrdimensionalen Differentialrechnung

3.1 Der Umkehrsatz

- Weil die Ableitung einer differenzierbaren Abbildung diese lokal sehr gut approximiert, gibt sie zum Beispiel Auskunft auf die Frage nach der lokalen Umkehrbarkeit der Funktion.

Lemma 3.1.1

Für endlich-dimensionale Banachräume V, W gleicher Dimension sei

$$\text{Iso}(V, W) := \{A \in L(V, W) \mid A \text{ invertierbar}\}.$$

Dann gilt:

- (1) $\text{Iso}(V, W)$ ist offen in $L(V, W)$.
- (2) Die Inversenabbildung

$$\text{inv} : \text{Iso}(V, W) \rightarrow L(W, V), A \mapsto A^{-1}$$

ist differenzierbar mit

$$D_A \text{inv}(B) = -A^{-1}BA^{-1}$$

- (3) Für $A \in \text{Iso}(V, W)$ und $v \in V$ gilt

$$\|A(v)\| \geq \frac{1}{\|A^{-1}\|} \|v\|, \quad (37)$$

wobei $\|A^{-1}\|$ die Operatornorm bezeichnet.

(i) Beweis. Die Behauptungen (i), (ii) folgen aus dem Beispiel ?? mit Hilfe eines Isomorphismus $\Phi : V \rightarrow W$.

Zu (iii). Es ist

$$\|v\| = \|A^{-1}(A(v))\| \leq \|A^{-1}\| \cdot \|A(v)\|.$$

Daraus folgt (37). □

SATZ 3.1.1: UMKEHRSATZ

Seien $G \subset V$ offen und $f : V \supset G \rightarrow W$ stetig differenzierbar, d.h. Df existiert und ist stetig. Sei $p \in G$ und sei

$$D_p f : V \rightarrow W \text{ invertierbar.}$$

Dann ist f bei p lokal invertierbar mit stetig differenzierbarem Inversen.

Genauer: Es gibt eine offene Umgebung U von p in G , so dass gilt

- (1) $f|_U$ ist injektiv,
- (2) $f(U)$ ist offen in W ,

(3) $(f|_U)^{-1} : f(U) \rightarrow V$ ist stetig differenzierbar und für alle $x \in U$ gilt

$$D_{f(x)}(f|_U)^{-1} = (D_x f)^{-1}.$$

Bemerkung Aus der letzten Formel folgt: Ist f sogar k -mal stetig differenzierbar, so ist auch die lokale Umkehrung k -mal stetig differenzierbar.

DEFINITION 3.1.2

Eine k -mal stetig differenzierbare Abbildung mit einem k -mal stetig differenzierbaren Inversen heißt ein C^k -Diffeomorphismus. C^k -Diffeomorphismus

Eine stetig differenzierbare Abbildung mit invertierbarem Differential ist also lokal ein C^1 -Diffeomorphismus.

Beweis. (Beweis des Umkehrsatzes) Zu (i). Lokale Injektivität von f bei p .

Wir setzen $F := D_p f$ und $\beta := \frac{1}{\|F^{-1}\|}$.

Idee: Seien x, y nah bei p . Dann ist

$$\begin{aligned} \|f(y) - f(x)\| &= \|(f(y) - f(p)) - (f(x) - f(p))\| \\ &\approx \|D_p f(y-p) - D_p f(x-p)\| = \|F(y-x)\| \stackrel{(37)}{\geq} \beta \|y-x\|. \end{aligned} \tag{38}$$

Aus $y \neq x$ "folgt" dann also $f(y) \neq f(x)$.

Um das zu präzisieren, müssen wir das \approx -Zeichen quantitativ kontrollieren. Der Approximationsfehler ist

$$\|f(y) - f(x) - F(y-x)\| = \|(f(y) - F(y)) - (f(x) - F(x))\| = \|\varphi(y) - \varphi(x)\|$$

mit

$$\varphi(x) := f(x) - F(x).$$

Offenbar ist φ stetig differenzierbar und $D_p \varphi = D_p f - F = 0$. Also existiert ein $\delta > 0$, so dass

$$U = U_\delta(p) \subset G$$

und

$$\|D_\xi \varphi\| \leq \frac{\beta}{3} \text{ für alle } \xi \in U.$$

Dann ist nach dem Schrankensatz

$$\|\varphi(y) - \varphi(x)\| \leq \sup_{\xi \in U} \|D_\xi f - F\| \|y-x\| \leq \frac{\beta}{3} \|y-x\|.$$

Der Approximationsfehler in (38) ist also maximal $\frac{1}{3}$ der rechten Seite. Nach Dreiecksungleichung gilt (nachrechnen!) $\|f(x) - f(y)\| \geq \|F(y) - F(x)\| - \|\varphi(x) - \varphi(y)\|$, also ist nach der letzten Ungleichung aus (38)

$$\|f(y) - f(x)\| \geq \frac{2}{3} \beta \|y-x\| \tag{39}$$

und $f|_U$ injektiv. Nach dem Lemma ist weiter

$$D_x f \text{ invertierbar für } x \in U. \tag{40}$$

Zu (ii). Offenheit von $f(U)$.

Seien U wie oben und $x \in U$. Wir müssen zeigen, dass es ein $\varepsilon > 0$ gibt, so dass

$$U_\varepsilon(f(x)) \subset f(U).$$

Wähle zunächst $r > 0$ mit

$$K := \{y \in V \mid \|y - x\| \leq r\} \subset U.$$

Nach (39) gilt

$$\|y - x\| = r \implies \|f(y) - f(x)\| \geq \frac{2}{3}\beta r, \quad (41)$$

d.h. die Randpunkte von K werden durch f auf Punkte abgebildet, die mindestens den Abstand $\frac{2}{3}\beta r$ von $f(x)$ haben. Wir wollen zeigen, dass

$$U_{\frac{1}{3}\beta r}(f(x)) \subset f(K) \subset f(U). \quad (42)$$

Sei also $z \in U_{\frac{1}{3}\beta r}(f(x))$. Sei $y^* \in K$ ein Punkt, in dem die stetige Funktion $\|f(y) - z\|$ auf dem kompakten K ihr Minimum annimmt. Wir wollen zeigen, dass $f(y^*) = z$; dann ist (42) bewiesen.

Zunächst ist

$$\|y^* - x\| < r. \quad (43)$$

Sonst wäre nach Definition von K nämlich $\|y^* - x\| = r$, und nach (41) folgte mit der Dreiecksungleichung

$$\|f(y^*) - z\| \geq \frac{1}{3}\beta r.$$

Aber das steht wegen $\|z - f(x)\| < \frac{1}{3}\beta r$ im Widerspruch zur Wahl von y^* .

Wir nehmen nun an, dass

$$f(y^*) \neq z. \quad (44)$$

Wegen der Invertierbarkeit von $D_{y^*}f$ existiert dann ein $v \neq 0$ mit

$$D_{y^*}f(v) = z - f(y^*) \neq 0.$$

Geht man von y^* in Richtung v , so bleibt man für eine Weile in K , und das f -Bild bewegt sich in Richtung $z - f(y^*)$, also in Richtung auf z zu. Daher liegt für kleine positive t der Punkt $f(y^* + tv)$ näher an z als $f(y^*)$, und wir erhalten einen Widerspruch zur Wahl von y^* .

Wir präzisieren das:

Wähle $\delta_1 > 0$ so klein, dass $y^* + tv \in K$ für alle $t \in [0, \delta_1]$. Dann ist

$$\begin{aligned} f(y^* + tv) - z &= f(y^*) - z + tD_{y^*}f(v) + R(y^* + tv) \\ &= (f(y^*) - z)(1 - t) + R(y^* + tv). \end{aligned}$$

Es gibt ein $t \in]0, \delta_1[$, so dass

$$\|R(y^* + tv)\| \leq \frac{\|z - f(y^*)\|}{2\|v\|} \|tv\| = \frac{t}{2} \|z - f(y^*)\|,$$

also

$$\|f(y^* + tv) - z\| \leq (1-t)\|f(y^*) - z\| + \frac{t}{2}\|z - f(y^*)\| < \|f(y^*) - z\|$$

im Widerspruch zur Wahl von y^* . Damit war die Annahme (44) falsch, und es gilt $f(y^*) = z$, also $U_{\frac{1}{3}\beta r}(f(x)) \subset U$.

Zu (iii). Stetige Differenzierbarkeit der lokalen Umkehrabbildung.

Sei $g := (f|_U)^{-1} : f(U) \rightarrow V$. Seien $z, w \in f(U)$ und $x := g(z)$. Dann haben wir

$$\underbrace{f(g(w))}_{=w} = \underbrace{f(g(z))}_{=z} + D_x f(g(w) - g(z)) + R(g(w))$$

oder

$$D_x f(g(w) - g(z)) - (w - z) = -R(g(w)).$$

mit $\lim_{y \rightarrow x} \frac{R(y)}{\|y - x\|} = 0$. Wegen (40) ist $D_x f$ invertierbar. Es folgt

$$g(w) = g(z) + (D_x f)^{-1}(w - z) \underbrace{-(D_x f)^{-1}(R(g(w)))}_{=: \tilde{R}(w)}.$$

Wir wollen zeigen, dass

$$\lim_{w \rightarrow z} \frac{\tilde{R}(w)}{\|w - z\|} = 0. \quad (45)$$

Wegen der Injektivität von g und nach (39) gilt für $w \neq z$

$$0 < \|g(w) - g(z)\| \leq \frac{3}{2\beta} \|w - z\|.$$

Insbesondere ist g stetig, und aus

$$\frac{\tilde{R}(w)}{\|w - z\|} = -(D_x f)^{-1} \left(\underbrace{\frac{R(g(w))}{\|g(w) - g(z)\|}}_{\rightarrow 0 \text{ für } w \rightarrow z} \right) \underbrace{\frac{\|g(w) - g(z)\|}{\|w - z\|}}_{\leq \frac{3}{2\beta}}$$

folgt die Behauptung (45). Also ist g differenzierbar und

$$D_{f(x)} g = D_z g = (D_x f)^{-1}.$$

Schließlich ist $z \mapsto g(z) \mapsto D_{g(z)} f \mapsto (D_{g(z)} f)^{-1}$ als Komposition stetiger Abbildungen wieder stetig. Damit haben wir die stetige Differenzierbarkeit der Umkehrabbildung gezeigt. \square

Bemerkung Die Formel für die Ableitung folgt auch aus

$$(f|_U)^{-1} \circ f|_U = \text{id}$$

mit der Kettenregel:

$$D_{f(x)}(f|_U)^{-1} \circ D_x(f|_U) = D_x \text{id} = \text{id}.$$

Beispiel 3.1.3 Die Abbildung $f : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}^2$ mit $f(x, y) := (x^2 - y^2, 2xy)$ hat die Funktionalmatrix

$$f'(x, y) = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}.$$

Sie ist deshalb stetig differenzierbar und $D_{(x,y)}f$ ist für alle $(x, y) \in \mathbb{R}^2 \setminus \{0\}$ invertierbar. Also besitzt f um jeden Punkt lokal ein stetig differenzierbares Inverses. Aber f ist nicht global invertierbar, weil z.B. $f(1, 1) = f(-1, -1)$. Es ist

$$f(\{(x, y) \mid x > 0 \text{ und } y > 0\}) = \{(x, y) \mid y > 0\}$$

und

$$g(u, v) = \frac{1}{\sqrt{2}} \left(\sqrt{\sqrt{u^2 + v^2} + u}, \sqrt{\sqrt{u^2 + v^2} - u} \right), \quad v > 0$$

ist das Inverse von $f|_{\{(x, y) \mid x > 0 \text{ und } y > 0\}}$. Die Formel für die Ableitung der Inversen liefert

$$g'(f(x, y)) = (f'(x, y))^{-1} = \frac{1}{2(x^2 + y^2)} \begin{pmatrix} x & y \\ -y & x \end{pmatrix}.$$

Zum Beispiel ergibt sich für $x = y = 1$, also $(u, v) = (0, 2)$,

$$g'(0, 2) = \frac{1}{4} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Beispiel 3.1.4 (Stetige Polarkoordinaten) Die Polarkoordinaten in der Ebene sind nicht eindeutig, die Winkelkoordinate ist nur bis auf ein ganzzahliges Vielfaches von 2π bestimmt. Und wenn man die Eindeutigkeit mit ‘‘Gewalt’’ erzwingt, indem man zum Beispiel verlangt, dass $\varphi \in [-\pi, \pi[$, so wird die Winkelkoordinate auf der negativen x -Achse unstetig.

Wir wollen aber überlegen: Eine stetige Kurve $c : [a, b] \rightarrow \mathbb{R}^2 \setminus \{0\}$ kann man auch in Polarkoordinaten mit stetiger Winkelfunktion beschreiben: Ist $c(a) = \|c(a)\|(\cos \varphi_0, \sin \varphi_0)$, so existiert genau eine stetige Funktion $\Phi : [a, b] \rightarrow \mathbb{R}$ mit $\Phi(a) = \varphi_0$ und

$$\begin{aligned} c(t) &= \|c(t)\|(\cos \Phi(t), \sin \Phi(t)) \\ &= \|c(t)\|e^{i\Phi(t)} \text{ in komplexer Notation.} \end{aligned} \tag{46}$$

Wir betrachten die Abbildung

$$\begin{aligned} f : \mathbb{R}^2 &\supset G := \{(r, \varphi) \mid r > 0\} \rightarrow \mathbb{R}^2 \setminus \{0\} \\ (r, \varphi) &\mapsto (r \cos \varphi, r \sin \varphi) \end{aligned}$$

Dann ist

$$f'(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}.$$

Rechnen Sie nach, dass das für alle $(r, \varphi) \in G$ invertierbar ist. Also ist f nach dem Umkehrssatz lokal invertierbar. Wir wissen natürlich mehr: Die Abbildung f ist surjektiv auf $\mathbb{R}^2 \setminus \{0\}$, und mittels Arcus-Funktionen lassen sich lokale Umkehrabbildungen explizit hinschreiben. Weil das wegen der erforderlichen Fallunterscheidungen

mühsam ist, wählen wir nun zu jedem $p = (r, \varphi) \in G$ eine offene Umgebung U_p , die von f diffeomorph auf eine offene Menge $V_p := f(U_p) \subset \mathbb{R}^2 \setminus \{0\}$ abgebildet wird. Dann ist $(V_p)_{p \in G}$ eine offene Überdeckung von $\mathbb{R}^2 \setminus \{0\}$, und wegen der Stetigkeit von c ist $(c^{-1}(V_p))_{p \in G}$ eine offene Überdeckung von $[a, b]$. Nach dem Lebesgue-Lemma existiert eine Zerlegung

$$a = t_0 < t_1 < \dots < t_n = b,$$

so dass jedes $[t_{j-1}, t_j]$ in einem der $c^{-1}(V_p)$ enthalten ist. Wir wählen zu jedem j ein solches p , und schreiben $f_j := f|_{U_p}$.

Wir definieren nun rekursiv

$$\begin{aligned}\Phi(a) &:= \varphi_0 \\ \Phi(t) &:= (f_j^{-1}(c(t)))_2 - (f_j^{-1}(c(t_{j-1})))_2 + \Phi(t_{j-1}) \text{ für } t \in]t_{j-1}, t_j].\end{aligned}$$

Dabei bedeutet der untere Index $(.)_2$ die 2. Komponente (eben die φ -Komponente). Offenbar ist dann $\Phi|_{]t_{j-1}, t_j]}$ stetig, und weil außerdem

$$\lim_{t \searrow t_{j-1}} \Phi(t) = \Phi(t_{j-1}),$$

ist $\Phi : [a, b] \rightarrow \mathbb{R}$ stetig. Wir zeigen, dass (46) gilt. Nehmen wir an, dass das bereits für $t \leq t_{j-1}$ erfüllt ist. Dann folgt für $t_{j-1} < t \leq t_j$:

$$\begin{aligned}\|c(t)\|e^{i\Phi(t)} &= \|c(t)\| \exp\left(i\left((f_j^{-1}(c(t)))_2 - (f_j^{-1}(c(t_{j-1})))_2 + \Phi(t_{j-1})\right)\right) \\ &= \|c(t)\| \exp\left(i(f_j^{-1}(c(t)))_2\right) \frac{\|c(t_{j-1})\| \exp(i\Phi(t_{j-1}))}{\|c(t_{j-1})\| \exp\left(i(f_j^{-1}(c(t_{j-1})))_2\right)} \\ &= c(t) \frac{c(t_{j-1})}{c(t_{j-1})} = c(t).\end{aligned}$$

Zur Eindeutigkeit von Φ . Wir nehmen an, dass

$$\|c(t)\|e^{i\Phi(t)} = c(t) = \|c(t)\|e^{it\Phi(t)} \text{ für alle } t \in [a, b].$$

Dann folgt

$$e^{i(\Phi(t)-t\Phi(t))} = 1 \text{ für alle } t \in [a, b],$$

also

$$\Phi(t) - t\Phi(t) \in \{2k\pi \mid k \in \mathbb{Z}\} \text{ für alle } t \in [a, b].$$

Wenn Φ und $t\Phi$ stetig sind mit $\Phi(a) = \varphi_0 = t\Phi(a)$, so folgt daraus $\Phi = t\Phi$.

3.2 Implizite Funktionen

- Ist F linear, so ist $F(x, y) = F((x, 0) + (0, y)) = F(x, 0) + F(0, y)$ und die Frage, ob sich die Gleichung $F(x, y) = 0$ nach $y = y(x)$ auflösen lässt, ist einfach die Frage nach der Umkehrbarkeit von $F(0, \cdot)$. Wir lernen im Satz über implizite Funktionen die Antwort auf die entsprechende Frage für differenzierbares F .

Problem: Seien V_0, V_1, W endlich-dimensionale Banachräume und $f : V_0 \times V_1 \rightarrow W$.

Unter welchen Voraussetzungen hat die Gleichung

Voraussetzungen hat
die Gleichung

$$f(x, y) = 0 \quad (47)$$

für jedes $x \in V_0$ genau jedes $x \in V_0$ genau eine Lösung $y \in V_1$?

eine Lösung $y \in V_1$? Unter diesen Umständen existiert dann eine eindeutig bestimmte Funktion $g : V_0 \rightarrow V_1$, für die für alle $x \in V_0$ gilt

$$f(x, g(x)) = 0. \quad (48)$$

Man sagt dann auch, dass (47) sich nach einer Funktion $y = g(x)$ eindeutig auflösen implizit definiert oder dass g durch (48) implizit definiert wird.

Geometrisch bedeutet das, dass man das Niveau $f = 0$ als Graphen $\{(x, g(x)) \mid x \in V_0\}$ einer Funktion $g : V_0 \rightarrow V_1$ beschreibt, also durch V_0 parametrisiert: Jeder Punkt auf dem 0-Niveau liegt über genau einem Punkt von V_0 .

Im Fall $W = \mathbb{R}^m$ hat f die Komponentenfunktionen f_1, \dots, f_m . Man hat also m Gleichungen, die Dimension von W ist die Anzahl der gegebenen Gleichungen. Ebenso kann man die Dimension von V_1 als die Anzahl der gesuchten Unbekannten y_i ansehen. Es ist also wohl vernünftig, $\dim V_1 = \dim W$ zu wählen.

Beispiel 3.2.1 Sei $f = F : V_0 \times V_1 \rightarrow W$ linear und sei $\dim V_1 = \dim W$. Dann hat man

$$F(x, y) = F((x, 0) + (0, y)) = F(x, 0) + F(0, y),$$

d.h. F liefert zwei lineare Abbildungen

$$\begin{aligned} F(\cdot, 0) &: V_0 \rightarrow W, \\ F(0, \cdot) &: V_1 \rightarrow W. \end{aligned}$$

Dann ist (47) genau dann für jedes $x \in V_0$ eindeutig lösbar, wenn die lineare Abbildung $F(0, \cdot) : V_1 \rightarrow W$ invertierbar ist. Die Gleichung

$$0 = F(x, y) = F(x, 0) + F(0, y)$$

ist nämlich äquivalent zu

$$F(0, y) = -F(x, 0).$$

Das ist höchstens dann eindeutig lösbar, wenn $F(0, \cdot)$ injektiv ist. Nach der Dimensionsvoraussetzung ist in diesem Fall aber $F(0, \cdot)$ bijektiv und die Gleichung tatsächlich für jedes x eindeutig lösbar. Man findet

$$g(x) = -F(0, \cdot)^{-1}(F(x, 0)).$$

Im Fall $V_0 = \mathbb{R}^n$ und $V_1 = W = \mathbb{R}^m$ ist F gegeben durch eine $m \times (n+m)$ -Matrix der Form

$$\left(\underbrace{F^{(1)}}_n \mid \underbrace{F^{(2)}}_m \right),$$

und $F(0, \cdot)$ wird repräsentiert durch die quadratische $m \times m$ -Matrix $F^{(2)}$, die also invertierbar sein muss.

Wenn wir dieses Ergebnis von linearen Abbildungen auf differenzierbare Abbildungen verallgemeinern wollen, ist es plausibel, dass wir nur ein **lokales** Ergebnis erhalten. Experimentieren Sie ein bisschen mit dem Fall $V_0 = V_1 = W = \mathbb{R}$ und

$$f(x, y) := x - y^2.$$

SATZ 3.2.1: ÜBER IMPLIZITE FUNKTIONEN

Seien V_0, V_1, W endlich-dimensionale Banachräume, $G \subset V_0 \times V_1$ offen und $f : V_0 \times V_1 \supset G \rightarrow W$ stetig differenzierbar.

Sei $(p, q) \in G$ mit

$$f(p, q) = 0, \quad (49)$$

$$D_{(p,q)} f(0, \cdot) : V_1 \rightarrow W \text{ invertierbar.} \quad (50)$$

Beachte, dass damit $\dim V_1 = \dim W$.

Dann lässt sich

$$f(x, y) = 0 \quad (51)$$

in einer Umgebung von (p, q) eindeutig nach einer stetig differenzierbaren Abbildung $y = g(x)$ auflösen. Genauer:

Es gibt offene Umgebungen U_0 von p in V_0 und U_1 von q in V_1 mit folgenden Eigenschaften:

- ① $U_0 \times U_1 \subset G$, und zu jedem $x \in U_0$ existiert genau ein $y \in U_1$ mit

$$f(x, y) = 0.$$

- ② Die nach (i) eindeutig bestimmte Funktion $g : U_0 \rightarrow U_1$ mit

$$f(x, g(x)) = 0$$

ist stetig differenzierbar.

- ③ Für alle $x \in U_0$ ist $D_{(x,g(x))} f(0, \cdot) : V_1 \rightarrow W$ invertierbar und für $v \in V_0$ ist

$$D_x g(v) = - (D_{(x,g(x))} f(0, \cdot))^{-1} \circ D_{(x,g(x))} f(v, 0). \quad (52)$$

Bemerkung 3.2.2 Im Fall $V_0 = \mathbb{R}^n, V_1 = W = \mathbb{R}^m$ werden die linearen Abbildungen

$$D_{(x,y)} f(0, \cdot) : \mathbb{R}^m \rightarrow \mathbb{R}^m \quad \text{bzw.} \quad D_{(x,y)} f(\cdot, 0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

repräsentiert durch die Matrizen

$$\left(\frac{\partial f_i}{\partial y_j}(x, y) \right)_{i,j=1,\dots,m} \quad \text{bzw.} \quad \left(\frac{\partial f_i}{\partial x_j}(x, y) \right)_{i=1,\dots,m; j=1,\dots,n}$$

Beweis. (Beweis zum Satz über implizite Funktionen) Die Idee. Die Gleichung $f(x, y) = 0$ ist genau dann eindeutig nach y auflösbar, wenn dasselbe für die Gleichung

$$h(x, y) := (x, f(x, y)) = (x, 0)$$

gilt. h erweist sich nach dem Umkehrsatz als lokal invertierbar, und die gesuchte Lösungsfunktion g ist dann gegeben durch

$$(x, g(x)) = h^{-1}(x, 0),$$

also durch die zweite Komponente von $h^{-1}(., 0)$.

A. Vorbemerkung. Da $V_0 \times V_1$ endlich-dimensional ist, sind alle Normenäquivalent, und wir verwenden der Einfachheit halber die Norm

$$\|(v, w)\| = \sup(\|v\|, \|w\|).$$

Das hat den Vorteil, dass

$$U_\varepsilon((p, q)) = U_\varepsilon(p) \times U_\varepsilon(q) \text{ für } (p, q) \in V_0 \times V_1 \text{ und } \varepsilon > 0.$$

Analog verfahren wir gleich mit dem Raum $V_0 \times W$.

B. Reduktion auf den Umkehrsatz. Wir setzen die obige Beweisidee um und definieren die Abbildung

$$h : V_0 \times V_1 \supset G \rightarrow V_0 \times W, (x, y) \mapsto (x, f(x, y))$$

zwischen gleich-dimensionalen Vektorräumen. Es gilt

$$D_{(x,y)}h(v, w) = (v, D_{(x,y)}f(v, w)), \quad (53)$$

und deshalb ist mit f auch h stetig differenzierbar. Weiter ist

$$D_{(p,q)}h \text{ invertierbar}, \quad (54)$$

denn

$$\begin{aligned} 0 = D_{(p,q)}h(v, w) &\stackrel{(53)}{\iff} v = 0 \text{ und } D_{(p,q)}f(v, w) = 0 \\ &\iff v = 0 \text{ und } D_{(p,q)}f(0, w) = 0 \iff v = 0 \text{ und } w = 0 \end{aligned}$$

nach Voraussetzung.

C. Anwendung des Umkehrsatzes. Nach dem Umkehrsatz existiert $\varepsilon > 0$, so dass

$$\begin{aligned} U &:= U_\varepsilon(p) \times U_\varepsilon(q) \subset G, \\ h|U &\text{ injektiv,} \\ h(U) &\text{ offen,} \\ (h|U)^{-1} &\text{ stetig differenzierbar.} \end{aligned}$$

Da $h(U)$ offen und

$$(p, 0) = (p, f(p, q)) = h(p, q) \in h(U) \subset V_0 \times W,$$

existiert $\delta > 0$ mit $\delta < \varepsilon$ und

$$U_\delta(p) \times U_\delta(0) = U_\delta((p, 0)) \subset h(U) \subset V_0 \times W.$$

Wir setzen nun

$$U_0 := U_\delta(p), \quad U_1 := U_\varepsilon(q),$$

und behaupten, dass diese das Gewünschte leisten.

Zu (i). Zunächst ist

$$U_0 \times U_1 \subset U_\varepsilon(p) \times U_\varepsilon(q) = U \subset G. \quad (55)$$

Ist $x \in U_0$, so ist $(x, 0) \in U_\delta((p, 0)) \subset h(U)$. Darum existiert nach dem Umkehrsatz genau ein $(\tilde{x}, y) \in U$ mit $h(\tilde{x}, y) = (x, 0)$.

Offenbar ist

- $\tilde{x} = x$,
- $y \in U_1$ nach Definition von U und U_1 , und
- $f(x, y) = 0$ nach Definition von h .

Also existiert zu jedem $x \in U_0$ ein $y \in U_1$ mit $f(x, y) = 0$. Wir bezeichnen dieses y mit $g(x)$.

Sind $y_1, y_2 \in U_1$ mit $f(x, y_1) = 0 = f(x, y_2)$, so folgt $h(x, y_1) = (x, 0) = h(x, y_2)$, also $y_1 = y_2$.

Damit ist (i) bewiesen.

Zu (ii). Für $x \in U_0$ haben wir eben gezeigt, dass

$$h(x, g(x)) = (x, f(x, g(x))) = (x, 0).$$

Bezeichnen wir also mit $\pi : V_0 \times V_1 \rightarrow V_1, (v, w) \mapsto w$ die Projektion, so ist

$$g(x) = \pi \circ (h|U)^{-1}(x, 0).$$

Daher ist g stetig differenzierbar.

Zu (iii). Nach (54) ist $D_{(x,y)}h$ für $(x, y) \in U$ invertierbar. Nach (51) haben wir dann

$$D_{(x,y)}h(0, w) = (0, D_{(x,y)}f(0, w)) = 0 \iff w = 0.$$

Daraus folgt, dass für $(x, y) \in U$ auch $D_{(x,y)}f(0, .)$ injektiv und damit invertierbar ist. Insbesondere ist also für alle $x \in U_0$

$$D_{(x,g(x))}f(0, .) \text{ invertierbar.} \quad (56)$$

Nun differenzieren wir

$$\varphi : U_0 \ni x \xrightarrow{\alpha} (x, g(x)) \mapsto f(x, g(x))$$

nach der Kettenregel. Wir erhalten

$$\begin{aligned} D_x \varphi(v) &= D_{(x,g(x))}f(D_x \alpha(v)) \\ &= D_{(x,g(x))}f(v, D_x g(v)) \\ &= D_{(x,g(x))}f(v, 0) + D_{(x,g(x))}f(0, D_x g(v)). \end{aligned} \quad (57)$$

Andererseits ist $\varphi = 0$, also $D_x \varphi = 0$. Damit folgt aus (56) und (57) die Formel in (iii). \square

Beispiel 3.2.3 Seien $V_0 = V_1 = W = \mathbb{R}$ und $f : V_0 \times V_1 = G \rightarrow \mathbb{R}$, $(x, y) \mapsto x - y^2$. Dann ist

$$D_{(p,q)}f(v, w) = v - 2qw.$$

In $(p, q) = (0, 0)$ ist die Voraussetzung über die Invertierbarkeit der Ableitung also nicht erfüllt, wohl aber in allen Punkten (q^2, q) mit $q \neq 0$. In der Nähe dieser Punkte lässt sich $f^{-1}(\{0\}) = \{(x, y) | x - y^2 = 0\}$ also lokal als Graph schreiben.

3.3 Der Rangsatz

- Der Rangsatz beinhaltet in gewisser Weise die Quintessenz der linearen Approximation differenzierbarer Abbildungen.

In der Linearen Algebra betrachtet man folgendes Problem: Eine lineare Abbildung

$$F : V \rightarrow W$$

zwischen zwei \mathbb{R} -Vektorräumen der Dimensionen n und m kann man durch eine nachdem man $(m \times n)$ -Matrix darstellen, nachdem man in V und W Basen gewählt hat. Die und W Basen gewählt Darstellungsmatrix hängt wesentlich von den gewählten Basen ab, und man kann fragen, ob man sie durch geschickte Wahl der Basen besonders einfach gestalten kann. Tatsächlich kann man immer die folgende Form erreichen

$$\begin{pmatrix} 1 & \dots & 0 & 0 & \dots & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & & & \vdots \\ 0 & \dots & 1 & 0 & \dots & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & \dots & 0 \\ \vdots & & \vdots & \vdots & & & \vdots \\ 0 & \dots & 0 & 0 & \dots & \dots & 0 \end{pmatrix}.$$

Die Zahl der Einsen ist dabei der Rang r der linearen Abbildung, d.h. die Dimension von $F(V)$. Dieses Resultat kann man auch so formulieren:

Ist $F : V \rightarrow W$ wie oben, so existiert Isomorphismen $\Phi : V \rightarrow \mathbb{R}^n$ und $\Psi : W \rightarrow \mathbb{R}^m$, so dass

$$\Psi \circ F \circ \Phi^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

gegeben ist durch

$$\Psi \circ F \circ \Phi^{-1}(x_1, \dots, x_n) = (x_1, \dots, x_r, 0 \dots, 0).$$

Koordinaten Isomorphismen Φ und Ψ nennt man auch **Koordinaten**. In geeigneten Koordinaten sieht also jede lineare Abbildung vom Rang r aus wie

$$(x_1, \dots, x_n) \mapsto (x_1, \dots, x_r, 0 \dots, 0).$$

Wir übertragen das nun lokal auf C^k -Abbildungen, vgl. Definition ???. An die Stelle der linearen Koordinatenabbildung $\Phi : V \rightarrow \mathbb{R}^n$ tritt jetzt ein C^k -Diffeomorphismus

$$\Phi : V \supset U \rightarrow \tilde{U} \subset \mathbb{R}^n$$

(also eine bijektive C^k -Abbildung mit C^k -Inversem) zwischen offenen Umgebungen und U von $p \in G$ und \tilde{U} von $\Phi(p) = 0$ in \mathbb{R}^n und analog für Ψ . Diese Diffeomorphismen nennt man ebenfalls (krummlinige) **Koordinaten**.

SATZ 3.3.1: RANGSATZ

Seien V, W Banachräume der Dimensionen n und m , $G \subset V$ offen, und

$$f : V \supset G \rightarrow W \quad k\text{-mal stetig differenzierbar}, 1 \leq k \leq +\infty.$$

f sei von konstantem Rang r , d.h. der Rang von $D_x f : V \rightarrow W$ sei $= r$ unab-

hängig von $x \in G$. Dann gilt: Zu jedem $p \in G$ existiert C^k -Diffeomorphismen

$$\Phi : V \supset U_1 \rightarrow \tilde{U}_1 \subset \mathbb{R}^n$$

und

$$\Psi : W \supset U_2 \rightarrow \tilde{U}_2 \subset \mathbb{R}^m$$

offener Umgebungen von p bzw. $f(p)$ auf offene Umgebungen von $0 = \Phi(p)$ in \mathbb{R}^n bzw. von $\Psi(f(p)) = 0 \in \mathbb{R}^m$, so dass

$$f(U_1) \subset U_2$$

und

$$\Psi \circ f \circ \Phi^{-1}(x_1, \dots, x_n) = (x_1, \dots, x_r, 0, \dots, 0) \quad \text{für alle } x \in \tilde{U}_1.$$

C^k -Abbildungen von konstantem Rang r sind also in geeigneten C^k -Koordinaten von der Form

$$(x_1, \dots, x_n) \mapsto (x_1, \dots, x_r, 0, \dots, 0).$$

Konvention. Um die Notation übersichtlich zu halten, schreiben wir zum Beispiel:

Sei $g : V \rightarrow W$ ein lokaler Diffeomorphismus bei p ,

wenn g auf einer offenen Umgebung von $p \in V$ (nicht notwendig aber auf ganz V) definiert und C^k -differenzierbar ist, und eine (eventuell kleinere) offene Umgebung von p diffeomorph auf eine offene Umgebung von $g(p)$ in W abbildet.

Wegen des zu Beginn dieses Abschnittes angeführten Satzes aus der linearen Algebra genügt es, folgende Version des Rangsatzes zu beweisen:

SATZ 3.3.2: RANGSATZ, 2. VERSION

Seien V, W endlich-dimensionale Banachräume, sei $G \subset V$ offen, $p \in G$ und sei

$$f : V \supset G \rightarrow W \in C^k \quad \text{und von konstantem Rang } r.$$

Dann existiert lokale C^k -Diffeomorphismen

$$\varphi : V \rightarrow V \text{ bei } p \text{ mit } \varphi(p) = 0,$$

$$\psi : W \rightarrow W \text{ bei } f(p) \text{ mit } \psi(f(p)) = 0,$$

für die auf einer offenen Umgebung von $0 \in V$

$$\psi \circ f \circ \varphi^{-1} = D_p f$$

gilt. In geeigneten lokalen Koordinaten um p und $f(p)$ sieht f also aus wie seine Ableitung, d.h. wie eine lineare Abbildung vom Rang r .

Beweis. Vorbereitung. Durch Translationen (also C^∞ -Diffeomorphismen) in V und W können wir erreichen, dass $p = 0$ und $f(p) = 0$. Das setzen wir im folgenden voraus. Um den Beweis übersichtlich zu halten, benutzen wir die obige Konvention und verzichten auf die explizite Kontrolle der Definitionsbereiche.

Wir definieren

$$V_2 := \text{Kern } D_0 f, \quad W_1 := \text{Bild } D_0 f = D_0 f(V),$$

und wählen zu V_2 und W_1 komplementäre Unterräume, so dass also

$$V = V_1 \oplus V_2, \quad W = W_1 \oplus W_2.$$

Das Differential $D_0 f$ bildet dann also den r -dimensionalen Raum V_1 isomorph auf W_1 ab. Entsprechend der Zerlegung bezeichnen wir für $x \in V$ die Komponenten in V_1 bzw. V_2 mit x_1 bzw. x_2 und entsprechend für $y \in W$. Insbesondere ist $f = f_1 + f_2$ mit $f_i : G \rightarrow W_i$.

0. Schritt: Intuition: Die Formel $\psi \circ f \circ \varphi^{-1} = D_p f$ gilt, wenn wir $\psi := \text{id}$ und $\varphi := D_p f \circ f$ wählen. Jedoch passen die Definitions- und Zielbereich nicht so richtig und die Invertierbarkeit von $D_p f$ ist auch nicht gegeben. Also müssen wir φ ein bisschen „korrigieren“ und deswegen auch ψ ein wenig anders wählen.

1. Schritt: Konstruktion von φ . Die Komponentenabbildungen sind linear, und deshalb ist

$$F := D_0(f_1|_{V_1}) = (D_0 f|_{V_1})_1 : V_1 \rightarrow W_1$$

ein Isomorphismus. Daher ist nach dem Umkehrsatz

$$f_1|_{V_1} : V_1 \rightarrow W_1$$

ein lokaler Diffeomorphismus⁸. Dann ist auch $F^{-1} \circ f_1 : V_1 \rightarrow V_1$ ein lokaler Diffeomorphismus. Wir definieren $\varphi : V \rightarrow V$ durch

$$\varphi(x) := F^{-1}(f_1(x)) + x_2 \text{ für } x = x_1 + x_2.$$

Dann ist

$$D_0 \varphi(v_1 + v_2) = F^{-1}(D_0 f_1(v_1 + v_2)) + v_2 = v_1 + v_2,$$

also

$$D_0 \varphi = \text{id}_V, \tag{58}$$

und deshalb ist φ ein lokaler Diffeomorphismus.

Aus der Definition folgt für die V_1 -Komponente $\varphi_1(x) = F^{-1} \circ f_1(x)$, also $\varphi_1(\varphi^{-1}(x)) = F^{-1} \circ f_1(\varphi^{-1}(x))$ und

$$f_1(\varphi^{-1}(x)) = F(x_1). \tag{59}$$

2. Schritt: Konstruktion von ψ . Nun definieren wir $\psi : W \rightarrow W$ durch

$$\psi(y_1 + y_2) = y_1 + y_2 - f_2 \circ \varphi^{-1} \circ F^{-1}(y_1).$$

Dafür gilt

$$D_0 \psi(w_1 + w_2) = w_1 + \underbrace{w_2 - D_0(f_2 \circ \varphi^{-1} \circ F^{-1})(w_1)}_{\in W_2} = 0 \iff w_1 = 0 \text{ und } w_2 = 0.$$

⁸ Genauer: „... ein lokaler C^k -Diffeomorphismus bei 0.“, aber das unterdrücken wir in Zukunft: Alle unsere lokalen Diffeomorphismen und Abbildungen sind „bei 0“ und k -mal stetig differenzierbar.

Also $D_0\psi$ injektiv und damit bijektiv, und ψ ist ein lokaler Diffeomorphismus. Wir erhalten

$$\begin{aligned}\psi \circ f \circ \varphi^{-1}(x) &= \psi(f \circ \varphi^{-1}(x)) \\ &\stackrel{(59)}{=} \psi(F(x_1) + f_2 \circ \varphi^{-1}(x)) \\ &= F(x_1) + f_2 \circ \varphi^{-1}(x) - f_2 \circ \varphi^{-1} \circ F^{-1}(F(x_1)) \\ &= F(x_1) + f_2 \circ \varphi^{-1}(x) - f_2 \circ \varphi^{-1}(x_1)\end{aligned}$$

Wir sind also fertig, wenn wir zeigen können, dass für $x \in V$ nah bei 0

$$f_2 \circ \varphi^{-1}(x) = f_2 \circ \varphi^{-1}(x_1). \quad (60)$$

3. Schritt: Nachweis von (60). Das ist das eigentliche Herzstück des Beweises. Aus (59) folgt

$$\begin{aligned}D_x f(D_{\varphi(x)} \varphi^{-1}(v_1 + v_2)) &= D_{\varphi(x)}(f \circ \varphi^{-1})(v_1 + v_2) \\ &= D_{\varphi(x)}(f_1 \circ \varphi^{-1} + f_2 \circ \varphi^{-1})(v_1 + v_2) \\ &= F(v_1) + D_{\varphi(x)} f_2(v_1 + v_2).\end{aligned} \quad (61)$$

Wir betrachten nun die Projektion $\pi_1 : W \rightarrow W_1$, $y \mapsto y_1$. Aus der letzten Gleichung folgt

$$\pi_1(D_x f(V)) \supset \pi_1(D_x f(D_{\varphi(x)} \varphi^{-1}(V_1))) = F(V_1) = W_1.$$

Damit ist $\text{Rang}(D_x f) \geq \dim W_1 = r$ für alle Punkte x nah bei $p = 0$.

Gäbe es $v_2 \in V_2$ mit $D_{\varphi(x)}(f_2 \circ \varphi^{-1})(v_2) = w_2 \neq 0$, so wäre

$$w_2 \stackrel{(61)}{=} D_x f(D_{\varphi(x)} \varphi^{-1}(v_2)) \in D_x f(V) \quad \text{und} \quad \pi_1(w_2) = D_{\varphi(x)}(\pi_1 \circ f_2 \circ \varphi^{-1})(v_2) = 0.$$

Also wäre $w_2 \in \text{Kern } \pi_1|_{D_x f(V)}$ und nach Linearer Algebra

$$\dim D_x f(V) = \dim \text{Kern}(\pi_1) + \dim \text{Bild}(\pi_1) \geq r + 1$$

im Widerspruch zur Rangvoraussetzung über f , die wir hier zu ersten Mal benutzen.

Es folgt

$$D_{\varphi(x)}(f_2 \circ \varphi^{-1})|_{V_2} = 0,$$

d.h. $f_2 \circ \varphi^{-1}$ ist nach Korollar 2.5.6 lokal unabhängig von der V_2 -Komponenten und

$$f_2 \circ \varphi^{-1}(x) = f_2 \circ \varphi^{-1}(x_1). \quad \square$$

Wir halten noch ein Ergebnis aus diesem Beweis fest: Im letzten Schritt haben wir – ohne Benutzung der Konstanz des Ranges – gezeigt, dass für alle Punkte x nah bei p

$$D_p f(V) = W_1 \subset \pi_1(D_x f(V)),$$

Also ist der Rang von Df in Nachbarpunkten von p mindestens so groß wie in p .

Man sagt, er ist **unterhalbstetig**. Damit erhalten wir:

unterhalbstetig

Lemma 3.3.1

Für jede stetig differenzierbare Funktion ist der Rang unterhalb-stetig.

Beispiel 3.3.2 Auf der Menge der reellen invertierbaren $n \times n$ Matrizen betrachten wir die Abbildung

$$f : M(n \times n, \mathbb{R}) \supset \mathrm{GL}(n, \mathbb{R}) \rightarrow M(n \times n, \mathbb{R})$$

mit $f(A) = AA^T$, wobei A^T die transponierte Matrix bezeichnet. Dafür gilt

$$D_A f(B) = BA^T + AB^T,$$

und diese Matrix ist symmetrisch (=selbstadjungiert)! Ist andererseits $C \in M(n \times n, \mathbb{R})$ symmetrisch, so folgt

$$D_A f\left(\frac{1}{2}C(A^{-1})^T\right) = \frac{1}{2}C(A^{-1})^T A^T + \frac{1}{2}A(A^{-1})C^T = C.$$

Also ist für alle $A \in \mathrm{GL}(n, \mathbb{R})$ das Bild von $D_A f$ der $\frac{n(n+1)}{2}$ -dimensionale Raum aller symmetrischen Matrizen und f ist von konstantem Rang $\frac{n(n+1)}{2}$.

Korollar 3.3.3

Sei $f : V \subset G \rightarrow W$ stetig differenzierbar.

- (1) Ist f eine Immersion, d.h. $D_p f$ für alle p injektiv, so ist f lokal injektiv.
- (2) Ist f eine Submersion, d.h. $D_p f$ für alle p surjektiv, so ist f eine offene Abbildung, d.h. f bildet offene Mengen in offene Mengen ab.

Beweis. Selbst. □

4 Mannigfaltigkeiten

- Wir lernen mit den Mannigfaltigkeiten eine Verallgemeinerung des Flächenbegriffs auf beliebige Dimension (und Kodimension) kennen.
- Beispiele sind vor allem die “Niveaus” von Abbildungen, wie die höherdimensionalen Sphären, aber auch viel abstraktere Räume, wie etwa die orthogonalen Matrizen.
- Der Tangentialraum ist eine lineare Approximation der Mannigfaltigkeit und ermöglicht, auch für Funktionen auf Mannigfaltigkeiten die Ableitung als lineare Abbildung zu definieren.
- Als Anwendung behandeln wir Extrema unter Nebenbedingungen.

Eine m -dimensionale Mannigfaltigkeit im Banachraums V ist eine Teilmenge $M \subset V$, die in geeigneten krummlinigen Koordinaten (für V !) lokal so aussieht wie ein m -dimensionaler Untervektorraum:

DEFINITION 4.0.1

Seien $m, k \in \mathbb{N}$, $k > 0$. Eine Teilmenge $M \subset V$ eines n -dimensionalen Banachraums heißt eine m -dimensionale C^k -(Unter)mannigfaltigkeit, wenn es zu jedem Punkt $p \in M$ eine offene Umgebung U von p in V und einen C^k -Diffeomorphismus $\varphi : U \rightarrow \varphi(U)$ auf eine offene Teilmenge $\varphi(U) \subset \mathbb{R}^n$ gibt, so dass gilt:

$$M \cap U = \varphi^{-1}(\mathbb{R}^m \cap \varphi(U)),$$

d.h.

$$M \cap U = \{x \in U \mid \varphi_{m+1}(x) = \dots = \varphi_n(x) = 0\}. \quad (62)$$

Dabei betrachten wir also $\mathbb{R}^m \subset \mathbb{R}^n$ als den Unterraum aller Punkte, deren letzte $n - m$ Koordinaten verschwinden.

Eine große Klasse von Beispielen liefert der folgende

SATZ 4.0.1: GLEICHUNGSDEFINIERTE UNTERMANNIGFALTIGKEITEN

Seien V und W Banachräume endlicher Dimension. Seien $G \subset V$ offen und $g : G \rightarrow W \in C^k$, $k > 0$, vom konstanten Rang r , $0 < r < n := \dim V$ und $q \in g(G)$. Dann ist

$$M := g^{-1}(\{q\})$$

eine $n - r$ -dimensionale C^k -Mannigfaltigkeit.

Im Fall $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ bzw. $g : \mathbb{R}^3 \rightarrow \mathbb{R}$ ist M also eine Niveaumkurve bzw. -fläche. Gleichungsdefinierte Untermannigfaltigkeiten kann man also auch als **Niveaumannigfaltigkeiten** bezeichnen.

Beweis. Sei $p \in M$. Nach dem Rangsatz existiert C^k -Diffeomorphismen

$$\Phi : V \supset U_1 \rightarrow \tilde{U}_1 \subset \mathbb{R}^n$$

und

$$\Psi : W \supset U_2 \rightarrow \tilde{U}_2 \subset \mathbb{R}^m \quad (m = \dim W)$$

offener Umgebungen von p bzw. $q = g(p)$ auf offene Umgebungen von $\Phi(p) = 0$ in \mathbb{R}^n bzw. von $\Psi(q) = 0$ in \mathbb{R}^m , so dass

$$g(U_1) \subset U_2$$

und

$$\Psi \circ g \circ \Phi^{-1}(x_1, \dots, x_n) = (x_1, \dots, x_r, 0, \dots, 0) \quad \text{für alle } x \in \tilde{U}_1. \quad (63)$$

Dann gilt für $p' \in U := U_1$

$$\begin{aligned} p' \in M &\iff g(p') = q \\ &\iff \Psi(g(p')) = 0 \\ &\iff \Psi \circ g \circ \Phi^{-1} \circ \Phi(p') = 0 \\ &\stackrel{(63)}{\iff} \Phi_1(p') = \dots = \Phi_r(p') = 0. \end{aligned}$$

Bis auf die Nummerierung der Koordinatenfunktionen ist das die Definitionsgleichung (62). \square

Beispiel 4.0.2 Die Abbildung

$$g : \mathbb{R}^{n+1} \rightarrow \mathbb{R}, \quad (x_1, \dots, x_{n+1}) \mapsto \left(\sum_{i=1}^{n+1} x_i^2 \right) - 1$$

hat die Funktionalmatrix

$$g'(x_1, \dots, x_{n+1}) = 2(x_1, \dots, x_{n+1}),$$

und weil \mathbb{R} eindimensional ist, ist $D_x g$ surjektiv für alle $x \neq 0$. Daher ist die Einheitssphäre

$$S^n := \{x \mid g(x) = 0\} = \left\{ (x_1, \dots, x_{n+1}) \mid \sum x_i^2 = 1 \right\}$$

eine n -dimensionale C^∞ -Untermannigfaltigkeit des \mathbb{R}^{n+1} .

Beispiel 4.0.3 Wir betrachten im n^2 -dimensionalen Vektorraum der quadratischen n -reihigen Matrizen die Menge

$$\mathcal{O}(n) = \{A \in M(n \times n) \mid AA^t = E\}$$

orthogonalen Matrizen. Nach Beispiel ?? ist das eine C^∞ -Untermannigfaltigkeit der Dimension $n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}$. Die orthogonalen Matrizen bilden außerdem bezüglich der Matrixmultiplikation eine Gruppe. Die Gruppenoperationen sind Liegruppen offenbar differenzierbar und $\mathcal{O}(n)$ ist eine sogenannte Liegruppe.

DEFINITION 4.0.4 (TANGENTIALRAUM)

Sei M eine m -dimensionale Mannigfaltigkeit im n -dimensionalen Banachraum V , sei $p \in M$ und $\varphi : U \rightarrow \mathbb{R}^n$ ein Koordinatensystem dazu wie in der Definition ???. Dann ist also

$$M \cap U = \varphi^{-1}(\mathbb{R}^m \cap \varphi(U)),$$

und wir definieren den Tangentialraum $T_p M$ an M in p durch

$$T_p M := D_{\varphi(p)} \varphi^{-1}(\mathbb{R}^m).$$

Das ist also ein m -dimensionaler Vektorraum und eine lineare Approximation für M in der Nähe von p .

Auf dem nebenstehenden Bild ist eigentlich nicht $T_p M$ dargestellt, sondern der nach p verschobene Tangentialraum, weil das unserer anschaulichen Vorstellung eher entspricht. Zum Rechnen ist natürlich der Vektorunterraum $T_p M$ angenehmer als der parallele affine Unterraum.

Damit der Tangentialraum wohldefiniert ist, müssen wir zeigen, dass er nicht vom gewählten Koordinatensystem abhängt. Sei also $\tilde{\varphi} : \tilde{U} \rightarrow \mathbb{R}^n$ ein weiteres Koordinatensystem um p wie in der Definition ???. Wir können o.E. annehmen, dass $U = \tilde{U}$. Weil φ und $\tilde{\varphi}$ lokale Diffeomorphismen sind, folgt aus $\tilde{\varphi}^{-1}(\mathbb{R}^m \cap \varphi(U)) = M \cap U = \varphi^{-1}(\mathbb{R}^m \cap \varphi(U))$, dass

$$\tilde{\varphi} \circ \varphi^{-1}(\mathbb{R}^m \cap \varphi(U)) \subset \mathbb{R}^m$$

und deshalb

$$D_p \tilde{\varphi} \circ D_{\varphi(p)} \varphi^{-1}(\mathbb{R}^m) \subset \mathbb{R}^m,$$

also

$$D_{\varphi(p)} \varphi^{-1}(\mathbb{R}^m) \subset (D_p \tilde{\varphi})^{-1}(\mathbb{R}^m) = D_{t\varphi(p)}(t\varphi^{-1})(\mathbb{R}^m).$$

Durch Vertauschen von φ und $\tilde{\varphi}$ ergibt sich die umgekehrte Inklusion, also Gleichheit der Räume.

Beispiel 4.0.5 Ist $M = g^{-1}(\{q\}) \subset V$ eine gleichungsdefinierte Untermannigfaltigkeit wie im Satz 4.0.1, so gilt für $p \in M$ und ein Koordinatensystem $\varphi : U \rightarrow \mathbb{R}^n$ um p , dass $M \cap U = \varphi^{-1}(\mathbb{R}^m \cap \varphi(U))$, also $g \circ \varphi^{-1}(\mathbb{R}^m \cap \varphi(U)) = \{q\}$ und daher

$$D_p g(T_p M) = D_p g(D_{\varphi(p)} \varphi^{-1}(\mathbb{R}^m)) = 0.$$

Weil der Rang von g aber gerade $\dim V - \dim M$ ist, folgt

$$T_p M = \text{Kern } D_p g = (D_p g)^{-1}(\{0\}).$$

Das ist die linearisierte Version von

$$M = g^{-1}(\{q\}).$$

Auf Mannigfaltigkeiten kann man ‘‘Analysis betreiben’’, insbesondere die Differenzierbarkeit von Funktionen erklären. Das Differential an einer Stelle $p \in M$ ist dann eine lineare Abbildung auf dem Tangentialraum $T_p M$.

Wir betrachten dazu nur ein

Beispiel 4.0.6 (Extrema auf Mannigfaltigkeiten) Seien $G \subset V$ offen und $M \subset G$ ein Mannigfaltigkeit. Sei $f : G \rightarrow \mathbb{R}$ eine differenzierbare Funktion. Wir suchen lokale Extrema der Funktion $f|_M : M \rightarrow \mathbb{R}$. Sei $p \in M$ und $\varphi : U \rightarrow \mathbb{R}^n$ ein Koordinatensystem um p wie in der Mannigfaltigkeitsdefinition, $\varphi(p) = 0$. Dann ist $M \cap U = \varphi^{-1}(\mathbb{R}^m)$. Hat also $f|_M$ in p ein lokales Extremum, so hat $f \circ \varphi^{-1}|_{\mathbb{R}^m \cap \varphi(U)}$ in 0 ein lokales Extremum. Deshalb ist

$$D_0(f \circ \varphi^{-1})(\mathbb{R}^m) = D_p f(T_p M) = 0. \quad (64)$$

Notwendig für lokale Extrema der Einschränkung $f|_M$ von f ist also das Verschwinden der Einschränkung der Ableitung auf den Tangentialraum an M .

Ist $M = g^{-1}(\{q\})$ gleichungsdefiniert, so bedeutet (64), dass

$$\text{Kern } D_p g \subset \text{Kern } D_p f.$$

Die im Beispiel zuletzt betrachtete Situation ist unter dem Namen **Extremwerte** Extremwerte **unter Nebenbedingungen** berühmt. Sei $g : V \supset G \rightarrow W$ stetig differenzierbar und Nebenbedingung $q \in g(G)$. Sei weiter $f : G \rightarrow \mathbb{R}$ differenzierbar. Wir suchen lokale Extrema der Funktion f unter der Nebenbedingung $g = q$, d.h. lokale Extrema von $f|_{g^{-1}(\{q\})}$. Die Menge

$$\tilde{G} := \{p \in G \mid D_p g \text{ ist surjektiv}\}$$

ist nach Lemma ?? eine offene Teilmenge und $g^{-1}(\{q\}) \cap \tilde{G}$ eine Mannigfaltigkeit M der Dimension $\dim V - \dim W$. Hat $f|_{g^{-1}(\{q\})}$ ein lokales Extremum in $p \in M$, so gilt dort also die notwendige Bedingung (65). Typischerweise ist in den Anwendungen die Menge $g^{-1}(\{q\}) \setminus M$ der sogenannten singulären Punkte eine endliche Punktmenge, die man dann noch gesondert untersuchen muss.

Wir geben noch eine Variante von (65), die für die explizite Berechnung lokaler Extrema unter Nebenbedingungen hilfreich ist:

Es ist ein Standardproblem der lineare Algebra, den Kern einer linearen Abbildung zu bestimmen, also zu prüfen, ob (65) gilt. Aber meistens kennt man p gar nicht, sondern will die Extremalstellen erst finden. Das führt in der Regel auf nicht-lineare Gleichungssysteme, die schwer zu lösen sind. Bei der Bestimmung der Punkte vom zweiten Typ ist aber das folgende Lemma hilfreich:

Lemma 4.0.7 (Lagrange-Multiplikatoren)

Sei G offen in $V = \mathbb{R}^n$ und seien $f : G \rightarrow \mathbb{R}$ und $g = (g_1, \dots, g_m) : G \rightarrow \mathbb{R}^m$ differenzierbar bzw. stetig differenzierbar.

Dann ist (65) äquivalent dazu, dass es reelle Zahlen $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ gibt (sog. Lagrange-Multiplikatoren), so dass für alle $j \in \{1, \dots, n\}$

$$\partial_j f(p) = \sum_{i=1}^m \lambda_i \partial_j g_i(p). \quad (66)$$

Beweis. Bezeichnen wir die Funktionalmatrizen mit $f'(p)$ bzw. $g'(p)$, die Transposition mit $(\dots)^T$ und setzen wir $\lambda := (\lambda_1, \dots, \lambda_n)$, so ist (66) äquivalent zu

$$f'(p) = \lambda g'(p) \quad \text{oder} \quad f'(p)^T = g'(p)^T \lambda^T.$$

Dieses lineare Gleichungssystem ist genau dann lösbar, wenn die erweiterte Matrix $(g'(p)^T, f'(p)^T)$ denselben Rang wie $g'(p)^T$ hat, wenn also die Matrix $\begin{pmatrix} g'(p) \\ f'(p) \end{pmatrix}$ denselben Rang wie die Matrix $g'(p)$ hat. Weil beide dieselbe Anzahl von Spalten haben, ist das genau dann der Fall, wenn die Kerne dieser beiden Matrizen gleiche Dimension haben. Weil aber

$$\text{Kern } g'(p) \supset \text{Kern } \begin{pmatrix} g'(p) \\ f'(p) \end{pmatrix} = \text{Kern } g'(p) \cap \text{Kern } f'(p),$$

ist das genau dann der Fall, wenn $\text{Kern } D_p g \subset \text{Kern } D_p f$. □

Rezept. Zur Bestimmung der Kandidaten p für Stellen lokaler Extrema von

$$f : \mathbb{R}^n \supset G \rightarrow \mathbb{R}$$

unter der Nebenbedingungen $g = 0$ mit $g : G \rightarrow \mathbb{R}^m$ sucht man

-
- ① alle Punkte p mit $g(p) = 0$, in denen $D_p g(\mathbb{R}^n) \neq \mathbb{R}^m$ (singuläre Punkte),
 ② alle Lösungen p, λ von

$$g_1(p) = 0, \dots, g_m(p) = 0,$$

$$\partial_j f(p) = \sum_{i=1}^m \lambda_i \partial_j g_i(p), \quad j = 1, \dots, m.$$

Das sind $m + n$ Gleichungen für die $n + m$ Variablen $p_1, \dots, p_n, \lambda_1, \dots, \lambda_m$.

Die λ 's kann man wieder vergessen.

In typischen Problemen ist $m < n$, und die so gefundene Kandidatenmenge diskret oder sogar endlich.

Beispiel 4.0.8 Wir betrachten das Problem,

$$f(x, y, z) = xyz$$

unter der Nebenbedingung

$$x^2 + y^2 + z^2 \leq 1$$

zu maximieren, also das größte achsenparallele Quader in der Einheitskugel B zu finden. (Dessen Volumen ist dann $8|xyz|$, vgl. Abbildung.)

Beachten Sie, dass hier die Nebenbedingung durch eine [Ungleichung](#) gegeben ist. Die Kugel B ist kompakt, und weil f stetig ist, nimmt es auf B sein Maximum an. Das kann nicht in einem inneren Punkt geschehen, weil wir sonst alle Seiten des Quaders ein wenig vergrößern können und immer noch in der Kugel B bleiben:

Wenn $x^2 + y^2 + z^2 < 1$, ist auch $|x|^2 + |y|^2 + |z|^2 < 1$ und es gibt $\varepsilon > 0$ mit

$$(|x| + \varepsilon)^2 + (|y| + \varepsilon)^2 + (|z| + \varepsilon)^2 < 1.$$

Dafür ist aber

$$f(|x| + \varepsilon, |y| + \varepsilon, |z| + \varepsilon) = (|x| + \varepsilon)(|y| + \varepsilon)(|z| + \varepsilon) > |xyz| \geq xyz = f(x, y, z).$$

Ein anderes Argument liefert dasselbe Ergebnis: Läge das Maximum in einem inneren Punkt (x, y, z) so wäre

$$f'(x, y, z) = (yz \ xz \ xy) = (0 \ 0 \ 0).$$

Dann wäre aber $f(x, y, z) = 0$ das Maximum. Jedoch nimmt f offenbar auch positive Werte an.

Also wird das Maximum auf dem Rand angenommen, ist also ein Maximum unter der Nebenbedingung

$$g(x, y, z) := x^2 + y^2 + z^2 - 1 = 0.$$

Die Funktionalmatrix

$$g'(x, y, z) = (2x \ 2y \ 2z)$$

ist $\neq (0 \ 0 \ 0)$ für alle Punkte, die die Nebenbedingung erfüllen. Daher existiert keine singulären Punkte.

Wir lösen nach dem Rezept:

$$x^2 + y^2 + z^2 - 1 = 0,$$

und

$$\begin{aligned}yz &= \lambda 2x, \\xz &= \lambda 2y, \\xy &= \lambda 2z.\end{aligned}$$

Multipliziert man diese letzteren Gleichungen mit x, y, z , addiert und verwendet die Nebenbedingung, so hat man

$$3xyz = 2\lambda.$$

Einsetzen von λ in die obigen Gleichungen liefert

$$x^2 = y^2 = z^2 = \frac{1}{3}$$

oder zwei Koordinaten sind 0, die dritte dann wegen der Nebenbedingung ± 1 . Die letzteren Punkte liefern aber $f = 0$ und scheiden daher für ein Extremum aus. Mögliche Extrema liegen also in den Punkten

$$(\pm \frac{1}{\sqrt{3}}, \pm \frac{1}{\sqrt{3}}, \pm \frac{1}{\sqrt{3}})$$

mit voneinander unabhängigen Vorzeichen. Die entsprechenden Funktionswerte sind

$$f = \pm \frac{1}{3\sqrt{3}}.$$

Die positiven sind die Maxima, die negativen die Minima.

Als eine weiteres Anwendung für die Methode der Lagrange-Multiplikatoren beweisen wir im nächsten Beispiel die früher behauptete Abschätzung der l^p -Normen gegeneinander, vgl. Beispiel ??.

Beispiel 4.0.9 Wir erinnern an die Definition der l^p -Norm auf \mathbb{R}^n :

$$\|x\|_p := \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}$$

Wir zeigen: Für $1 \leq p \leq q$ und alle $x \in \mathbb{R}^n$ gilt

$$\|x\|_q \leq \|x\|_p \leq n^{\frac{1}{p} - \frac{1}{q}} \|x\|_q.$$

(67)

Wir zeigen das durch vollständige Induktion über n .

$n = 1$. Trivial.

$n - 1 \implies n$. Es genügt zu zeigen: Für alle $x = (x_1, \dots, x_n)$ gilt

$$\sum_{i=1}^n |x_i|^q = 1 \implies 1 \leq \|x\|_p \leq n^{\frac{1}{p} - \frac{1}{q}}.$$

Die Voraussetzung impliziert $|x_i| \leq 1$ und deshalb $|x_i|^q \leq |x_i|^p$ für alle i . Also folgt die linke Ungleichung, wir müssen nur noch die rechte beweisen. Offenbar können wir uns dabei auf die kompakte Menge

$$\left\{ x = (x_1, \dots, x_n) \mid \sum_{i=1}^n x_i^q = 1 \text{ mit } x_i \geq 0 \text{ für alle } i \right\}$$

beschränken. Ist wenigstens ein $x_i = 0$, so liegt x in einem \mathbb{R}^{n-1} . Nach Induktionsvoraussetzung gilt für solche x also

$$\|x\|_p \leq (n-1)^{\frac{1}{p}-\frac{1}{q}} \leq n^{\frac{1}{p}-\frac{1}{q}}.$$

Daher genügt es zu zeigen, dass die differenzierbare Funktion $f(x) := \sum_{i=1}^n x_i^p$ auf der Menge $\{x = (x_1, \dots, x_n) \mid x_i > 0\}$ unter der Nebenbedingung

$$g(x) := \sum_{i=1}^n x_i^q = 1$$

das Maximum $\left(n^{\frac{1}{p}-\frac{1}{q}}\right)^p$ besitzt. Die notwendige Bedingung für ein Extremum ist die Existenz eines λ mit

$$\frac{\partial f}{\partial x_j} = px_j^{p-1} = \lambda \frac{\partial g}{\partial x_j} = \lambda qx_j^{q-1}$$

oder

$$x_j^{p-q} = \lambda \frac{q}{p}$$

für alle j . Daraus folgt $x_1 = \dots = x_n$, nach der Nebenbedingung also $x_1 = \dots = x_n = \frac{1}{n^{1/q}}$. Der Funktionswert an dieser Stelle ist

$$f(x) = n \frac{1}{n^{p/q}} = n^{1-\frac{p}{q}} = \left(n^{\frac{1}{p}-\frac{1}{q}}\right)^p$$

und damit das (eindeutig bestimmte) Maximum von f .

5 Differentialgleichungen

5.1 Lineare Differentialgleichungen

17.07.2018 Seien V ein n -dimensionaler BANACH-Raum, $I \subset \mathbb{R}$ ein Intervall, $g : I \rightarrow V$ und $F : I \rightarrow L(V, V)$ stetige Abbildungen.

Wir betrachten die lineare Differentialgleichung

$$\dot{x}(t) = F(t)x(t) + g(t), \quad t \in I, \quad (68)$$

d.h. wir suchen alle Lösungen $x \in \mathcal{C}^1(I, V)$ der Gleichung (68).

Die dazugehörige **homogene Differentialgleichung** erhalten wir durch $g \equiv 0$. Sie wird durch alle $x \in \text{Ker}(L)$ gelöst, wobei L der *lineare* Operator ist, der durch

$$L : \mathcal{C}^1(I, V) \rightarrow \mathcal{C}^0(I, V), \quad (Lx)(t) := \dot{x}(t) - F(t)x(t), \quad t \in I \quad (69)$$

gegeben ist.

Da der Kern einer linearen Abbildung immer ein Vektorraum ist, hat unsere Lösungsmenge nun mehr 'Struktur'.

Also sind alle Lösungen der **inhomogenen** ($g \neq 0$) von (68) von der Form

partikuläre Lösung $x + x_p$ mit $x \in \text{Ker}(L)$, x_p ist eine **partikuläre Lösung**

Die partikuläre Lösung ist irgendeine Lösung von (68). Also gilt

Alle Lösungen der inhomogenen DGL sind die Summe einer partikulären Lösung und einer Lösung des homogenen DGL.⁹

5.2 Der Hauptsatz über lineare Differentialgleichungen

💡 Problem

Finde eine Abbildung $x : I \rightarrow \mathbb{R}^n \in \mathcal{C}^1$, welche (68) erfüllt.

Wir suchen stets das maximal 'Lösungsintervall' I , welche in diesem Fall (68) nach dem folgenden Satz stets durch $I = J$ gegeben ist.

SATZ 5.2.1: HAUPTSATZ ÜBER LINEARE DIFFERENTIALGLEICHUNGEN (EXISTENZ & EINDEUTIGKEIT)

- (1) Das Anfangswertproblem (68) ist zusammen mit einer Anfangswertbedingung $x(t_0) = x_0$ auf I eindeutig lösbar. Anfangswertbedingung
- (2) Es gilt $\dim(\text{Ker}(L)) = \dim(V)$ für den Lösungsraum der homogenen Differentialgleichung. Für alle Lösungen $x_1, \dots, x_n \in \text{Ker}(L)$ gilt^a

$$x_1, \dots, x_n \text{ linear unab. } \iff x_1(t), \dots, x_n(t) \text{ linear unab. } \forall t \in I$$

- (3) Sind (x_1, \dots, x_n) eine Basis von $\text{Ker}(L)$, also ein Fundamentalsystem der Differentialgleichung, so ist jede Lösung des obigen Anfangswertproblems von der Form

$$\sum_{i=1}^n c_i(t)x_i(t) \quad \text{mit } c_i \in \mathcal{C}^1(I, \mathbb{R}) \quad \forall i \in \{1, \dots, n\}.$$

Solche Funktionen lösen das Anfangswertproblem, wenn $\sum_{i=1}^n \dot{c}_i(t)x_i(t) = g(t)$ mit $t \in I$.

^aDie lineare Unabhängigkeit der Lösungen wird in dem obigen nicht besonders konkret definiert. Konkret ist eine Aussage von (2):

$$\left(\sum_i c_i y_i(t) = 0 \quad \forall t \iff c_i = 0 \quad \forall i \right) \iff \left(\sum_i c_i y_i(t) = 0 \text{ für ein } t \iff c_i = 0 \quad \forall i \right),$$

da bei ist die Richtung \implies klar, die Rückrichtung ist die Aussage.

Beispiel 5.2.1 (ohne Anfangswertbedingung keine eindeutige Lösung)

Seien $F : \mathbb{R} \rightarrow \mathbb{R}$, $t \mapsto 0$ sowie $g(t)$ stetig mit $g|_{\mathbb{R} \setminus [-1, 1]} \equiv 0$. Betrachte die Differentialgleichung

$$\dot{x} = F(t) \cdot x + g(t) = g(t)$$

und dessen Lösung $G(t) := \int_0^t g(x)dx$, welche wegen der Konstante der Integration nicht eindeutig ist. Ist eine Anfangswertbedingung gegeben, kann diese eindeutig bestimmt werden.

Bemerkung 5.2.2 Aus 2. folgt: Sind $x_1(t), \dots, x_n(t)$ linear unabhängig für ein $t \in I$ mit $x_1, \dots, x_n \in \text{Ker}(L)$, dann sind alle Lösungen der homogenen DGL von der Form

$$\sum_{i=1}^n c_i x_i \quad \text{mit } c_i \in \mathbb{R} \quad \forall i \in \{1, \dots, n\}.$$

Bemerkung 5.2.3 (Variation der Konstanten) Dieses Verfahren ergibt sich direkt aus 3. des obigen Satzes, und besteht daraus, c_1, \dots, c_n zu finden, sodass $\sum_{i=1}^n \dot{c}_i x_i = g$ gilt.

Dieses Verfahren ist nicht immer erfolgreich, klappt jedoch immer, wenn $F(t)$ konstant ist (vgl. Kapitel 5.3).

Beweis. (Skizze, da nur Verbesserung des PICARD-LINDELÖF-Beweises) (1)

Für $z \in I$ gilt

$$\text{Anfangswertproblem} \iff x(t) = x_0 + \int_{t_0}^t F(\tau) x(\tau) + g(\tau) d\tau =: \Phi(x)(t). \quad (70)$$

Also ist das Anfangswertproblem äquivalent zu der Fixpunktgleichung $x = \Phi(x)$ mit $x \in C^0(I, V)$. Ist x stetig, so ist der Integrand in (70) stetig, und somit als Funktion von t differenzierbar (???).

Wir wählen auf $C^0(I, V)$ die Metrik, die von der Norm

$$\|f\|_\alpha := \sup_{t \in I} \|x(t)\| \cdot e^{-\alpha \cdot |t - t_0|} \quad \text{mit } \alpha \in \mathbb{R}.$$

induziert wird.

Fall 1: I ist kompakt.

Nun zeigt man (vgl. Ferus-Skript)

- (i) Der Vektorraum $C^0(I, V)$ ist für jedes α mit dieser Norm vollständig.
- (ii) Die Abbildung $\Phi : C^0(I, V) \rightarrow C^0(I, V)$ ist wohldefiniert.
- (iii) Die Abbildung Φ ist eine Kontraktion

Aus (3) erhält man, dass $\alpha > \sup_{t \in I} \|F(t)\|_{\text{op}}$ seien muss.

Nach dem Fixpunktsatz von BANACH existiert genau eine Lösung des Anfangswertproblems in I , und nicht auf einem kleineren Intervall.

Fall 2: I ist nicht kompakt.

Ist I nicht beschränkt, so wähle man kompakte Intervalle $I_1 \subset I_2 \subset \dots$ mit $I \supset \bigcup_{i=1}^{\infty} I_i$ und $t_0 \in I_k$ für alle $k \in \mathbb{N}$.

Dann kann auf jedem I_k das Anfangswertproblem nach Fall 1 eindeutig gelöst werden. Somit kann man aus diesen 'Teillösungen' die ganze Lösung $x(t)$ zusammenbasteln, indem man $x(t) = x_k(t)$ falls $t \in I_k$ definiert. Da alle Lösungen $x_k(t_0) = \xi_0$ gemeinsam haben und jeweils eindeutig sind, ist auch $x(t)$ eindeutig bestimmt.

- (2) Sei $t_1 \in I$ ein beliebiger 'Zeitpunkt'. Die Abbildung

$$\kappa_{t_1} : \text{Ker}(L) \rightarrow V, \quad x \mapsto x(t_1)$$

ist linear und bijektiv, wie eine Anwendung von i auf $t_0 = t$ und $g \equiv 0$ ergibt. Somit ist die Abbildung κ ein Vektorraumisomorphismus und die Dimension beider Vektorräume ist gleich.

Die anderen Aussagen folgen, da der Isomorphismus lineare Unabhängigkeit erhält.

- ③ Seien (x_1, \dots, x_n) eine Basis von $\text{Ker}(L)$ und $c_1, \dots, c_n \in \mathcal{C}^1(I, V)$. Setze nun $x := \sum_{i=1}^n c_i x_i$. Dann gilt nach der Produktregel und aufgrund der Linearität von F

$$\begin{aligned} (Lx)(t) &= \dot{x} - F(t)x(t) = \frac{d}{dt} \sum_{i=1}^n c_i(t)x_i(t) - F(t) \sum_{i=1}^n c_i(t)x_i(t) \\ &= \sum_{i=1}^n \dot{c}_i(t)x_i(t) + \sum_{i=1}^n c_i(t) \underbrace{[\dot{x}_i(t) - F(t)x_i(t)]}_{=(Lx_i)(t)=0} \\ &= \sum_{i=1}^n \dot{c}_i(t)x_i(t). \end{aligned}$$

Löst $x(t)$ die DGL, so gilt $(Lx)(t) = g(t)$, somit folgt die zweite Behauptung.

Für jedes t sind $x_1(t), \dots, x_n(t) \in V$ linear unabhängig nach ②. Eine Lösung x können wir als $\sum_{i=1}^n \dot{c}_i(t)x_i(t)$ mit geeigneten Koordinaten $\dot{c}_i(t)$ für $i \in \{1, \dots, n\}$ schreiben. Die Lösungen $\dot{c}_1(t), \dots, \dot{c}_n(t)$ sind rationale Funktionen der $x_1(t), \dots, x_n(t)$, 'erben' also deren Stetigkeit. Dann erhalten wir die c_i durch Integration der \dot{c}_i . \square

Beispiel 5.2.4 Seien $V = \mathbb{R}^2$ sowie

$$\begin{aligned} \dot{x}_1 &= x_1 + 3x_2 + 2 \cos^2(t) \\ \dot{x}_2 &= 3x_1 + x_2 + 2 \sin^2(t) \end{aligned}$$

Wir können dieses System umformen:

$$\dot{x} = \underbrace{\begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}}_{=:A} x(t) + \underbrace{\begin{pmatrix} 2 \cos^2(t) \\ 2 \sin^2(t) \end{pmatrix}}_{=:g(t)}.$$

Zuerst lösen wir die homogene DGL. Die Eigenwerte von A sind $\lambda_1 := 4$ und $\lambda_2 := -2$ zu den Eigenvektoren $v_1 := (1, 1)^T$ und $v_2 := (1, -1)^T$ und somit ist A diagonalisierbar.

Also sind $(x_i(t) := e^{\lambda_i \cdot t} \cdot v_i)_{i \in \{1, 2\}}$ eine Basis von $\text{Ker}(L)$.

Nun suchen wir mit der Methode 'Variation der Konstanten' eine partikuläre Lösung.

Ansatz: $g(t) = \sum_{i=1}^2 \dot{c}_i(t)x_i(t)$. Wir erhalten

$$2 \begin{pmatrix} \cos^2(t) \\ \sin^2(t) \end{pmatrix} = \begin{pmatrix} e^{4t} & e^{-2t} \\ e^{4t} & -e^{-2t} \end{pmatrix} \begin{pmatrix} \dot{c}_1(t) \\ \dot{c}_2(t) \end{pmatrix}$$

Durch Abziehen der Gleichungen voneinander und einsetzen erhält man

$$\begin{aligned} \dot{c}_1(t) &= e^{-4t} \implies c_1(t) = -\frac{1}{4}e^{-4t}, \\ \dot{c}_2(t) &= \cos(2t) \cdot e^{2t} \implies c_2(t) = \frac{1}{4}(\sin(2t) + \cos(2t)) \cdot e^{2t}. \end{aligned}$$

5.3 Lineare Differentialgleichungen mit konstanten Koeffizienten

15.10.2018



Differentialgleichung Finden einer Lösung $x : \mathbb{R} \rightarrow \mathbb{R}^n$ der linearen Differentialgleichung

$$\dot{x} = A \cdot x(t) \quad \text{mit } t \in \mathbb{R}, A \in \mathbb{R}^{n \times n}, \quad (71)$$

wobei \dot{x} die komponentenweise Ableitung und A die konstante Koeffizienten-Koeffizientenmatrix matrix ist.

Fall 1: Die Matrix A ist eine Diagonalmatrix.

Sei $A := \text{diag}(\lambda_1, \dots, \lambda_n)$. Dann erfüllt $x(t) = (x_1(t), \dots, x_n(t))$ die Gleichung $\dot{x}_i(t) = \lambda_i x_i(t)$ für alle $i \in \{1, \dots, n\}$. Somit folgt

$$x_i(t) = e^{\lambda_i \cdot t} x_i(x_0) \quad (72)$$

für alle $i \in \{1, \dots, n\}$, $t \in \mathbb{R}$ und mit einem Anfangswert $x_0 \in \mathbb{R}$.

Fall 2: Die Matrix A ist ähnlich zu einer Diagonalmatrix.

Dann existiert eine invertierbare Matrix $P \in \text{GL}(n, \mathbb{R})$, sodass $A = P^{-1}AP$ gilt.

Dann folgt

$$\dot{x} = Ax \iff (Px)' = P\dot{x} = APx$$

Für $y := Px$ gilt also $\dot{y} = Ay$ und wir erhalten die Lösungen aus Fall 1.

Fall 3: Das charakteristische Polynom $\chi_A(\lambda)$ zerfällt über \mathbb{R} in Linearfaktoren und besitzt die Eigenwerte $\lambda_1, \dots, \lambda_n$.

Jordan Dann existiert eine Jordanische Normalform $A = P^{-1}JP$, wobei J eine "Diagonale Normalmatrix" mit Jordanblöcken ist.

Beispiel 5.3.1 Die Matrix $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ ist in Jordannormalform mit einem Block der Länge eins und einem Block der Länge zwei und ist nicht äquivalent zu $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

Hauptvektoren Spalten von P bestehen aus den Hauptvektoren, d.h. Vektoren v mit $(A - \lambda I)^n v = 0$.

Nehmen wir an, dass $A \in \mathbb{R}^{n \times n}$ nur aus einem Jordanblock besteht, also den Eigenwert λ auf der Diagonale und Einsen auf der Nebendiagonale hat. Dann existiert eine zugehörige Basis (v_1, \dots, v_n) aus Hauptvektoren, sodass also $P = (v_1 \dots v_n)$: $PA = AP$, also für alle i : $Av_i = \lambda v_i + v_{i+1}$, d.h. für alle i gilt $(A - \lambda I)v_i = v_{i+1}$. Also gilt

$$v_i = (A - \lambda I)^{i-1} v_1.$$

Nun können wir die Lösung

$$x(t) = e^{\lambda t} \sum_{j=0^{n-1}} \frac{t^j}{j!} \underbrace{(A - \lambda I)^j v_1}_{=v_{j+1}}$$

für $\dot{x}(t) = Ax(t)$ verifizieren.

Es gilt

$$\begin{aligned}\dot{x}(t) &= \lambda x(t) + e^{\lambda t} \sum_{j=1} n \frac{t^{j-1}}{(j-1)!} v_{j+2} = \lambda x(t) + e^{\lambda t} \sum_{j=0} n - 1 \frac{t^j}{j!} (A - \lambda I) v_{j+1} \\ &\stackrel{L}{=} \lambda x(t) + (A - \lambda I) e^{\lambda t} \sum_{j=0} n - 1 \frac{t^j}{j!} v_{j+1} = \lambda x(t) + (A - \lambda I) x(t) = Ax.\end{aligned}$$

Ein Weg, diese Form von $x(t)$ herzuleiten, ist über die Exponentialreihe für Matrizen:

$$\exp : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}, A \mapsto \sum_{n=0}^{\infty} \frac{A^n}{n!}$$

Man setzt (analog zu $n - 1$) an: (???)

$$\begin{aligned}x(t) &= \exp(tA)v = \exp(t\lambda I + t(A - \lambda I))v = e^{\lambda t} \exp(t(A - \lambda I))v \\ &= e^{\lambda t} \sum_{j=0}^n \frac{t^j}{j!} (A - \lambda I)^j v\end{aligned}$$

Nun gilt $(A - \lambda I)^j v = 0$ wenn $j \geq n$ ist, wenn v ein Hauptvektor ist, also ist die Reihe endlich. Die Jordanische Normalform gibt uns also eine Basis aus Hauptvektoren (v_1, \dots, v_n) , dann ist (x_1, \dots, x_n) mit $x_i(t) = \exp(tA)v_i$ eine Basis des Lösungsraums von (71).

SATZ 5.3.1

Sei $A \in \mathbb{R}^{n \times n}$ mit Eigenwerten $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, und χ_A zerfalle über \mathbb{R} . Sei $\lambda \in \{\lambda_1, \dots, \lambda_n\}$ und sei $v \in \mathbb{R}$ ein Hauptvektor zur Stufe $k \in \mathbb{N} \setminus \{0\}$, d.h. $(A - \lambda I)^k v = 0$.

Dann ist $x(t) = e^{\lambda t} \sum_{j=0}^{n-1} \frac{t^j}{j!} (A - \lambda I)^j v$ eine Lösung der Differentialgleichung (71).

Bemerkung 5.3.2 Der Fall, dass A diagonalisierbar ist, ist in diesem Satz enthalten, da wir dann $k = 1$ wählen können und dann $x(t) = v$ übrig bleibt.

KOROLLAR 5.3.3

Sei $A \in \mathbb{R}^{n \times n}$ mit einer Basis (v_1, \dots, v_n) aus Hauptvektoren der Stufe $k_1, \dots, k_n \in \mathbb{N}_{>0}$ zu den Eigenwerten $\lambda_1, \dots, \lambda_n$, dann bilden die Funktionen

$$x_i(t) = e^{\lambda_i t} \sum_{j=0}^{n-1} \frac{t^j}{j!} (A - \lambda_i I)^j v_i$$

für $i \in \{1, \dots, n\}$ eine Basis des Lösungsraum von (71).

Beweis. Benutzt Satz 5.2.1, denn die $x_i(0)$ sind linear unabhängig. \square

Beispiel 5.3.4 Sei die Matrix

$$\begin{pmatrix} 0 & 1 & -1 \\ -2 & 3 & -1 \\ -1 & 1 & 1 \end{pmatrix}$$

Die Eigenwerte sind $\lambda_1 = \lambda_2 = 1$ und $\lambda_3 = 2$. Dann gilt $\text{Ker}(A - \lambda_1 I) = \text{span}(1, 1, 0)$. Also gilt $\dim(\text{Ker}(A - \lambda I)) = 1 < 2$. Wir brauchen also einen Hauptvektor.

Dazu suchen wir ein Vektor $v_2^{(1)}$ mit $(A - \lambda_2)^2 v_2^{(1)} = 0$, aber $(A - \lambda_2 I)v_2^{(1)} \neq 0$. Dazu betrachten wir $(A - \lambda_1)v_2 = v_1 \neq 0$ und erhalten $v_2 = (0, 0, -1)$.

Also hat (71) die drei linear unabhängigen Lösungen

$$x_1(t) := e^t(1, 1, 0)^T, \quad x_2(t) := e^t((0, 0, -1)^T + t(1, 1, 0)^T), \quad x_3(t) = e^{2t}(0, 1, 1)^T.$$

Übung am 15.10.18

DEFINITION 5.3.5 (LINEARE DIFFERENTIALGLEICHUNGEN MIT KONSTANTEN KOEFFIZIENTEN)

Wir betrachten den Fall \mathbb{R}^1 . Dann ist

$$\sum_{k=0}^n \alpha_k y^{(k)} = 0 \quad \text{mit } a_n \neq 0$$

eine homogene lineare Differentialgleichung n -ter Ordnung mit konstanten Koeffizienten $(a_i)_{i \in \{1, \dots, n\}} \in \mathbb{R}$.

Beispiel 5.3.6 (Hinführung zur Lösungstheorie mit der Schwingungsgleichung)

Wir betrachten die aus Analysis I bekannte Schwingungsgleichung

$$y'' + y = 0, \tag{73}$$

und können die reellwertige Lösungen $y_1(t) := \sin(t)$ und $y_2 := \cos(t)$ erkennen.

Wir wollen nun aus y_1 und y_2 die allgemeine Lösung von (73) konstruieren.

Behauptung: $y(t) = \alpha \cdot y_1(t) + \beta \cdot y_2(t)$ mit $\alpha, \beta \in \mathbb{R}$ ist die allgemeine Lösung von (73).

Es ist klar, dass die speziellen Lösungen y_1 und y_2 linear unabhängig sind und den gesamten Raum $\mathcal{C}_c(\mathbb{R}, \mathbb{R})$ aufspannen, und somit eine Fundamentalsystem bilden. Weil die Gleichung (73) eine lineare Differentialgleichung zweiter Ordnung ist, besteht das Fundamentalsystem aus zwei speziellen Lösungen.

Wir zeigen nun, dass $y(t)$ die Gleichung (73) löst.

$$y'' + y = A \sin''(t) + B \cos''(t) + A \sin(t) + B \cos(t) = A(-\sin(t) + \sin(t)) + B(-\cos(t) + \cos(t)) = 0,$$

da beide Terme verschwinden, weil y_1 und y_2 die speziellen Lösungen von (73) sind.

Warum klappt dieser unsystematischer Ansatz?

Lösungstheorie für lineare Differentialgleichungen mit konstanten Koeffizienten

💡 Ansatz

Weil die Ableitungen von y skalare Vielfachen von y sind, versuchen wir $y(t) = e^{\lambda \cdot t}$. Für den Fall $n = 1$ ist nämlich

$$\sum_{k=0}^n \alpha_k y^{(k)} = 0 \iff y' = -\frac{\alpha_0}{\alpha_1} \cdot y.$$

Wir setzen nun $y(t)$ in die allgemeine lineare Differentialgleichungen mit

konstanten Koeffizienten ein.

$$0 \stackrel{!}{=} \sum_{k=0}^n \alpha_k (e^{\lambda \cdot t})^{(k)} = \left(\underbrace{\sum_{k=0}^n \alpha_k \cdot \lambda^k}_{=: \chi(\lambda)} \right) \cdot e^{\lambda \cdot t}.$$

DEFINITION 5.3.7 (CHARAKTERISTISCHES POLYNOM EINER LINEAREN DIFFERENTIALGLEICHUNG)

Das Polynom in

$\chi(\lambda)$ nennt man **charakteristisches Polynom** einer linearen Differentialgleichung.

Da $e^{\lambda \cdot t} \neq 0$ für alle λ und t ist, können wir das Lösen der Differentialgleichung auf das Finden von Nullstellen λ_i von χ zurückführen.

Dies können wir nun auf die Schwingungsgleichung anwenden. Das charakteristische Polynom von (73) ist $\lambda^2 + 1 = 0 \iff \lambda_{1,2} = \pm i$. Also existieren zwei Lösungen $y_{1,2}(t) := e^{\pm i \cdot t}$. Aus den komplexen Lösungen $y_{1,2}(t)$ erhält man die reellwertigen Lösungen durch Streicher einer Lösung von jedem Paar komplex konjugierter Nullstellen:

$$\tilde{y}_1 := \Re(e^{i \cdot t}) = \cos(t) \quad \text{und} \quad \tilde{y}_2 := \Im(e^{i \cdot t}) = \sin(t)$$

e^{it} nennt man auch die **verbliebene Lösung**.

Beispiel 5.3.8 (Verdeutlichung an linearen DGLs zweiter Ordnung)

Die normierte allgemeine Form einer linearen Differentialgleichung zweiter Ordnung mit konstanten Koeffizienten ist

$$y'' + ay' + b = 0.$$

Dann ist das charakteristische Polynom gegeben durch $\chi(\lambda) = \lambda^2 + a\lambda + b$. Fall 1: $\chi(\lambda)$ hat zwei komplexe Nullstellen $\lambda_1, \bar{\lambda}_1 \in \mathcal{C}$. Das Lösungsverfahren haben wir gerade am Fall der Schwingungsgleichung gezeigt.

Fall 2: $\chi(\lambda)$ hat zwei reelle Nullstellen $\lambda_1, \lambda_2 \in \mathbb{R}$ mit $\lambda_1 \neq \lambda_2$. Dann ist $(y_{1,2} = e^{\lambda_{1,2} \cdot t})$ das linear unabhängige Fundamentalsystem und die allgemeine Lösung $y = \mu y_1 + \rho y_2$.

Fall 3: Ist $\lambda_1 \in \mathbb{R}$ die einzige Nullstelle von χ , so ist

$$y_1(t) = e^{\lambda_1 \cdot t} \quad \text{und} \quad y_2(t) = t \cdot e^{\lambda_1 \cdot t}$$

das linear unabhängige Fundamentalsystem.

Wenn wir y_2 in die Differentialgleichung einsetzen, erhalten wir

$$\begin{aligned} y_2'' + ay'_2 + by_2 &= e^{\lambda_1 \cdot t} [t\lambda_1^2 + 2\lambda_1] + a \cdot e^{\lambda_1 \cdot t} [\lambda_1 t + 1] + b \cdot e^{\lambda_1 \cdot t} [t] \\ &= e^{\lambda_1 \cdot t} [t\lambda_1^2 + 2\lambda_1 + a(\lambda_1 \cdot t + 1) + b(t)] \\ &= t \cdot e^{\lambda_1 \cdot t} \underbrace{[\lambda_1^2 + \lambda_1 a + b]}_{\chi(\lambda)} + e^{\lambda_1 \cdot t} [2\lambda_1 + a] \end{aligned}$$

Weil $\chi(\lambda)$ nach Voraussetzung gleich Null ist, und dann auch ihre erste Ableitung $\chi'(\lambda) = 0$ ist, ist also auch y_2 eine Lösung der Differentialgleichung.

5.3 LINEARE DIFFERENTIALGLEICHUNGEN MIT KONSTANTEN KOEFFIZIENTEN

2. Newton'sches Gesetz

Die Änderung der Bewegung (Beschleunigung a) eines Körpers mit Masse m ist der Einwirkung der bewegenden Kraft F proportional.

$$F = ma \quad (1)$$

Problem: Beschreibe die Auslenkung $y(t)$ des Körpers von einem Referenzpunkt.

Umschreiben von (1) in eine Differentialgleichung

a ist die Änderungsrate der Geschwindigkeit v , d.h. $a = \frac{d}{dt}v = v'$

v ist die Änderungsrate von $y(t)$, d.h. $v = \frac{dy}{dt} = y'$

Die Kraft F hängt von Zeit t , Geschwindigkeit und Auslenkung ab, d.h. $F = F(t, y, v) = F(t, y, y')$. Zusammen liefert dies

$$F(t, y, y') = my''$$

Beispiel: vibrierende Feder

Betrachte eine Feder an einem Balken:



0

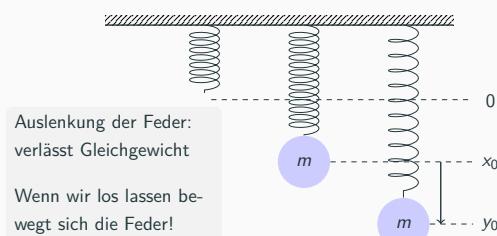
Befestige Gewicht m :
System im Gleichgewicht in Höhe x_0

Kräfte am Gewicht: **Gravitation** mg und **Rückstellkraft** $R(y)$ abhängig von der Auslenkung y .

Im Gleichgewicht bewegt sich die Feder nicht.

Beispiel: vibrierende Feder

Betrachte eine Feder an einem Balken:



Kräftewirkung in Bewegung: **Dämpfung** $D(v)$ abhängig von Geschwindigkeit v , **äußere Kräfte** $F(t)$, $R(y)$, mg .

Ein Modell für die Auslenkung x der Feder

Da $v = y'$ und $a = v' = y''$.

gilt nach dem 2. Newton'schen Gesetz (1)

$$\begin{aligned} ma &= \text{Summe aller Kräftewirkung auf das Gewicht} \\ &= R(y) + mg + D(v) + F(t) \end{aligned} \quad (2)$$

Umschreiben von (2) liefert

$$my'' = R(y) + mg + D(y') + F(t). \quad (3)$$

Für manche Federn gilt das **Hooke'sche Gesetz**: $R(y) = -\omega^2 y$ mit ω konstant für kleine y .

Nehme an, dass das Hooke'sche Gesetz gilt, dann wird aus (2)

$$my'' = -\omega^2 y + mg + D(y') + F(t). \quad (4)$$

harmonischer Oszillator

Auf dem Übungsblatt betrachten wir den **gedämpften harmonischen Oszillator**

$$y'' + 2\kappa y' + \omega^2 y = f(t) \quad (5)$$

mit $\kappa, \omega \in \mathbb{R}$ und $f(t)$ gegeben.

Die Parameter nennt man (inspiriert aus der Physik):

1. κ , Dämpfungsconstante, ω , natürliche Frequenz, $f(t)$, äußere Kraft

Beispiel: Die vibrierende Feder

$$my'' = -\omega y + mg + D(y') + F(t)$$

wobei nun $D(y') = -2\kappa y'$ für $\kappa \in [0, \infty)$ und somit (für $m = 1$)

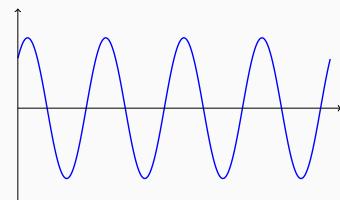
$$y'' + 2\kappa y' + \omega^2 y = F(t) + g \quad (6)$$

Einfacher harmonischer Oszillator

Ist $\kappa = 0 = F(t) + g$ so ist eine Lösung von (6):

$$y(t) = \cos(\omega t) + \sin(\omega t)$$

Die Lösung ist periodisch mit Periode ω .



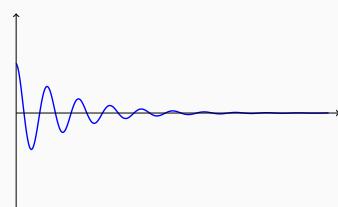
Keine Dämpfung: Lösung oszilliert mit natürlicher Frequenz (hier $\omega = 4$)

Gedämpfter harmonischer Oszillator

Nun mit Dämpfung $\kappa > 0$ (aber $F(t) + g = 0$), so dass $\kappa^2 < \omega^2$.

Man¹ findet als Lösung von (6) in diesem Fall:

$$y(t) = e^{-\kappa t} \cos(\sqrt{\omega^2 - \kappa^2} t - \varphi_0)$$



geringe Dämpfung: Lösung oszilliert, fällt aber exponentiell gegen 0 ab (hier gezeichnet $\kappa = 1, \omega = 10$)

¹d.h. Ihr wenn Ihr den Übungszettel macht.

Beispiel 5.3.9 (Die Airy'sche Differentialgleichung) Die Gleichung $y'' - xy = f(x)$ ist eine lineare Differentialgleichung mit nicht-konstanten Koeffizienten, da die Koeffizienten von x , aber nicht von $y(x)$ oder ihren Ableitungen abhängen dürfen.

5.4 Skalare lineare Differentialgleichungen höherer Ordnung

Umschreiben in eine System erster Ordnung

Wir wollen Lösungen der Differentialgleichung

$$x^{(n)}(t) + \sum_{k=0}^{n-1} f_k(t)x^{(k)}(t) = g(t), \quad (74)$$

finden, wobei $f_0, \dots, f_{n-1}, g : J \rightarrow \mathbb{R}$ stetige Funktionen sind. Wollen wir eine eindeutige Lösung, benötigen wir n Anfangsbedingungen

$$x(t_0) = \xi_0, \dots, x^{(n-1)}(t_0) = \xi_{n-1}.$$

Wir können dieses Anfangswertproblem in ein System von n Differentialgleichungen erster Ordnung überführen. Dafür setzen wir $y_1 := x, \dots, y_n := x^{(n-1)}$ und $y := (y_1, \dots, y_n)^T$ und betrachten

$$\dot{y} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \vdots & 0 & \ddots & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \\ -f_0 & -f_1 & \dots & \dots & f_{n-1} \end{pmatrix} y + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ g(t) \end{pmatrix}. \quad (75)$$

mit

$$y(t_0) = (y_1(t_0) \dots y_n(t_0))^T = (\xi_0 \dots \xi_{n-1})^T.$$

Somit löst y genau dann (75), wenn x das Anfangswertproblem (74) löst. Ist umgekehrt $y : J \rightarrow \mathbb{R}^n$ eine Lösung von (75), so löst $x := y_1$ das Anfangswertproblem (74).

16.10.2018

Wir wollen nun reelle Lösungen der Differentialgleichung (71) finden, wenn $A \in \mathbb{R}^{n \times n}$ auch komplexe Eigenwerte hat. Hier können wir immer noch unsere Theorie für reelle Eigenwerte benutzen, wollen aber aus den nun komplexen Lösungen reelle konstruieren.

Zunächst gilt

$$\bar{\dot{x}} = \dot{\bar{x}} = \overline{Ax} = \bar{A}\bar{x} = A\bar{x}$$

Somit löst $x : J \rightarrow \mathcal{C}^n$ genau dann (71), wenn \bar{x} die DGL löst. Somit können wir die konjugiert komplex auftretenden Eigenwerten von A einen vernachlässigen.

Weil es sich bei (71) um eine lineare Differentialgleichung handelt, sind auch Linearkombinationen von Lösungen wieder Lösungen. Somit sind also auch

$$\Re(x)(t) = \frac{1}{2} (x(t) + \bar{x}(t)) \in \mathbb{R} \quad \text{und} \quad \Im(x)(t) = \frac{1}{2i} (x(t) - \bar{x}(t)) \in \mathbb{R}.$$

Lösungen von (71).

Aus der linearen Algebra wissen wir, dass Eigenvektoren zu konjugierten Eigenwerten konjugiert zueinander sind.

SATZ 5.4.1: ALGORITHMUS FÜR DAS FINDEN EINER REELLEN LÖSUNGSBASIS VON (71)

- ① Berechne die Eigenwerte von A und lasse von Paaren konjugiert komplexer Eigenwerte jeweils einen weg.
- ② Zu jedem verbleibenden Eigenwert λ mit algebraischer Vielfachheit k berechne k Hauptvektoren $v_1, \dots, v_k \in \mathbb{C}^n$. Dann erhalten wir k Lösungen der Form

$$x_{\lambda, \kappa}(t) = e^{\lambda t} \sum_{j=0}^k \frac{t^j}{j!} (A - \lambda E)^j v_\kappa \quad \text{mit } \kappa \in \{1, \dots, k\}.$$

- ③ Die entstehenden nicht reellen Lösungen zerlege in Real- und Imaginärteil. Dies liefert insgesamt n linear unabhängige reelle Lösungen.

[Bsp 202]

Wir betrachten auf $J = \mathbb{R}$ die homogene lineare Differentialgleichung mit konstanten Koeffizienten 22.10.18

$$x^{(n)} + a_1 x^{(n-1)} + \dots + a_{n-1} x^{(1)} + a_n x = 0 \quad (76)$$

mit konstanten a_0, \dots, a_{n-1} und der gesuchten Lösung $x : \mathbb{R} \rightarrow \mathbb{R}$.

Dann ist $\chi(\lambda) = \lambda^n + a_1 \lambda^{n-1} + \dots + a_n$ das **charakteristische Polynom** von (76). Seien $\lambda_1, \dots, \lambda_m \in \mathbb{C}$ die verschiedenen Nullstellen von $x(\lambda_i) = 0$ und $k_i \in \mathbb{N}_{>0}$ die Vielfachheit von λ_i in $\chi(\lambda)$.

SATZ 5.4.2: REZEPT FÜR DAS LÖSEN SKALARER DGLs

- ① Die Funktionen $x_{i,j} := t^j e^{\lambda_i t}$ mit $i \in \{1, \dots, n\}$ und $j \in \{1, \dots, k_i - 1\}$ bilden eine (evtl. komplexe) Lösungsbasis von (76) in der Menge aller Funktion von \mathbb{R} nach \mathbb{C} .
- ② Wenn für $\lambda_i \in \mathbb{C} \setminus \mathbb{R}$ der Term $t^j e^{\lambda_i t}$ bzw. $t^j e^{\bar{\lambda}_i t}$ durch $t^j e^{t \cdot \Re(\lambda_i)} \cdot \cos(t \cdot \Im(\lambda_i))$ und $t^j e^{t \cdot \Re(\lambda_i)} \cdot \sin(t \cdot \Im(\lambda_i))$ ersetzt wird, entsteht eine Basis aus reellwertigen Lösung von (76) von \mathbb{R} nach \mathbb{R} .

Beweis. ① Wir zeigen zunächst, dass jedes $x_{i,j}$ eine Lösung ist.

Es gilt

$$x_{i,j}^{(n)} + a_1 x_{i,j}^{(n-1)} + \dots + a_{n-1} x_{i,j}^{(1)} + a_n x_{i,j} = \chi(D) x_{i,j}$$

Weil $\chi(\lambda)$ in über \mathbb{C} in Linearfaktoren zerfällt, gilt

$$\chi(D) x_{i,j} = \left(\prod_{\ell=1}^m (D - \lambda_\ell)^{k_\ell} \right) x_{i,j}$$

Nach Lemma 206 gleich das

$$(D - \lambda_1)^{k_1} \dots (D - \lambda_{i-1})^{k_{i-1}} (D - \lambda_{i+1})^{k_{i+1}} \dots (D - \lambda_m)^{k_m} \underbrace{(D - \lambda_i)^{k_i} (t^j e^{\lambda_i t})}_{=0} = 0.$$

Nun zeigen wir, dass die $x_{i,j}$ linear unabhängig sind. Seien dazu $\alpha_{i,j} \in \mathbb{C}$ mit

$$0 = \sum_{i=1}^m \sum_{j=0}^{k_i-1} \alpha_{i,j} x_{i,j}(t) = \sum_{i=1}^m e^{\lambda_i t} \cdot p_i(t),$$

wobei $p_i(t)$ ein Polynom von Grad $\leq k_i - 1$ ist.

Dann gilt aber auch mit Lemma 206

$$0 = (D - \lambda_2)^{k_2} \dots (D - \lambda_m)^{k_m} \left(\sum_{i=1}^m e^{\lambda_i t} p_i(t) \right) = q_1(t) e^{\lambda_1 t}$$

für eine Polynom $q_1(t)$ mit $\deg(q_1(t)) = \deg(p_1)$. Daraus folgt $p_1 \equiv 0$. Analog folgt $p_i \equiv 0 \forall i \in \{2, \dots, m\}$.

Da nun gilt $\sum_{j=0}^{k_i-1} a_{i,j} t^j = 0$ für alle i gilt, folgt $\alpha_{i,j} = 0$ für alle $j \in \{1, \dots, k_i - 1\}$.

(2) Ist wie im Beweis des Satzes im Kasten auf Seite 128. \square

Übung 22.11.18

DEFINITION 5.4.1 (ALGEBRA)

Sei $(V, +, \cdot)$ ein Vektorraum und $m : V \times V \rightarrow V$ bilinear und verträglich mit $+$ und \cdot . Dann ist (V, m) eine Algebra.

Lemma 5.4.2 (206 im Ferus)

Sei $D = \frac{d}{dt}$ der Differentialoperator und somit $D^k = \frac{d^k}{dt^k}$ für ein $k \in \mathbb{N}$. Dann ist die "Einsetzung" $E(\varphi) = \varphi(D)$ definiert durch

$$E : \mathbb{R}[x] \rightarrow L(\mathcal{C}_c^\infty(\mathbb{R}, \mathbb{R}), \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})), \quad \varphi(x) = \sum_{i=1}^n a_i x_i \mapsto \sum_{i=1}^n a_i D^i$$

ein Algebren-Homomorphismus.

Beweis. Wir zeigen, dass der Operator E linear (klar.) ist, und dass für alle Polynom $\varphi, \psi \in \mathbb{R}[x]$ gilt $E(\varphi \cdot \psi) = E(\varphi) \circ E(\psi)$.

Seien $\varphi(x) = \sum_{i=0}^n a_i x^i$ und $\psi(x) = \sum_{j=0}^m b_j x^j$ beliebige Polynom in $\mathbb{R}[x]$. Dann gilt

$$\begin{aligned} E(\varphi \cdot \psi)(x) &= E(\varphi(x) \cdot \psi(x)) = \left(\sum_{i=0}^n a_i x^i \right) \left(\sum_{j=0}^m b_j x^j \right) = \sum_{k=0}^{m+n} \left(\sum_{i+j=k} a_i b_j \right) x^k \\ &= \sum_{k=0}^{m+n} \sum_{i+j=k} a_i b_j D^k = \sum_{k=0}^{m+n} \sum_{i+j=k} a_i b_j D^i \circ D^j = \left(\sum_{i=0}^n a_i D^i \right) \circ \left(\sum_{j=0}^m b_j D^j \right) \\ &= E(\varphi) \circ E(\psi). \end{aligned} \quad \square$$

Lemma 5.4.3 (206b im Ferus)

Seien $p \neq 0$ ein Polynom, $k \in \mathbb{N}$ und $\mu, \kappa \in \mathbb{C}$ komplexe Zahlen.

Dann ist

$$(D - \mu)^k (p(t) e^{\kappa t}) = q(t) e^{\kappa t}$$

mit einem Polynom q , für welches gilt

$$\text{grad}(q) = \begin{cases} \text{grad}(p), & \text{wenn } \mu \neq \kappa, \\ \text{grad}(p) - k, & \text{wenn } \mu = \kappa. \end{cases}$$

Dabei gilt $\text{grad}(q) < 0 \iff q \equiv 0$.

Beweis. Wir beweisen das Lemma mit vollständiger Induktion über $k \in \mathbb{N}$.

Induktionsanfang: ($n = 0$). Dann gilt $(D - \mu)^0 = \text{id}$. Mit $p = q \implies \text{grad}(p) = \text{grad}(q) - 0$ ist die Behauptung richtig.

Induktionsschritt: $n \rightarrow n + 1$. Mit Hilfe der Produktregel (*) gilt

$$\begin{aligned} (D - \mu)^{n+1} (p(t)e^{\kappa t}) &= (D - \mu) \circ (D - \mu)^n (p(t)e^{\kappa t}) \\ &\stackrel{\text{IV}}{=} (D - \mu)(q(t)e^{\kappa t}) \\ &\stackrel{(*)}{=} \tilde{q}(t)e^{\kappa t} + \underbrace{\kappa q(t)e^{\kappa t} - \mu q(t)e^{\kappa t}}_{=: \tilde{p}(t)}. \end{aligned}$$

Dann gilt $\text{grad}(\tilde{q}) = \text{grad}(q) - 1$ und $\text{grad}(\tilde{p}) = \begin{cases} 0 & \kappa = \mu \\ \text{grad}(q) & \text{sonst.} \end{cases}$.

Somit gilt

$$(D - \mu)^{n+1} (p(t)e^{\kappa t}) = (\tilde{q}(t) + \kappa q(t) - \mu q(t)) e^{\kappa t}$$

und dieses Polynom hat den gewünschten Grad. \square

Interessante Fakten aus Kahoot

Lemma 5.4.4

Jede Lösung x von $x' = Ax$ mit $A \in \mathbb{R}^{n \times n}$ ist beliebig oft stetig differenzierbar.

Beweis. Sei $x \in \mathcal{C}_c^1$. Da aber A linear und somit beliebig oft differenzierbar ist, ist $x' \in \mathcal{C}_c^2$. Dieses Argument kann iterativ anwenden und erhält $x \in \mathcal{C}_c^\infty$. \square

Lemma 5.4.5

Alle Lösungen von $x' = Ax$ lassen sich immer zu Lösungen auf ganz \mathbb{R} fortsetzen.

Beweis. Selbst. \square

Inhomogene lineare DGLs zweiter Ordnung (mit nicht konstanten Koeffizienten)

Ansatz

Sei die Differentialgleichung

$$y''(t)p(t) + q(t)y(t) = g(t).$$

gegeben. Wir benötigen zwei lineare unabhängige Lösungen u_1 und u_2 der homogenen Gleichung. Dann ist die allgemeine Lösung der homogenen Gleichung

$$y(t) = c_1(t)u_1 + c_2(t)u_2,$$

wobei c_1 und c_2 bis auf Konstanten durch folgenden Ansatz bestimmt sind:

$$W(t) \begin{pmatrix} \dot{c}_1(t) \\ \dot{c}_2(t) \end{pmatrix} = \begin{pmatrix} 0 & g(t) \end{pmatrix}.$$

Da die WRONSKI-Matrix nach Vorlesung invertierbar ist, erhält man

$$\begin{pmatrix} \dot{c}_1(t) \\ \dot{c}_2(t) \end{pmatrix} = \frac{1}{\det(W)} \begin{pmatrix} \dot{u}_2(t) & -u_2(t) \\ -\dot{u}_1(t) & u_1(t) \end{pmatrix} \begin{pmatrix} 0 & g(t) \end{pmatrix} = \frac{1}{\det(W)} \begin{pmatrix} -u_2(t)g(t) \\ u_1(t)g(t) \end{pmatrix}$$

Durch Integration erhält man also

$$\boxed{c_1(t) = -\frac{1}{\det(W)} \int_{t_0}^t u_2(s)g(s)ds \quad \text{und} \quad c_2(t) = \frac{1}{\det(W)} \int_{t_0}^t u_1(s)g(s)ds.} \quad (77)$$

Beispiel 5.4.6 Wir suchen die Lösung der Gleichung $y''(t) + y(t) = \tan(t)$. Die homogenen Lösungen sind durch $y_1(t) = \sin(t)$ und $y_2(t) = \cos(t)$ gegeben und es gilt $W_D(y_1, y_2) = -1$. Nach (77) gilt

$$c_1(t) = -\cos(t) \quad \text{und} \quad c_2 = -\sin(t) - \frac{1}{2} \ln \left| \frac{\sin(t) + 1}{\sin(t) - 1} \right|$$

Also ist die allgemeine Lösung gegeben durch

$$y(t) = c_1(t)u_1 + c_2(t)u_2 = -\cos(t)\sin(t) + \underbrace{\left(-\sin(t) - \frac{1}{2} \ln \left| \frac{\sin(t) + 1}{\sin(t) - 1} \right| \right)}_{\text{partikuläre Lösung}} + A\sin(t) + B\cos(t)$$

Methode der unbestimmten Koeffizienten

Problem

Sei die Differentialgleichung $\sum_{i=0}^n a_i y^{(i)} = g(t) \neq 0$ mit konstanten Koeffizienten a_i und einem 'schönen' g , z.B. ein Polynom, eine Exponentialfunktion oder $A\sin(t) + B\cos(t)$ gegeben.

$g(t)$	Ansatz
$\sum_{k=0}^n a_k x^k$	$\sum_{k=0}^n a_k x^k$
$\sin(t)$ o. $\cos(t)$	$A\sin(t) + B\cos(t)$
$e^{\lambda t}$	$e^{\lambda t}$.

Beispiel 5.4.7 Wir wollen die DGL $y'' - y' - 2y = 2e^{-2t}$ lösen. Weil die Exponentialfunktion sich beim Ableiten reproduziert, wählen wir als Ansatz $y(t) = Ce^{-2t}$, wobei C konstant und unbestimmt ist.

Einsetzen liefert $C = \frac{1}{2}$ und $y(t) = \frac{1}{2}e^{-2t}$ ist eine Lösung der inhomogenen Gleichung.

Die Matrixexponentialfunktion

Lemma 5.4.8

Es gilt $\exp(tA)' = A \cdot \exp(tA)$ für eine quadratische Matrix $A \in \mathbb{R}^{n \times n}$ und ein beliebiges $t \in \mathbb{R}$.

Beweis. Gilt $\exp(tA) = A \cdot \exp(tA)$, so löst $\exp(tA)$ die Differentialgleichung $\dot{x} = Ax$. Ist zusätzlich ein Anfangswertproblem mit $x(0) = x_0$ gegeben, so folgt aus $\exp(0 \cdot A) = I$, dass $x(t) = \exp(tA) \cdot x_0$ das Anfangswertproblem löst.

Seien $x(t) := \exp(tA)A$ und $y(t) := A\exp(tA)$. Dann gilt $x(0) = y(0) = A$ sowie

$$\dot{x}(t) = A\exp(tA) \cdot A = Ax(t) \quad \text{und} \quad \dot{y}(t) = A\exp(tA) \cdot A = Ay(t).$$

Somit lösen die Funktionen $x(t)$ und $y(t)$ die Differentialgleichung $\dot{u} = Au$ mit der selben Anfangswertbedingung. Da ein Anfangswertproblem eindeutig lösbar ist gilt $x(t) = y(t)$. \square

Korollar 5.4.9 (gilt das wirklich??)

Die Matrizen A und e^{tA} kommutieren für jedes $t \in \mathbb{R}$.

Lemma 5.4.10

Für quadratische Matrizen $A, B \in \mathbb{R}^{n \times n}$ gilt $\exp(A)\exp(B) = \exp(A+B)$, wenn A und B kommutieren.

Beweis. (von mir) Seien A und B quadratische kommutierenden Matrizen.

Wir setzen $X(t) := e^{tA}e^{tB}$. Diese Funktion ist die Lösung des Anfangswertproblems

$$\begin{cases} X'(t) = (A+B)X(t) & \forall t > 0 \\ X(0) = E \end{cases},$$

denn es gilt aufgrund des letzten Lemmas und der Kommutativität der Matrizen

$$X'(t) = Ae^{tA}e^{tB} + e^{tA}Be^{tB} = (A+B)e^{tA}e^{tB} = (A+B)X(t).$$

Ferner gilt $X(0) = e^0e^0 = E$. Aufgrund der Eindeutigen Lösbarkeit dieser Art von Anfangswertproblems ist $X(t) = e^{tA}e^{tB}$ die einzige Lösung und es gilt

$$e^{tA}e^{tB} = e^{t(A+B)},$$

weil Anfangswertproblems der Form $X'(t) = AX(t)$ mit $X(0) = Z$ immer von Ze^{tA} gelöst werden. Mit $t = 1$ folgt die Behauptung. \square

TEIL II

Analysis III: Maß- und Integrationstheorie

6 Das RIEMANN-Integral für Funktionen

$$\mathbb{R}^d \rightarrow \mathbb{R}$$

Erinnerung zum Integrationsbegriff von RIEMANN

Eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ heißt genau dann **RIEMANN-integrierbar**, wenn für alle $\varepsilon > 0$ eine Zerlegung von $[a, b]$ und dazu zwei Treppenfunktionen φ und ψ mit $\varphi \leq f \leq \psi$ existieren und

$$\sum_{i=1}^n (t_i - t_{i-1})(\varphi(t_i) - \psi(t_i)) = \int_a^b |\psi(x) - \varphi(x)| dx < \varepsilon$$

gilt, falls φ und ψ auf $(t_{i-1}, t_i]$ konstant sind ^a und $a = t_0 < t_1 < \dots < t_n = b$ die Zerlegung ist.

Das ist äquivalent zu

$$\exists I \in \mathbb{R} \quad \forall \varepsilon > 0 \quad \exists \mathcal{Z} : a = t_0 < t_1 < \dots < t_n = b$$

$$\forall \xi_i \in (t_{i-1}, t_i) : \left| I - \underbrace{\sum_{i=1}^n f(\xi_i)(t_i - t_{i-1})}_{=: \text{RIEMANN-Summe mit Unterteilung } \mathcal{Z \text{ und Stützstellen } \xi_i}} \right| < \varepsilon.$$

RIEMANN-Summe

=: RIEMANN-Summe mit

Unterteilung \mathcal{Z}

und Stützstellen ξ_i

Dann ist $I = \int_a^b f(x) dx$ das **RIEMANN-Integral** von f über $[a, b]$.

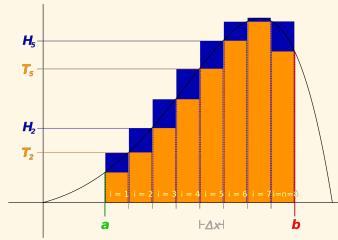


Abb. 28: RIEMANN-Integral mit Unter- und Obersummen.

Aus Analysis I wissen wir, dass stetige Funktionen RIEMANN-integrierbar sind.

^aalternativ kann man auch nur fordern, dass die Treppenfunktionen auf (t_{i-1}, t_i) konstant sind und dann die Mittelpunkte dieser Intervalle verwenden.

6.1 Das HAARSche Maß

Nun betrachten wir für $a_i, b_i \in \mathbb{R}$ für alle $i \in \{1, \dots, d\}$ und $d \in \mathbb{N}$ den kompakten achsenparallelen Quader $Q := \times_{i=1}^d [a_i, b_i]$ und eine stetige Funktion $f : Q \rightarrow \mathbb{R}$

Wir wollen $\int_Q f(x)dx$ definieren.

Dazu überlegen wir uns (HA 7b), dass (hier für $d = 2$) die Abbildung

$$x_2 \mapsto \int_{a_1}^{b_1} f(x_1, x_2) dx_1$$

stetig ist, sodass wir

$$\boxed{\int_Q f(x)dx := \int_{a_d}^{b_d} \left(\dots \int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} f(x_1, \dots, x_d) dx_1 \right) dx_2 \right) \dots dx_d} \quad (78)$$

definieren dürfen.

SATZ 6.1.1: VON FUBINI FÜR STETIGE FUNKTIONEN

Für jede Permutation i_1, \dots, i_d von $1, \dots, d$ gilt

$$\int_Q f(x)dx = \int_{a_{i_d}}^{b_{i_d}} \left(\dots \left(\int_{a_{i_2}}^{b_{i_2}} \left(\int_{a_{i_1}}^{b_{i_1}} f(x_1, \dots, x_d) dx_{i_1} \right) dx_{i_2} \right) \dots \right) dx_{i_d} \quad (79)$$

Beweis. In der großen Übung. □

Nun wollen wir diesen Integralbegriff für stetige Funktionen auf kompakten Quadern auf stetige Funktionen von ganz $\mathbb{R}^d \rightarrow \mathbb{R}$ erweitern.

DEFINITION 6.1.1 (TRÄGER EINER FUNKTION)

Der **Träger** einer Funktion $f : \mathbb{R}^d \rightarrow \mathbb{R}$ ist der Abschluss der Nichtnullstellenmenge von f :

$$\text{supp}(f) := \overline{\{x \in \mathbb{R}^d : f(x) \neq 0\}}.$$

DEFINITION 6.1.2 (STETIGEN FUNKTION MIT KOMPAKTEM TRÄGER)

Die Menge der stetigen Funktionen mit **kompaktem Träger** bezeichnen wir mit

$$\mathcal{C}_c(\mathbb{R}^d) := \{f : \mathbb{R}^d \rightarrow \mathbb{R} \text{ stetig, } \text{supp}(f) \text{ kompakt}\}.$$

Für $f \in \mathcal{C}_c(\mathbb{R}^d)$ können wir also definieren

$$\int_{\mathbb{R}^d} f(x)dx := \int_Q f(x)dx,$$

wobei der achsenparallele Quader Q den Träger von f enthält.

Lemma 6.1.3

Diese Definition hängt nicht von der Wahl des Quaders ab.

Beweis. (unvollständig) Wenn für einen zweiten Quader \tilde{Q} auch der achsenparallele Quader $Q \cap \tilde{Q}$ ausreicht, und somit alles in $Q \setminus (Q \cap \tilde{Q})$ bzw. $\tilde{Q} \setminus (Q \cap \tilde{Q})$ also nur die konstante Nullfunktion integriert und somit nicht zum Wert des Integral beiträgt. \square

Schreibweise: $\int_{\mathbb{R}^d} f(x) dx = \int_{\mathbb{R}^d} f(x_1, \dots, x_d) dx = \int f(x) d^d x = \int f(\underline{x}) dx$.

DEFINITION 6.1.4 (TRANSLATIONSINVARIANZ)

Funktional Ein Funktional¹⁰ $J : \mathcal{C}_c(\mathbb{R}^d) \rightarrow \mathbb{R}$ heißt translationsinvariant, falls für alle $x \in \mathbb{R}^d$ gilt

$$J(\theta_x f) = J(f),$$

Shift-Operator wobei für ein $y \in \mathbb{R}^d$ der Shift-Operator als

$$\theta_x(y) = y - x \quad \text{oder} \quad \theta_x f(y) = f(y - x)$$

definiert ist.

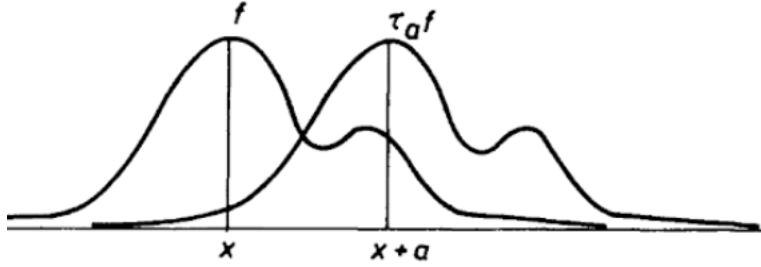


Abb. 29: Wirkung des Shift-Operators τ_a auf eine Funktion f . [Quelle: Forster]

Lemma 6.1.5 (Produktfunktionen)

Es seien $\varphi_1, \dots, \varphi_d \in \mathcal{C}_c(\mathbb{R})$ gegeben. Dann liegt die Funktion

$$f := \bigotimes_{i=1}^d \varphi_i : \mathbb{R}^d \rightarrow \mathbb{R}, \quad (x_1, \dots, x_d) \mapsto \prod_{i=1}^d \varphi_i(x_i)$$

in $\mathcal{C}_c(\mathbb{R}^d)$ und es gilt $\int_{\mathbb{R}^d} f(x) dx = \prod_{i=1}^d \int_{\mathbb{R}} \varphi_i(x_i) dx_i$.

Beweis. Es genügt zu zeigen, dass $\text{supp}(f)$ kompakt ist, da f als Komposition stetiger Funktionen stetig ist. Nun gilt

$$\begin{aligned} \text{supp}(f) &= \overline{\left\{ x = (x_1, \dots, x_d)^T \in \mathbb{R}^d : \prod_{i=1}^d \varphi_i(x_i) \neq 0 \right\}} \\ &= \overline{\left\{ x = (x_1, \dots, x_d)^T \in \mathbb{R}^d : \varphi_i(x_i) \neq 0 \forall i \in \{1, \dots, d\} \right\}} \\ &= \overline{\prod_{i=1}^d \{x \in \mathbb{R} : \varphi_i(x) \neq 0\}} = \prod_{i=1}^d \text{supp}(\varphi_i). \end{aligned}$$

Produkte kompakter Mengen sind kompakt, somit folgt die Behauptung.

Sei nun $S_i := \text{supp}(\varphi_i)$. Dann gilt

$$\begin{aligned}\int_{\mathbb{R}^d} f(x) dx &= \int_{\text{supp}(f)} \prod_{i=1}^d \varphi_i(x_i) dx = \int_{S_d} \left(\dots \left(\int_{S_1} \prod_{i=1}^d \varphi_i(x_i) dx_1 \right) \dots \right) dx_d \\ &= \left(\int_{S_1} \varphi_1(x_1) dx_1 \right) \left(\int_{S_d} \left(\dots \left(\int_{S_2} \prod_{i=2}^d \varphi_i(x_i) dx_2 \right) \dots \right) dx_d \right)\end{aligned}$$

Durch weitere $d - 1$ -fache Anwendung des letzten Schritts folgt die Behauptung. \square

SATZ 6.1.2: AXIOMATISCHE INTEGRAL-CHARAKTERISIERUNG

Das Funktional $I : \mathcal{C}_c(\mathbb{R}^d) \rightarrow \mathbb{R}$, $f \mapsto \int f(x) dx$ ist linear, monoton und translationsinvariant.

translationsinvariant

Beweis. Die Linearität und Monotonie folgt unmittelbar aus den vergleichbaren Eigenschaften des eindimensionalen RIEMANN-Integrals. Wir zeigen nun die Translationsinvarianz.

Sei zunächst $n = 1$ und $f \in \mathcal{C}_c(\mathbb{R}^1)$. Dann existiert ein kompaktes Intervall $[\alpha_1, \beta_1]$, sodass $\text{supp}(f) \subset [\alpha_1, \beta_1]$ gilt. Dann gilt $\text{supp}(\theta_y(f)) \subset [\alpha_1 + y, \beta_1 + y]$ und mit der Substitutionsregel aus Analysis I

$$\int_{\mathbb{R}} f(x) dx = \int_{\alpha_1}^{\beta_1} f(x) dx = \int_{\alpha_1+y}^{\beta_1+y} f(x-y) dx = \int_{\alpha_1+y}^{\beta_1+y} \theta_y(f(x)) dx = \int_{\mathbb{R}} \theta_y(f)(x) dx.$$

Für $n > 1$ folgt wie bei dem Nachweis der anderen Eigenschaft durch n -fache Anwendung des obigen Arguments. \square

Unser nächstes großes Ziel ist es, zu zeigen, dass I durch diese drei Eigenschaften (plus Normierung) eindeutig festgelegt ist.

Lemma 6.1.6 (eine gewisse Stetigkeitseigenschaft)

Seien $J : \mathcal{C}_c(\mathbb{R}^d) \rightarrow \mathbb{R}$ ein lineares monotones Funktional sowie $f, f_1, f_2, \dots \in \mathcal{C}_c(\mathbb{R})$, sodass ein achsenparalleler kompakter Quader Q existiert, sodass $\text{supp}(f_1), \text{supp}(f_2), \dots \subset Q := \prod_{i=1}^d \mathcal{I}_i = \prod_{i=1}^d [a_i, b_i]$ und $\lim_{n \rightarrow \infty} \|f_n - f\|_{\infty} = 0$ gilt.

Dann gilt $J(f_n) \xrightarrow{n \rightarrow \infty} J(f)$.

Beweis. Sei $\Phi : \mathbb{R}^d \rightarrow [0, 1]$ eine stetige Abbildung mit kompaktem Träger und der Eigenschaft $\Phi|_Q = 1$. Dann gilt

$$-\|f_n - f\|_{\infty} \cdot \Phi \leq f_n - f \leq \|f_n - f\|_{\infty} \cdot \Phi,$$

denn für alle $x \notin Q$ gilt $f_k(x) = 0$ und somit auch $0 = \lim_{k \rightarrow \infty} f_k(x) = f(x)$ ¹¹ und somit gilt $\text{supp}(f) \subset Q$ und für alle $x \notin Q$

$$-\|f_k - f\|_{\infty} \Phi(x) \leq 0 = f_k(x) - f(x) \leq \|f_k - f\|_{\infty} \Phi(x)$$

und für $x \in Q$

$$-\|f_k - f\|_{\infty} \Phi(x) = -\|f_k - f\|_{\infty} \leq f_k(x) - f(x) \leq \|f_k - f\|_{\infty} = \|f_k - f\|_{\infty} \Phi(x).$$

¹¹Alternativ: Es gilt $f_k(x) = 0$ für jedes $k \in \mathbb{N}$ und für alle $x \in \mathbb{R}^d \setminus Q$. Sei nun $\varepsilon > 0$ beliebig, dann folgt aus der gleichmäßigen Konvergenz von $f_k \rightarrow f$ $\|f_k(x) - f(x)\| = \|f(x)\| < \varepsilon$ für alle $x \in \mathbb{R}^d \setminus Q$, somit muss $f_k(x) - f(x) = f(x) = 0$ für alle $x \in \mathbb{R}^d \setminus Q$ gelten.

Durch Anwendung von J erhält man

$$-\underbrace{\|f_n - f\|_\infty \cdot \overbrace{J(\Phi)}^{\geq 0}}_{\xrightarrow{n \rightarrow \infty} 0} \leq J(f_n) - J(f) \leq \underbrace{\|f_n - f\|_\infty \cdot \overbrace{J(\Phi)}^{\geq 0}}_{\xrightarrow{n \rightarrow \infty} 0}.$$

Das Lemma sagt also aus, dass aus der Konvergenz $f_k \rightarrow f$ in dem metrischen Raum $(\mathcal{C}_c(\mathbb{R}^d), \|\cdot\|_\infty)$ in dem metrischen Raum $(\mathbb{R}, |\cdot|)$ die Konvergenz $J(f_k) \rightarrow J(f)$ folgt. \square

Bemerkung 6.1.7 So eine Funktion Φ erhalten wir, indem wir $\Phi = \bigotimes_{i=1}^d \varphi_i(x_i)$ setzen, wobei für jedes $\varphi \in \mathcal{C}_c(\mathbb{R})$ $\varphi_i(x) \in [0, 1]$ und $\varphi_i|_{\mathcal{I}_i} \equiv 1$ gilt.

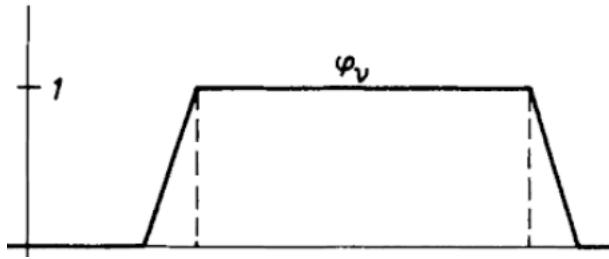


Abb. 30: Beispiel für so eine Funktion φ . [Quelle: Forster]

23.10.2018

Bemerkung 6.1.8 Um zu erkennen, wie wichtig der Quader Q für das obige Lemma ist, konstruieren wir ein Gegenexample für die obige Aussage.

Sei dafür

$$f_n(x) = \begin{cases} \frac{1}{n}, & \text{wenn } |x| \leq n, \\ 0, & \text{sonst.} \end{cases} \quad (80)$$

Nun gilt $\|f_n - f\|_\infty = \frac{1}{n} \xrightarrow{n \rightarrow \infty} 0$, also konvergiert die Folge $(f_n)_{n \in \mathbb{N}}$ gegen die konstante Nullfunktion. Wir finden jedoch keine Quader Q , außerhalb dessen die Funktionenfolge für alle n verschwindet.

Nun gilt aber $J(f_n) = (n - (-n)) \cdot \frac{1}{n} = 2$ für alle $n \in \mathbb{N}$, also konstant und somit gilt auch nicht $J(f_n) \xrightarrow{n \rightarrow \infty} J(f) = J(0) = 0$.

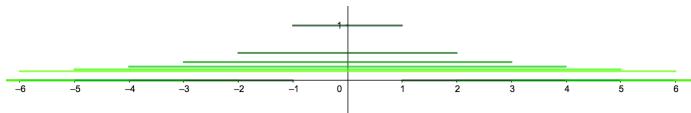


Abb. 31: f_j für $j \in \{1, \dots, 6\}$. [Quelle: GeoGebra6]

Im Folgenden wollen wir beliebige Funktionen in $\mathcal{C}_c(\mathbb{R}^d)$ mit Hilfe spezieller Funktionen besonders einfacher Gestalt gleichmäßig approximieren. Dazu betrachten wir zunächst die reskalierte Version einer gewissen Zackenfunktion.

\mathcal{C}_c -Teilung der Eins

Wir definieren zunächst $\psi \in \mathcal{C}_c(\mathbb{R}^d)$ durch

$$\psi : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto (1 - |x|)_+ = \begin{cases} 1 - |x|, & \text{für } |x| \leq 1, \\ 0, & \text{sonst.} \end{cases}$$

Damit definieren wir nun für $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ die d -dimensionale Variante von ψ :

$$\Psi : \mathbb{R}^d \rightarrow [0, 1], x \mapsto \psi^{\otimes d}(x) := \prod_{i=1}^d \psi(x_i).$$

Also gilt $\text{supp}(\Psi) = [-1, 1]^d$ und $\Psi \in \mathcal{C}_c(\mathbb{R}^d)$.

Schließlich setzen wir für $\varepsilon > 0$

$$\Psi_\varepsilon(x) := \Psi\left(\frac{x}{\varepsilon}\right).$$

Dann ist $\Psi_1 \equiv \Psi$ und $\text{supp}(\Psi_\varepsilon) = [-\varepsilon, \varepsilon]^d$ der Würfel mit Kantenlänge 2ε .

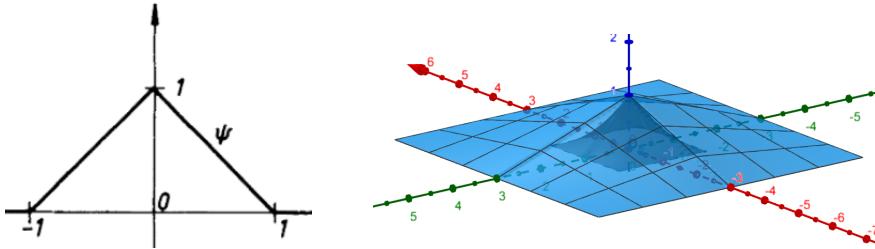


Abb. 32: Die Zacken- oder Hütchenfunktion ψ sowie $\Psi(x)$ (schwarz) und $\Psi_3(x)$ (blau) für $d = 2$. [Quelle: Forster und GeoGebra6]

Dann gilt für alle $x \in \mathbb{R}$

$$\sum_{k \in \mathbb{Z}} \theta_k(\psi(x)) = 1,$$

was man leicht an dem folgenden Bild sieht.

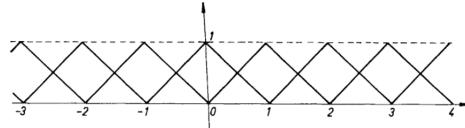


Abb. 33: [Quelle: Forster]

Analog gilt für alle $x \in \mathbb{R}^d$

$$\sum_{k \in \mathbb{Z}^d} \theta_k(\Psi(x)) = 1,$$

und analog für alle $x \in \mathbb{R}^d$ und jedes beliebige $\varepsilon > 0$

$$\sum_{k \in \mathbb{Z}^d} \theta_{k\varepsilon}(\Psi_\varepsilon(x)) = 1,$$

Unsere Teilung der Eins ist also die Familie $(\theta_{k\varepsilon}(\Psi_\varepsilon(x)))_{k \in \mathbb{Z}^d}$.

Nun ist $\text{supp}(\theta_{k\varepsilon}(\Psi_\varepsilon(x))) = k\varepsilon + [-\varepsilon, \varepsilon]^d$ der Würfel mit Mittelpunkt $k\varepsilon$ und Seitenlänge 2ε . Außerdem haben $\text{supp}(\theta_{k\varepsilon}(\Psi_\varepsilon(x)))$ und $(\theta_{\tilde{k}\varepsilon}(\Psi_\varepsilon(x)))$ einen nicht trivialen Schnitt¹², wenn $\|k - \tilde{k}\| \leq 1$ gilt.

Lemma 6.1.9

Seien $\varepsilon > 0$ beliebig und $J : \mathcal{C}_c(\mathbb{R}^d) \rightarrow \mathbb{R}$ ein lineares, translationsinvariantes Funktional. Dann gilt

$$J(\Psi_{\frac{\varepsilon}{2}}) = 2^{-d} J(\Psi_\varepsilon).$$

Beweis. Sei zunächst $d = 1$. Mithilfe der folgenden Skizze sieht man leicht, dass für jedes $\varepsilon > 0$ das folgende gilt:

$$\psi_\varepsilon = \frac{1}{2} \cdot \theta_{-\frac{\varepsilon}{2}}(\psi_{\frac{\varepsilon}{2}}) + \psi_{\frac{\varepsilon}{2}} + \frac{1}{2} \cdot \theta_{\frac{\varepsilon}{2}}(\psi_{\frac{\varepsilon}{2}})$$

Aufgrund der Linearität und Translationsinvarianz von J gilt nun

$$J(\psi_\varepsilon) = \frac{1}{2} \cdot J(\theta_{-\frac{\varepsilon}{2}}(\psi_{\frac{\varepsilon}{2}})) + J(\psi_{\frac{\varepsilon}{2}}) + \frac{1}{2} \cdot J(\theta_{\frac{\varepsilon}{2}}(\psi_{\frac{\varepsilon}{2}})) = 2 \cdot \frac{1}{2} \cdot J(\theta_{\frac{\varepsilon}{2}}(\psi_{\frac{\varepsilon}{2}})) + J(\psi_{\frac{\varepsilon}{2}}) = 2J(\psi_{\frac{\varepsilon}{2}}).$$

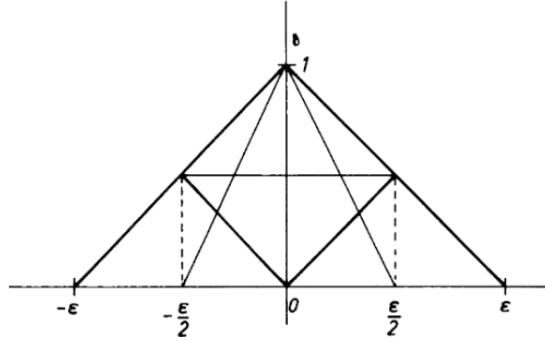


Abb. 34: [Quelle: Forster]

Für $d > 1$ gilt mit $\wp := \{-1, 0, 1\}$

$$\begin{aligned} \Psi_\varepsilon(x) &= \prod_{i=1}^d \psi_\varepsilon(x_i) = \prod_{i=1}^d \frac{1}{2} \cdot \theta_{-\frac{\varepsilon}{2}}(\psi_{\frac{\varepsilon}{2}}(x_i)) + \psi_{\frac{\varepsilon}{2}} + \frac{1}{2} \cdot \theta_{\frac{\varepsilon}{2}}(\psi_{\frac{\varepsilon}{2}}(x_i)) \\ &= \prod_{i=1}^d \sum_{\alpha_i \in \wp} \left(\frac{1}{2}\right)^{|\alpha_i|} \theta_{\frac{\alpha_i \cdot \varepsilon}{2}}(\psi_{\frac{\varepsilon}{2}})(x_i) \\ &= \sum_{\alpha_1, \dots, \alpha_d \in \wp} \left(\prod_{i=1}^d \left(\frac{1}{2}\right)^{|\alpha_i|} \right) \left(\prod_{i=1}^d \theta_{\frac{\alpha_i \cdot \varepsilon}{2}}(\psi_{\frac{\varepsilon}{2}})(x_i) \right). \end{aligned}$$

Mit $\alpha = (\alpha_1, \dots, \alpha_d)$ gilt also (Übung!)

$$\Psi_\varepsilon = \sum_{\alpha \in \wp^d} \left(\prod_{i=1}^d \left(\frac{1}{2}\right)^{|\alpha_i|} \right) \theta_{\frac{\alpha \cdot \varepsilon}{2}}(\Psi_{\frac{\varepsilon}{2}}(x)).$$

¹²Nullmengen zählen auch als trivialer Schnitt.

Somit folgt aufgrund der Linearität und Translationsinvarianz von J

$$\begin{aligned} J(\Psi_\varepsilon) &= \sum_{\alpha \in \wp^d} \left(\prod_{i=1}^d \left(\frac{1}{2} \right)^{|\alpha_i|} \right) J(\Psi_{\frac{\varepsilon}{2}}) = \prod_{i=1}^d \left(\sum_{\alpha_i \in \wp} \left(\left(\frac{1}{2} \right)^{|\alpha_i|} \right) \right) J(\Psi_{\frac{\varepsilon}{2}}) \\ &= \prod_{i=1}^d \left(\frac{1}{2} + 1 + \frac{1}{2} \right) J(\Psi_{\frac{\varepsilon}{2}}) = 2^d J(\Psi_{\frac{\varepsilon}{2}}). \end{aligned} \quad \square$$

Lemma 6.1.10 (Approximationslemma)

Für jedes $f \in \mathcal{C}_c(\mathbb{R}^d)$ gilt

$$\lim_{\varepsilon \searrow 0} \left\| f - \underbrace{\sum_{k \in \mathbb{Z}^d} f(k\varepsilon) \theta_{k\varepsilon}(\Psi_\varepsilon)}_{=:G_\varepsilon} \right\|_\infty = 0.$$

Beweis. Wohldefiniertheit: Da die Funktion f einen kompakten Träger hat, existieren nur endlich viele $k \in \mathbb{Z}^d$ mit $f(k\varepsilon) \neq 0$ und somit ist die Summe endlich.

Aussage: Sei $\eta > 0$ vorgegeben. Nun ist f auf einem genügend großen Kompaktum, welcher $\text{supp}(f)$ enthält, gleichmäßig stetig. Somit existiert ein $\varepsilon > 0$, sodass $|f(x) - f(y)| < \eta$ für alle $x, y \in \mathbb{R}^d$ mit $\|x - y\|_\infty \leq \varepsilon$ gilt.

Wir zeigen nun, dass $\|G_\varepsilon\|_\infty < \eta$ gilt. Sei dafür

$$\mathcal{A}(x) := \{k \in \mathbb{Z}^d : x \in \text{supp}(\theta_{k\varepsilon}(\Psi_\varepsilon))\}$$

für eine gegebenes $x \in \mathbb{R}$ die Menge aller k , sodass f einen nichttrivialen Beitrag zu der Summe liefert. Diese Menge ist nach dem obigen Argument der Wohldefiniertheit endlich.

Dann gilt, weil $f(x)\theta_{k\varepsilon}(\Psi_\varepsilon)$ die Teilung der Eins ist,

$$\|G_\varepsilon(x)\|_\infty = \left\| f(x) - \sum_{k \in \mathcal{A}(x)} f(k\varepsilon) \theta_{k\varepsilon}(\Psi_\varepsilon) \right\|_\infty \stackrel{\triangle\neq}{\leq} \sum_{k \in \mathcal{A}(x)} |f(x) - f(k\varepsilon)| \theta_{k\varepsilon}(\Psi_\varepsilon)$$

Nun ist $\text{supp}(\theta_{k\varepsilon}(\Psi_\varepsilon))$ ein Würfel um den Punkt $k\varepsilon$ mit Seitenlänge 2ε . Jeder Punkt in diesem Würfel ist also bezüglich der Supremumsnorm höchstens ε von dem Mittelpunkt entfernt. Also können wir die Voraussetzung von oben einsetzen: Weil $\|x - k\varepsilon\| \leq \varepsilon$ gilt, ist $|f(x) - f(k\varepsilon)| < \eta$ und somit gilt

$$|G_\varepsilon(x)| < \eta \sum_{k \in \mathcal{A}(x)} \theta_{k\varepsilon}(\Psi_\varepsilon) = \eta. \quad \square$$

SATZ 6.1.3: EINDEUTIGKEIT DES HAARSCHEN MASS

Sei $J : \mathcal{C}_c(\mathbb{R}^d) \rightarrow \mathbb{R}$ ein LMTF. Dann existiert ein $c \in \mathbb{R}$, sodass für alle $f \in \mathcal{C}_c(\mathbb{R}^d)$ gilt

$$J(f) = cI(f) = c \int_{\mathbb{R}^d} f(x) dx.$$

Beweis. Wir zeigen die Aussage mit $c = J(\Psi)$. Ist J monoton wachsend, so ist $c \geq 0$, ist J monoton fallend, so ist $c \leq 0$. (???)

n -malige Iteration von Lemma 6.1.9 ergibt

$$J(\Psi_{2^{-n}}) = 2^{-nd} J(\Psi) = c \cdot 2^{-nd} \stackrel{6.1.9}{=} c \cdot \frac{I(\Psi_{2^{-n}})}{I(\Psi)}$$

Nun ist trivialerweise $I(\psi) = 1$ und nach Hausaufgabe 8 gilt dann

$$I(\Psi) = \int_{R^n} \Psi(x) dx = \int \prod_{i=1}^d \psi(x_i) d(x_1, \dots, x_n) = \prod_{i=1}^d \int \psi(x_i) dx_i = 1^d = 1.$$

und somit ist $J(\Psi_{2^{-n}}) = c \cdot I(\Psi_{2^{-n}})$.

Sei nun $f \in \mathcal{C}_c(\mathbb{R}^d)$. Dann definieren wir

$$f_n := \sum_{k \in \mathbb{Z}^d} f(k2^{-n}) \theta_{k2^{-n}}(\Psi_{2^{-n}}).$$

Dann gilt

$$\begin{aligned} J(f_n) &= \sum_{k \in \mathbb{Z}^d} f(k2^{-n}) J(\Psi_{2^{-n}}) = \sum_{k \in \mathbb{Z}^d} f(k2^{-n}) c \cdot I(\Psi_{2^{-n}}) \\ &= c \cdot I \left(\sum_{k \in \mathbb{Z}^d} f(k2^{-n}) \theta_{k2^{-n}}(\Psi_{2^{-n}}) \right) = cI(f_n). \end{aligned}$$

Dann konvergiert die Folge $(f_n)_{n \in \mathbb{N}}$ gegen f im Sinne von Lemma 6.1.6 (einen geeigneten Quader findet man leicht, und Lemma 6.1.10 wird benutzt) also folgt im Grenzübergang $J(f) = c \cdot I(f)$ aus Lemma 6.1.6. \square

DEFINITION 6.1.11 ((UNNORMIERTES) HAARSCHES MASS)

Jedes LMTF: $\mathcal{C}_c(\mathbb{R}^d) \rightarrow \mathbb{R}$ heißt HAARSches Maß.

Bemerkung 6.1.12 Benannt nach dem ungarischen Mathematiker ALFRÉD HAAR, hilft es, Ergebnisse der Maßtheorie in der Gruppentheorie anwendbar zu machen. Der Name "Maß" hat eher historische Gründe, es wir natürlich vom LEBESGUE-Maß erzeugt. Manchmal ist es einfacher festzustellen, wie ein Maß gewisse Funktionenklassen integriert.

6.2 Die Transformationsformel

Dies ist die d -dimensionale Verallgemeinerung der Substitutionsregel aus Analysis I:

$$\int f(\varphi(t))\varphi'(t)dt = \int f(x)dx,$$

wobei φ eine bijektiv Abbildung ist.

Das machen wir Schrittweise: für lineare φ (also Matrizen), erst für diagonale und orthogonale Matrizen einzeln.

Lemma 6.2.1

Ist $J : \mathcal{C}_c(\mathbb{R}^d) \rightarrow \mathbb{R}$ ein lineares monotones translationsinvariantes Funktional (LMTF) und $A \in \mathrm{GL}(d, \mathbb{R})$ eine invertierbare Matrix, so ist die Abbildung $f \mapsto J(f \circ A)$ auch ein LMTF.

Beweis. Wohldefiniertheit. (Selbst) Da Kompositionen stetiger Funktionen stetig sind und $\mathrm{supp}(f \circ A) = A^{-1}(\mathrm{supp}(f))$ als stetiges Urbild einer abgeschlossenen Menge abgeschlossen ist und somit ist $\mathrm{supp}(f \circ A) \subset \mathrm{supp}(f \circ A)$ als abgeschlossene Teilmenge einer kompakten Menge kompakt, wir dürfen also $J(f \circ A)$ schreiben.

Aussage. Die Linearität und Monotonie ist klar.

Um die Translationsinvarianz zu zeigen, zeigen wir zunächst

$$(\theta_y f) \circ A = \theta_{A^{-1}y}(f \circ A) \quad \forall y \in \mathbb{R}^d$$

gilt, denn dann folgt, weil J translationsinvariant ist,

$$J((\theta_y f) \circ A) = J(\theta_{A^{-1}y}(f \circ A)) = J(f \circ A).$$

Für alle $x \in \mathbb{R}^d$ gilt

$$(\theta_y f \circ A)(x) = \theta_y f(Ax) \stackrel{\text{Def.}}{=} f(Ax - y) = f(A(x - A^{-1}y)) \stackrel{\text{Def.}}{=} \theta_{A^{-1}y}(f \circ A)(x). \quad \square$$

Dieser Satz (auch: Transformationssatz für orthogonale Matrizen) erhält seine Bezeichnung daher, dass nach dieser Definition lineare Abbildungen, die durch orthogonale Abbildung beschrieben werden, Objekte nur Dreh(spiegeln), aber nicht strecken oder stauchen, also nur "bewegen".

SATZ 6.2.1: BEWEGUNGSGINVARIANZ DES RIEMANN-INTEGRALS

Für jedes $f \in \mathcal{C}_c(\mathbb{R}^d)$ und jede orthogonale Matrix $A \in \mathcal{O}(d)$ gilt

$$\int_{\mathbb{R}^d} f(Ax)dx = \int_{\mathbb{R}^d} f(x)dx.$$

Beweis. Wir definieren $I(f) := \int_{\mathbb{R}^d} f(x)dx$. Dann ist die Abbildung

$$f \mapsto \int f(Ax)dx = I(f \circ A)$$

nach Lemma 6.2.1 eine LMTF, da auch $f \circ A \in \mathcal{C}_c$ ist.

Nach Satz 6.1.3 existiert eine Konstante $c \in \mathbb{R}$, sodass für alle $f \in \mathcal{C}_c(\mathbb{R}^d)$ gilt

$$I(f) = c \cdot I(f \circ A).$$

Wir betrachten $f(x) := (1 - \|x\|)_+$. Es gilt $f \in \mathcal{C}_c(\mathbb{R}^d)$, weil $\text{supp}(f) = [-1, 1]^d$ kompakt ist, und

$$\lim_{x \nearrow 1} f(x) = \lim_{x \nearrow 1} 1 - x = 0 = f(\pm 1) = \lim_{x \searrow -1} x - 1 = \lim_{x \searrow -1} f(x).$$

Ferner gilt $f \circ A = f$, weil orthogonale Matrizen die euklidische Norm erhalten und somit ist $c = 1$. \square

Lemma 6.2.2 (Transformationssatz für Diagonalmatrizen)

Seien $\alpha_1, \dots, \alpha_d \in \mathbb{R}_{>0}$ die Diagonaleinträge der Matrix $A := \text{diag}(\alpha_1, \dots, \alpha_d) \in \mathbb{R}^{d \times d}$. Dann gilt für alle $f \in \mathcal{C}_c(\mathbb{R}^d)$

$$\int_{\mathbb{R}^d} f(Ax) dx = \frac{1}{\prod_{i=1}^d \alpha_i} \int_{\mathbb{R}^d} f(x) dx = \frac{1}{|\det(A)|} \int_{\mathbb{R}^d} f(x) dx.$$

Beweis. (Selbst.) Da $f(Ax) \in \mathcal{C}_c(\mathbb{R}^d)$ ist, gilt mithilfe der Substitutionsregel

$$y(x_1) := \alpha_1 x_1 \implies dx_1 = \frac{dy_1}{\alpha_1}$$

aus Analysis I für einen kompakten achsenparallelen Quader $Q := \prod_{i=1}^d [a_i, b_i]$, welche den Träger von $f(Ax)$ enthält, und aufgrund der Linearität der eindimensionalen Integrale

$$\begin{aligned} \int_{\mathbb{R}^d} f(Ax) dx &= \int_{a_d}^{b_d} \left(\dots \int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} f(\alpha_1 x_1, \dots, \alpha_d x_d) dx_1 \right) dx_2 \right) \dots dx_d \\ &= \int_{a_d}^{b_d} \left(\dots \int_{a_2}^{b_2} \left(\int_{\alpha_1 a_1}^{\alpha_1 b_1} f(y_1, \alpha_2 x_2, \dots, \alpha_d x_d) \frac{dy_1}{\alpha_1} \right) dx_2 \right) \dots dx_d \\ &= \int_{a_d}^{b_d} \left(\dots \int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} \frac{1}{\alpha_1} f(x_1, \alpha_2 x_2, \dots, \alpha_d x_d) dx_1 \right) dx_2 \right) \dots dx_d. \end{aligned}$$

Durch weiteres $d-1$ -malige Anwenden des obigen Arguments folgt die Behauptung. \square

Bemerkung Es ist auch möglich $\alpha_i < 0$ zu benutzen ($\alpha_i = 0$ für ein $i \implies \det(A) = 0 \implies A$ nicht invertierbar und dann muss das Integral Null sein, da $\dim(\mathfrak{I}(\mathbb{R}^d)) = d-1$ gilt), aber die Notation lässt keine Regeln der Form

$$-\int_{-a}^{-b} f(x) dx = \int_a^b f(-y) dy$$

zu, weil wir im \mathbb{R}^1 von links nach rechts integrieren, hier benutzen wir aber nicht \int_a^b sondern $\int_{\mathbb{R}^d}$.

Bemerkung 6.2.3 Nun kann man einen einfachen Beweis für Satz 6.1.1 geben, indem eine Permutationsmatrix A mit $A(x_1, \dots, x_d) = (x_{i_1}, \dots, x_{i_d})$, welche orthogonal ist, in Satz 6.2.1 einsetzt.

Lemma 6.2.4 (Singulärwertzerlegung)

Für jede invertierbare Matrix $A \in \text{GL}(d, \mathbb{R})$ existieren positive Zahlen $\alpha_1, \dots, \alpha_d > 0$ und orthogonale Matrizen $S_1, S_2 \in \mathcal{O}(d)$, sodass gilt

$$A = S_1 \text{diag}(\alpha_1, \dots, \alpha_d) S_2.$$

Beweis. Die Matrix $A^T A$ ist nach Linearer Algebra symmetrisch, also orthogonal diagonalisierbar. Somit existieren eine orthogonale Matrix $S \in \mathcal{O}(d)$ und Zahlen $\alpha_1, \dots, \alpha_d \in \mathbb{R}$ mit

$$S^T A^T A S = \underbrace{\text{diag}(\alpha_1^2, \dots, \alpha_d^2)}_{=: D^2},$$

wobei die Diagonaleinträge von D^2 die positiven Eigenwerte von $A^T A$ sind, und somit als α_i^2 geschrieben werden können.¹³ Ohne Beschränkung der Allgemeinheit können wir annehmen, dass alle $\alpha_i > 0$ sind.

Sei nun $S_1 := ASD^{-1}$. Dann ist S_1 orthogonal, denn weil D eine Diagonalmatrix ist (*), gilt

$$S_1^T S_1 = D^{-1T} S^T A^T A S D^{-1} = D^{-1T} D^2 D^{-1} = D^{-1T} D \stackrel{(*)}{=} D^{-1} D = E.$$

Sei nun $S_2 = S^{-1} (= S^T) \in \mathcal{O}(d)$. Dann gilt

$$S_1 D S_2 = A S D^{-1} D S^{-1} = A. \quad \square$$

SATZ 6.2.2: TRANSFORMATIONSFORMEL FÜR BIJEKTIVE LINEARE KOORDINATIONSWECHSEL: $d(Ax) = |\det(A)|dx$

Für alle invertierbaren Matrizen $A \in \text{GL}(d, \mathbb{R})$ und alle $f \in \mathcal{C}_c(\mathbb{R}^d)$ gilt

$$\int_{\mathbb{R}^d} f(Ax) |\det(A)| dx = \int_{\mathbb{R}^d} f(y) dy$$

Beweis. Nach Lemma 6.2.4 existieren orthogonale Matrizen $S_1, S_2 \in \mathcal{O}(d)$ und Zahlen $\alpha_1, \dots, \alpha_d \in \mathbb{R}_{>0}$ mit $A = S_1 D S_2$ für $D := \text{diag}(\alpha_1, \dots, \alpha_d)$.

Dann ist nach dem Determinantenmultiplikationssatz $|\det(A)| = \det(D)$. Nach Lemma 6.2.2 gilt für alle $f \in \mathcal{C}_c(\mathbb{R}^d)$

$$I(f \circ D) \det(D) = I(f)$$

und Satz 6.2.1 zeigt, dass für alle $f \in \mathcal{C}_c(\mathbb{R}^d)$ gilt

$$I(f \circ S_1) = I(f) = I(f \circ S_2).$$

Somit folgt für alle $f \in \mathcal{C}_c(\mathbb{R}^d)$

$$\begin{aligned} I(f) &= I(f \circ S_1) = \det(D) \cdot I(f \circ S_1 \circ D) \\ &= |\det(A)| \cdot I(f \circ S_1 \circ D \circ S_2) = |\det(A)| \cdot I(f \circ A). \end{aligned} \quad \square$$

DEFINITION 6.2.5 (SPAT / PARALLELOTOP)

Für Spaltenvektoren $\alpha_1, \dots, \alpha_d \in \mathbb{R}^d$ und deren Matrix $A = (\alpha_1, \dots, \alpha_d) \in \mathbb{R}^{d \times d}$ nennt man

$$P(A) := P(\alpha_1, \dots, \alpha_d) := \left\{ \sum_{k=0}^d \lambda_k \alpha_k : \lambda_1, \dots, \lambda_d \in [0, 1] \right\}$$

¹³[Aus dem Forster] Alternativ gilt

$$\alpha_k = e_k^T D^2 e_k = e_k^T S^T A^T A S e_k = \|A S e_k\|^2 > 0.$$

Spat den von $\alpha_1, \dots, \alpha_d$ aufgespannten **Spat**.

[Bild: für $d = 2$: Parallelogramm]

Bemerkung 6.2.6 Man kann leicht sehen, dass das Volumen eines Spats Null ist, wenn zwei Vektoren linear abhängig sind.

Beispiel 6.2.7 (Volumen eines Spats / Parallelepipeds)

Wir wollen zeigen, dass das Volumen von $P(A)$ für eine invertierbare Matrix A , $\text{Vol}_d(P(A)) = |\det(A)|$ ist, wobei wir (in Abschnitt (2.6)) das Volumen eines Kompakts $K \subset \mathbb{R}^d$ definieren als

$$\text{Vol}_d(K) = \int_K 1 dx = \int_{\mathbb{R}^d} \mathbf{1}_K(x) dx,$$

wobei dieser Ausdruck zunächst nur für achsenparallele Quader definiert wurde.

Wir wissen (Übung!), dass $P(A) = A([0, 1]^d)$. Daher können wir mit Hilfe von Satz 6.2.2 für $A \in \text{GL}(d, \mathbb{R})$ definieren

$$\begin{aligned} \text{Vol}_d(P(A)) &= \int_{A([0, 1]^d)} 1 dx = \int_{\mathbb{R}^d} \mathbf{1}_{A([0, 1]^d)}(x) dx = \int_{\mathbb{R}^d} \mathbf{1}_{[0, 1]^d}(A^{-1}x) dx \\ &\stackrel{\text{Trafo}}{=} \frac{1}{|\det(A^{-1})|} \int \mathbf{1}_{[0, 1]^d}(x) dx = |\det(A)|. \end{aligned}$$

Nun bringe ich den Transformationsatz für eine C^1 -invertierbare Abbildung φ und Stelle der linearen Abbildung A . Aber vorher das helfende

Lemma 6.2.8

Seien $U, V \subset \mathbb{R}^d$ offene Teilmengen und sei $\varphi = (\varphi_1 \dots \varphi_n)^T : U \rightarrow V$ eine C^1 -invertierbare Funktion mit Differentialmatrix $D\varphi(x) = \left(\frac{\partial \varphi_i(x)}{\partial x_j} \right)_{i,j \in \{1, \dots, d\}}$. Dann gilt $(D\varphi(a))^{-1} = (D\varphi^{-1})(\varphi(a))$.

Beweis. Kettenregel. □

DEFINITION 6.2.9

Mit $\mathcal{C}_c(U)$ bezeichnen wir die Menge alle stetigen Funktionen mit kompaktem Träger in U .

Bemerkung 6.2.10 Für $f \in \mathcal{C}_c(U)$ liegt ihre Fortsetzung mit Null

$$\tilde{f}(x) := \begin{cases} f(x) & x \in U \\ 0 & x \in \mathbb{R}^d \setminus U \end{cases}$$

in $\mathcal{C}_c(\mathbb{R}^d)$, weil die Funktion f außerhalb des kompakten Trägers, welcher aufgrund der Offenheit von U einen positiven Abstand zu U^C hat, verschwindet, somit in U^C stetig mit 0 fortgesetzt werden kann.

Bemerkung 6.2.11 Ist $f \in \mathcal{C}_c(V)$, so ist $f \circ \varphi \in \mathcal{C}_c(U)$, man argumentiert wie bei der Wohldefiniertheit am Anfang dieses Kapitels.

SATZ 6.2.3: TRANSFORMATIONSSATZ FÜR $\mathcal{C}_c(\mathbb{R}^d)$ -FUNKTIONEN

Seien $U, V \subset \mathbb{R}^n$ offene Teilmengen und $\varphi : U \rightarrow V$ eine \mathcal{C}^1 -invertierbare Funktion. Dann gilt für jedes $f \in \mathcal{C}_c(V)$

$$\int_V f(y) dy = \int_U f(\varphi(x)) \cdot |\det(D\varphi(x))| dx. \quad (81)$$

Beweis. Wir approximieren f mit unserer Teilung der 1 und φ mit einem linearen Polynom, dem TAYLOR-Polynom erster Ordnung. Der Entwicklungspunkt ist hierbei der Punkt, um welchen die Zackenfunktion zentriert ist.

Sei

$$f_\varepsilon(y) := \sum_{k \in \mathbb{Z}^d} f(k\varepsilon) \theta_{k\varepsilon} \Phi_\varepsilon(y).$$

Dann gilt nach Lemma 6.1.10 und Lemma 6.1.6

$$\|f_\varepsilon - f\|_\infty \xrightarrow{\varepsilon \searrow 0} 0 \quad \text{und} \quad I(f_\varepsilon) \xrightarrow{\varepsilon \searrow 0} I(f).$$

Für die Approximation der rechten Seite seien

$$g(x) := f(\varphi(x)) |\det(D\varphi(x))| \quad \text{und} \quad g_\varepsilon(x) := f_\varepsilon(\varphi(x)) |\det(D\varphi(x))|$$

Dann gilt mit analoger Begründung

$$\frac{g_\varepsilon}{\|\cdot\|_\infty} \xrightarrow{\varepsilon \rightarrow 0} g \quad \text{und} \quad I(g_\varepsilon) \xrightarrow{\varepsilon \rightarrow 0} I(g).$$

Damit wir Lemma 6.1.6 überhaupt anwenden können, existiert für jedes $\varepsilon > 0$ ein Kompaktum, welche alle Träger von f_ε bzw. g_ε enthält.

Also ist noch zu zeigen, dass

$$|I(f_\varepsilon) - I(g_\varepsilon)| = \left| \int_V f_\varepsilon(y) dy - \int_U g_\varepsilon(x) dx \right| \xrightarrow{\varepsilon \searrow 0} 0.$$

gilt.

Da nur $\mathcal{O}(\varepsilon^{-d})$ Terme der Summe einen nichtrivialen Beitrag leisten, gilt

$$\begin{aligned} |I(f_\varepsilon) - I(g_\varepsilon)| &= \left| \sum_{k \in \mathbb{Z}^d} f(k\varepsilon) I(\theta_{k\varepsilon} \Psi(k\varepsilon)) - \sum_{k \in \mathbb{Z}^d} f(k\varepsilon) I((\theta_{k\varepsilon} \Psi(k\varepsilon)) \circ \varphi \cdot |\det(\varphi(x))|) \right| \\ &\leq \sum_{\substack{k \in \mathbb{Z}^d: \\ k \in \text{supp}(f)}} \|f\|_\infty |I(\theta_{k\varepsilon} \Psi(k\varepsilon)) - I((\theta_{k\varepsilon} \Psi(k\varepsilon)) \circ \varphi \cdot |\det(\varphi(x))|)| \\ &\leq \mathcal{O}(\varepsilon^{-d}) \sup_{\substack{k \in \mathbb{Z}^d: \\ k \in \text{supp}(f)}} |I(\theta_{k\varepsilon} \Psi(k\varepsilon)) - I((\theta_{k\varepsilon} \Psi(k\varepsilon)) \circ \varphi \cdot |\det(\varphi(x))|)|. \end{aligned}$$

Für ein festes $k \in \mathbb{Z}^d$ mit $k\varepsilon \in \text{supp}(f)$ definieren wir

$$a := \varphi^{-1}(k\varepsilon), \quad \text{und} \quad \mathcal{W}(k\varepsilon, \varepsilon) := \{x \in \mathbb{R}^d : \|x - k\varepsilon\| < \varepsilon\}.$$

Für die lineare Approximation von φ in $\varphi^{-1}(\mathcal{W}(k\varepsilon, \varepsilon))$ benötigen wir das TAYLOR-Polynom erster Ordnung

$$\lambda_a(x) := \varphi(a) + D\varphi(x - a) \quad (82)$$

Dann gilt (vgl. Analysis II)

$$\|\lambda_a(x) - \varphi(x)\|_\infty \leq o(\|x - a\|_\infty) \leq o(\varepsilon),$$

wobei die letzte Ungleichung für $x \in \varphi^{-1}(\mathcal{W}(k\varepsilon, \varepsilon)) \subset \varphi^{-1}(\mathcal{W}(a, C\varepsilon))$, wobei die Konstante $C > 0$ geeignet gewählt ist und aufgrund der Beschränktheit von φ' existiert, gilt.

Nach Satz 6.2.2¹⁴ gilt mit $h := \theta_{k\varepsilon} \Psi_\varepsilon$

$$I(h) = \int \underbrace{\Psi_\varepsilon(\lambda_a(x) - k\varepsilon) | \det(D\varphi(a))| dx}_{=\Psi_\varepsilon(D\varphi(a)(x-a))} = \int h(\lambda_a(x)) | \det(D\varphi(a))| dx$$

Also gilt für ein festes $k \in \mathbb{Z}^d$

$$\begin{aligned} & |I(\theta_{k\varepsilon} \Psi(k\varepsilon)) - I((\theta_{k\varepsilon} \Psi(k\varepsilon)) \circ \varphi \cdot |\det(\varphi(x))|)| \\ &= \left| \int h(\lambda_a(x)) | \det(D\varphi(a))| dx - \int h(\varphi(x)) \det(D\varphi(x)) dx \right| \\ &\leq \int_{W(a, C\varepsilon)} |h(\lambda_a(x)) - h(\varphi(x))| |\det(D\varphi(a))| dx \\ &\quad + \int_{W(a, C\varepsilon)} h(\varphi(x)) |\det(D\varphi(a)) - \det(D\varphi(x))| dx \end{aligned}$$

Durch scharfes Hinsehen erkennt man

$$\|h(x) - h(y)\| \leq \mathcal{O}\left(\frac{1}{\varepsilon}\right) \|x - y\|.$$

und somit gilt aufgrund der Stetigkeit von der Determinante

$$\begin{aligned} &\leq \int_{W(a, C\varepsilon)} \underbrace{\mathcal{O}\left(\frac{1}{\varepsilon}\right) \|\lambda_a(x) - \varphi(x)\|_\infty \cdot \mathcal{O}(1)}_{=o(1)} dx \\ &\quad + \int_{W(a, C\varepsilon)} \underbrace{h(\varphi(x))}_{\in [0,1]} \underbrace{|\det(D\varphi(a)) - \det(D\varphi(x))|}_{o(1)} dx \\ &\leq o(1)\mathcal{O}(1) \text{Vol}(\mathcal{W}(a, C\varepsilon)) + o(1) \text{Vol}(\mathcal{W}(a, C\varepsilon)) \\ &\leq o(\varepsilon^d) + o(\varepsilon^d). \quad \square \end{aligned}$$

Bemerkung 6.2.12 Alle $o(\dots)$ und $\mathcal{O}(\dots)$ hängen nicht von k ab!

Nun erweitern wir die Menge der Funktionen, die wir integrieren können.

DEFINITION 6.2.13

Sei $V \subset \mathbb{R}^d$ eine offene Teilmenge und $f : V \rightarrow [0, \infty)$ eine stetige Abbildung.
Dann definieren wir

$$\int_V f(y) dy = \lim_{n \rightarrow \infty} \int_V f_n(y) dy,$$

wobei $f_1, \dots \in \mathcal{C}_c(V)$ und diese Folge punktweise und gleichmäßig gegen f konvergiert, $f_n \nearrow f$, aber nur, falls die rechte Seite kleiner als Unendlich ist.

¹⁴Die Abbildung λ_a ist eigentlich nur linear affin.

Bemerkung 6.2.14 Man muss für jedes f die Existenz so einer Folge und die Unabhängigkeit der rechten Seite von der Wahl der Folge zeigen (vgl. 2.5.3 und 2.5.8 und Lemma 2.5.4 (ii)).

Bemerkung 6.2.15 Eine stetige Funktion $f : V \rightarrow \mathbb{R}$ zerlegt man in ihren Positivteil f_+ und f_- und definiert

$$\int_V f(y) dy = \int f_+(y) dy - \int_V f_-(y) dy$$

Bemerkung 6.2.16 Ist $d = 1$ und V ein Intervall, so ist dieses Integral das uneigentliche RIEMANN-Integral.

Korollar 6.2.17 (Transformationssatz für stetige Funktionen)

Seien $U, V \subset \mathbb{R}^d$ offene Teilmengen und $\varphi : U \rightarrow V$ eine invertierbare Funktionen. Dann gilt für jede stetige Funktion $f : V \rightarrow \mathbb{R}^d$, welche im Sinne der vorherigen Definition integrierbar ist,

$$\int_V f(y) dy = \int_U f(\varphi(x)) |\det(D\varphi(x))| dx.$$

Beweis. Sei o.B.d.A $f \geq 0$, da wir f sonst in f_+ und f_- zerlegen können.

Sei nun $(f_n)_{n \in \mathbb{N}}$ eine Folge in $\mathcal{C}_c(V)$, wobei $f_n \nearrow f$ gleichmäßig konvergiert. Dann gilt auch

$$\underbrace{f_n \circ \varphi |\det(D\varphi)|}_{\in \mathcal{C}_c(U)} \nearrow f \circ \varphi |\det(D\varphi)|.$$

Also gilt nach dem vorherigen Satz

$$\begin{aligned} \int_V f(y) dy &= \lim_{n \rightarrow \infty} \int_V f_n(y) dy \\ &= \lim_{n \rightarrow \infty} \int_U f_n \circ \varphi(x) |\det(D\varphi(x))| dx \\ &= \int_U f \circ \varphi(x) |\det(D\varphi(x))| dx. \end{aligned} \quad \square$$

Beispiel 6.2.18 (ebene Polarkoordinaten) Wir wollen die Transformation

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto r \cdot \begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \end{pmatrix}$$

im Licht des Transformationssatzes betrachten. Definiere dazu

$$\Phi^{-1} : \mathbb{R}^2 \setminus \underbrace{\{[0, \infty) \times \{0\}\}}_{:= S} \rightarrow (0, \infty) \times (0, 2\pi), \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \sqrt{x^2 + y^2} \\ \varphi(x, y) \end{pmatrix}.$$

Dann ist die Funktion

$$\Phi : (0, \infty) \times (0, 2\pi) \rightarrow \mathbb{R}^2 \setminus S, \quad (r, \varphi) \mapsto r \cdot \begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \end{pmatrix}$$

\mathcal{C}^1 -invertierbar.

Dann gilt

$$|\det(D\Phi(r, \varphi))| = \begin{vmatrix} \cos(\varphi) & -r \sin(\varphi) \\ \sin(\varphi) & r \cos(\varphi) \end{vmatrix} = r.$$

Nach dem obigen Korollar gilt also für jede integrierbare Funktion (im Sinne der obigen Definition)

$$\begin{aligned} \int_{\mathbb{R}^2} f(y) dy &= \int_{\mathbb{R}^2 \setminus S} f(y) dy = \int_{(0, \infty) \times (0, 2\pi)} f \begin{pmatrix} r \cos(\varphi) \\ r \sin(\varphi) \end{pmatrix} |det(D\Phi(r, \varphi))| d(r, \varphi) \\ &= \int_0^\infty r \int_0^{2\pi} f \begin{pmatrix} r \cos(\varphi) \\ r \sin(\varphi) \end{pmatrix} d(\varphi, r). \end{aligned}$$

Ist f rotationssymmetrisch ($f(x) = g(\|x\|)$ für alle x und eine Funktion $g : (0, \infty) \rightarrow \mathbb{R}$), so folgt

$$\int_{\mathbb{R}^2} f(y) dy = \int_0^\infty r \int_0^{2\pi} g(r) d(\varphi, r) = 2\pi \int_0^\infty r g(r) dr$$

Bemerkung 6.2.19 In einer beliebigen Dimension d würde man r^{d-1} anstatt r und anstatt 2π die Oberfläche der d -dimensionalen Einheitskugel einsetzen.

05.11.2018

Bemerkung 6.2.20 (Anwendungsbeispiel: GAUSSsche Normalverteilung)

Die Funktion $h(x) := \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ ist die Dichte der eindimensionalen Standard- oder GAUSSschen Normalverteilung also insbesondere eine Wahrscheinlichkeitsdichte, d.h. $\int_{\mathbb{R}} h(x) dx = 1$, da Wahrscheinlichkeitsdichten normiert sind.

Die Funktion $h(x)$ ist nicht elementar integrierbar. Die Normiertheit sieht man so:

$$\int_{\mathbb{R}^2} e^{-\|(x,y)\|^2_2} dx dy = 2\pi \int_0^\infty r e^{-r^2} dr = 2\pi \left[\frac{e^{-r^2}}{2} \right]_{r=0}^\infty = -\pi(0 - 1) = \pi.$$

und

$$\int_{\mathbb{R}^2} e^{-\|(x,y)\|^2_2} dx dy = \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-x^2-y^2} d(x, y) = \int_{\mathbb{R}} e^{-x^2} dx \int_{\mathbb{R}} e^{-y^2} dy = \left(\int_{\mathbb{R}} e^{-x^2} dx \right)^2$$

und somit $\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi}$.

6.3 Partielle Integration

Dies ist die Verallgemeinerung der bekannten eindimensionalen Regel auf d Dimensionen. Dies ist leicht, solange wir mit Funktionen mit kompaktem Träger arbeiten; Randterme werden später mit dem GAUSSSchen Integralsatz behandelt.

Notation.

Für eine offene Teilmenge $U \subset \mathbb{R}^d$ sei $\mathcal{C}^n(U)$ die Menge der n -mal stetig differenzierbaren Funktionen von U nach \mathbb{R} für $n \in \mathbb{N}_0 \cup \{\infty\}$. Es gilt $\mathcal{C}^0(U) = \mathcal{C}(U)$. Das Subskript c bedeutet jeweils mit kompakten Träger in U (vgl. oben).

Für offenen Mengen $U \subset V \subset \mathbb{R}^d$ sehen wir $\mathcal{C}_c^n(U)$ (durch Fortsetzung mit Null) als Teilmenge von $\mathcal{C}_c^n(V)$.

Glatte Teilung der Eins

Wir definieren

$$g : \mathbb{R} \rightarrow \mathbb{R}, t \mapsto \begin{cases} \exp\left(\frac{1}{t^2-1}\right), & \text{für } |t| < 1, \\ 0, & \text{sonst.} \end{cases}$$

Dann ist $g \in \mathcal{C}_c^\infty(\mathbb{R})$ ¹⁵ mit $\text{supp}(g) = [-1, 1]$.

Nun definieren wir

$$G : \mathbb{R} \rightarrow (0, \infty), t \mapsto \sum_{k \in \mathbb{Z}} g(t - k) \quad \text{und} \quad h(t) := \frac{g(t)}{G(t)}.$$

Dann ist auch $h \in \mathcal{C}_c^\infty(\mathbb{R})$ mit $\text{supp}(h) = \text{supp}(g)$. Ferner gilt

$$\sum_{k \in \mathbb{Z}} h(t - k) \equiv 1.$$

Somit ist $(\theta_k h)_{k \in \mathbb{Z}}$ eine glatte Teilung der Eins.

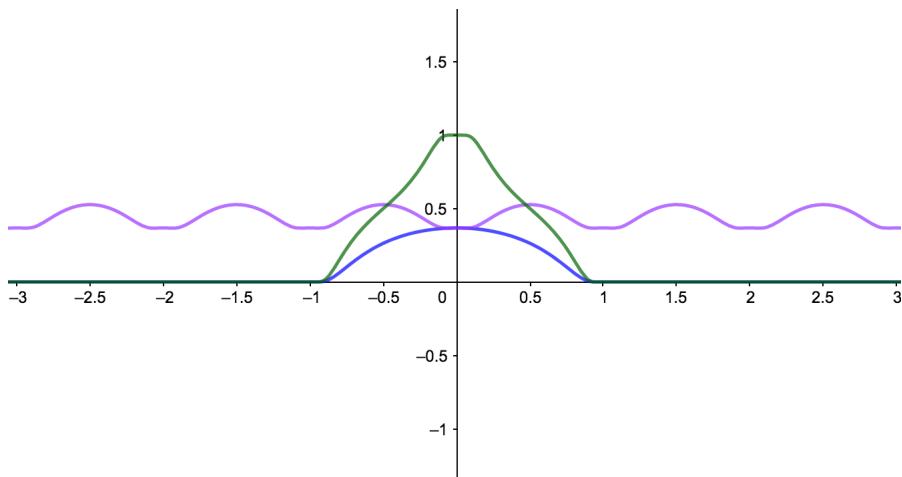


Abb. 35: g (blau), G (lila) und h (grün) für $k \in [-33, 33]$ anstatt ganz \mathbb{Z} . [Quelle: GeoGebra6]

¹⁵ HA, sieht man genauso wie in dem Beweis, dass die TAYLOR-Reihe der Abbildung $x \mapsto \exp(-x^{-2})$ in Null konstant gleich Null ist.

Setze nun für $x \in \mathbb{R}^d$, $k \in \mathbb{Z}^d$, $\varepsilon > 0$ mit $h_\varepsilon(x) := h\left(\frac{x}{\varepsilon}\right)$

$$\alpha_{k,\varepsilon}(x) := \prod_{i=1}^d h\left(\frac{x_i}{\varepsilon} - k_i \varepsilon\right) = \left[\bigotimes_{i=1}^d (\theta_{k_i \varepsilon} h)\right]\left(\frac{x}{\varepsilon}\right) = \theta_{k\varepsilon} h_\varepsilon^{\otimes d}(x).$$

Die Familie $(\alpha_{k,\varepsilon})_{k \in \mathbb{Z}}$ ist unsere d -dimensionale glatte Teilung der Eins.

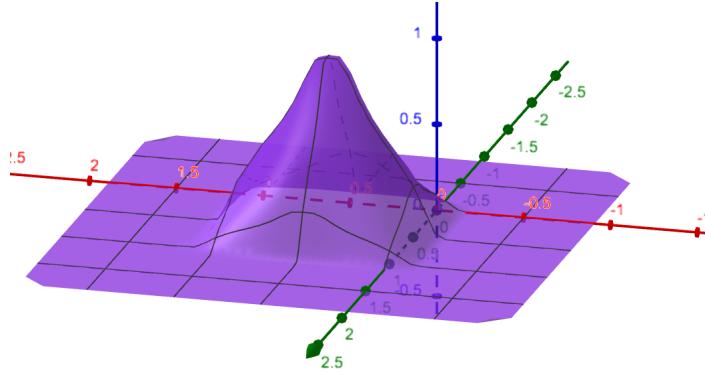


Abb. 36: $\alpha_{1,0.7}$ für $d = 2$. [Quelle: GeoGebra6]

Lemma 6.3.1

Seien $U \subset \mathbb{R}^d$ eine offene Teilmenge und $K \subset U$ eine kompakte Teilmenge und $n \in \mathbb{N} \cup \{\infty\}$. Dann existiert zu jeder Funktion $f \in C^n(U)$ eine Funktion $\beta \in C_c^n(U)$ mit $\beta|_K = f|_K$.

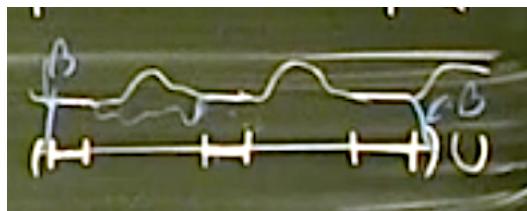


Abb. 37: [Quelle: Vorlesung]

Beweis. Sei o.B.d.A $f|_K \equiv 1$. Wähle nun $\varepsilon > 0$ so klein, dass jeder Würfel der Seitenlänge 2ε , der K trifft, in U enthalten ist.

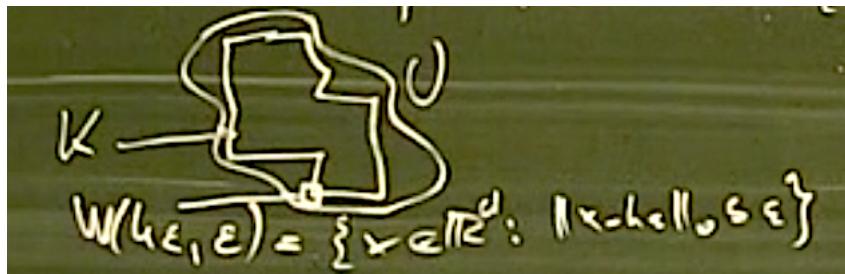


Abb. 38: [Quelle: Vorlesung]

Setze nun

$$\mathcal{P} := \{k \in \mathbb{Z}^d : \underbrace{\text{supp}(\alpha_{k,\varepsilon})}_{=\mathcal{W}(k\varepsilon, \varepsilon)} \cap K \neq \emptyset\}$$

als die Menge der $k \in \mathbb{Z}^d$, sodass die $\mathcal{W}(k\varepsilon, \varepsilon)$ -WürfelvK treffen. Dann erfüllt die Abbildung

$$\beta := \sum_{k \in \mathcal{P}} \alpha_{k, \varepsilon}$$

das Gewünschte denn: da \mathcal{P} endlich ist, summiert β endlich viele Funktionen mit kompakten Träger in U und somit hat auch β kompakten Träger in U . Dass $\beta|_K \equiv 1$ ist, sieht man, indem man nur über $k \in \mathcal{P}$ mit $x \in \mathcal{W}(k\varepsilon, \varepsilon)$ summiert. \square

Bemerkung 6.3.2 Nun kann man auch einen Beweis für die Existenz der in Definition 6.2.13 benutzten Folgen geben: Wir können die offene Menge U mit monoton wachsenden ineinander enthaltenen Kompakta $K_n \subset U$ ausschöpfen, für welche $K_n \subset K_{n+1}$ gilt und K_{n+1}° immer noch eine offene Umgebung von K_n ist. Dann wende man auf K_n und K_{n+1}° dieses Lemma an und erhält eine Funktion, die auf K_n das Gewünschte erfüllt und kleiner ist als die Funktion für K_{n+1} .

Lemma 6.3.3

Sei $U \subset \mathbb{R}^d$ eine offene Teilmenge, und $h \in \mathcal{C}(U)$ so gewählt, dass

$$\int_U h(x)g(x)dx = 0$$

für alle $g \in \mathcal{C}_c^\infty(U)$ gilt Dann gilt $h|_U \equiv 0$.

Bemerkung 6.3.4 (Lineare-Algebra-Sprechweise) Die obige Bedingung ist erfüllt, wenn die Vektoren g und h des Funktionenvektorraums $\mathcal{C}_c^\infty(U)$ bezüglich des Skalarprodukts $\langle f, g \rangle := \int_U f(x)g(x)dx$ senkrecht zueinander sind.

Beweis. Wir nehmen an, es existiert ein $a \in U$ mit $h(a) \neq 0$. Aufgrund der Stetigkeit von h muss es nun andere Punkte in einer Umgebung von a geben, auf denen h auch ungleich Null ist:

Sei o.B.d.A $h(a) > 0$. Dann existiert ein $\delta > 0$ und ein $\varepsilon > 0$ mit $\mathcal{W}(a, \varepsilon) \subset U$ und $h|_{\mathcal{W}(a, \varepsilon)} \geq \delta$.

Mit Hilfe der Funktion g vom Anfang des Kapitels 2.3 finden man ein $g \in \mathcal{C}_c^\infty(U)$ mit $\text{supp}(g) \subset \mathcal{W}(a, \varepsilon)$ und $g|_{\mathcal{W}(a, \frac{\varepsilon}{2})^\circ} > \frac{1}{2}$. Dann gilt

$$\begin{aligned} \int_U h(x)g(x)dx &= \int_{\mathcal{W}(a, \varepsilon)} h(x)g(x)dx \geq \delta \int_{\mathcal{W}(a, \frac{\varepsilon}{2})} g(x)dx \\ &\geq \frac{\delta}{2} \text{Vol}\left(\mathcal{W}\left(a, \frac{\varepsilon}{2}\right)\right) = \frac{\delta}{2} \cdot \varepsilon^d > 0. \end{aligned} \quad \square$$

Partielle Integration im \mathbb{R}^d

Erinnerung zur Notation

$$D_i := \frac{\partial}{\partial x_i} \text{ und } D_i^2 := \frac{\partial^2}{\partial x_i^2}.$$

Lemma 6.3.5 (partielle Integration)

Seien $U \subset \mathbb{R}^d$ eine offene Teilmenge und $i \in \{1, \dots, d\}$. Dann gelten

(1) $\int_U D_i \varphi(x)dx = 0$ für alle $\varphi \in \mathcal{C}_c^1(U)$ und

(2) $\int_U D_i f(x)g(x)dx = - \int_U f(x)D_i g(x)dx$ für alle $f, g \in \mathcal{C}_c^1(U)$.

Beweis. Die Aussage ② folgt aus der Aussage ① mit $\varphi := f \cdot g$ und ‘ausmultiplizieren’ von $D_i f(x)g(x)$ mit der Produktregel.

Für Aussage ① wählt man $R > 0$ so groß, dass $\text{supp}(\varphi) \subset [-R, R]^d$ ist. Dann kann man das d -dimensionale Integral in d Integrale über das kompakte Intervall $[-R, R]$ aufteilen und weiß, dass $\varphi(R) = \varphi(-R) = 0$ ist. \square

Beispiel 6.3.6 (LAPLACE-Operator) ¹⁶

Unter den Voraussetzungen des vorherigen Lemmas gilt für Funktionen $f \in \mathcal{C}^2(U)$ und $g \in \mathcal{C}_c^2(U)$.

$$\int_U D_i^2 f(x)g(x)dx = - \int_U D_i f(x)D_i g(x)dx = \int_U f(x)D_i^2 g(x)$$

Durch Summieren über alle $i \in \{1, \dots, d\}$ sieht man, dass der LAPLACE-Operator **selbstadjungiert** ist:

$$\int_U \Delta f(x)g(x)dx = \int_U f(x)\Delta g(x)dx.$$

Lineare Differentialoperatoren

Notation. (Multiindexschreibweise)

Für $\alpha := (\alpha_1, \dots, \alpha_d) \in \mathbb{N}_0^d$ seien

$$|\alpha| := \|\alpha\|_1 = \sum_{i=1}^d \alpha_i \quad \text{und} \quad D^\alpha := \prod_{i=1}^d D_i^{\alpha_i}.$$

DEFINITION 6.3.7

Für $m \in \mathbb{N}_{\geq n} \cup \{\infty\}$ und eine offene Teilmenge $U \subset \mathbb{R}^d$ heißt ein Operator

$$L : \mathcal{C}^m(U) \rightarrow \mathcal{C}^{m-n}(U), \quad f \mapsto \sum_{\substack{\alpha \in \mathbb{N}_0^d \\ |\alpha| \leq n}} a_\alpha D^\alpha f$$

mit $n \in \mathbb{N}_0$ und Koeffizientenfunktionen $a_\alpha \in \mathcal{C}^\infty(U)$ **linearer Differentialoperator** n -ter Ordnung.

Differentialoperator

07.11.2018

Beispiel 6.3.8 Jeder lineare Differentialoperator 0-ter Ordnung die Gestalt $L = a_0$, beschreibt also die Multiplikation mit einer Funktionen.

Für $n = 2$ ist ein spezieller Differential der LAPLACE-Operator

$$L := \Delta = \sum_{i=1}^d D_i^2 = \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}$$

Bemerkung 6.3.9 Der Raum der linearen Differentialoperatoren ist ein reeller Vektorraum über $\mathcal{C}^\infty(U)$.

¹⁶Definiert als die Summe der zweiten partielle Ableitungen: $\Delta := \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}$ ist für Stochastiker als Erzeuger der Brownschen Bewegung und für Analytiker für die Definition von Harmonizität von großer Bedeutung.

Lemma 6.3.10 (Komposition von Differentialoperatoren)

Seien $n_1, n_2, m \in \mathbb{N}$ mit $n_1 + n_2 \leq m$ und $U \subset \mathbb{R}^d$ eine offene Teilmenge. Ferner seien

$$L_1 : C^m(U) \rightarrow C^{m-n_1}(U) \quad \text{und} \quad L_2 : C^{m-n_1}(U) \rightarrow C^{m-(n_1+n_2)}(U)$$

zwei lineare Differentialoperatoren.

Dann ist

$$L_2 \circ L_1 : C^m(U) \rightarrow C^{m-(n_1+n_2)}(U)$$

ein linearer Differentialoperator der Ordnung $n_1 + n_2$.

Beweis. Durch vollständige Induktion über $|\alpha|$.

Da die linearen Differentialoperatoren einen Vektorraum bilden, genügt es, den Fall $L_1 := a_\alpha D^\alpha$ und $L_2 := D_i$ zu betrachten.

Induktionsschritt: $|\alpha| \rightarrow |\alpha| + 1$.

Für $f \in C^{|\alpha|+1}(U)$ gilt nach der Produktregel

$$(L_2 \circ L_1)(f) = L_2(L_1(f)) = D_i(a_\alpha D^\alpha)(f) = (D_i a_\alpha)(f) + a_\alpha D_i D^\alpha f,$$

welche ein linearer Differentialoperator der Ordnung $|\alpha| + 1$ ist. \square

Beispiel 6.3.11 (Die Komposition ist nicht kommutativ)

Seien $L_1 := D_i$ und $L_2 := x_j$.

Wäre die Komposition $L_1 \circ L_2$ kommutativ, müsste der **Kommutator** $[M, N]_{\text{Kommutator}} = M \circ N - N \circ M$ verschwinden. Jedoch gilt

$$\begin{aligned} [L_1, L_2]f &= (L_1 \circ L_2)(f) - (L_2 \circ L_1)(f) = (D_i x_j)(f) - (x_j D_i)(f) \\ &= \delta_{i,j} f + x_j(D_i f) - x_j(D_i f) = \delta_{i,j} f. \end{aligned}$$

Bemerkung 6.3.12 (Übergang Mechanik → Quantenmechanik) L_2 ist hier der Ortsoperator, also die Position eines Teilchens und L_1 den Impuls eines Teilchens.

Nach der **HEISENBERGSCHEN UNSCHÄRFERELATION** können wir nicht gleichzeitig die Position und den Impuls eines quantenmechanischen Teilchens messen.

SATZ 6.3.1: EXISTENZ UND EINDEUTIGKEIT DES ADJUNGIERTEN OPERATORS

Seien $U \subset \mathbb{R}^d$ eine offene Teilmenge und L ein linearer Differentialoperator n -ter Ordnung in U . Dann existiert genau ein linearer Differentialoperator L^* der Ordnung n , sodass

$$\int_U (L^* f) g dx = \int_U f (L g) dx \quad (83)$$

für alle $f \in C^n(U)$ und alle $g \in C_c^n(U)$ gilt.

DEFINITION 6.3.13 (ADJUNGIERTER OPERATOR)

Man nennt den obigen Operator L^* den zu L adjungierten Operator.

adjungierten Operator

Lemma 6.3.14 (Rechenregeln für adjungierte Operatoren)

In der Situation des obigen Satzes gilt

$$\textcircled{1} \quad (\lambda L)^* = \lambda L^* \text{ für alle } \lambda \in \mathbb{R}.$$

$$\textcircled{2} \quad (L_1 + L_2)^* = L_1^* + L_2^*.$$

$$\textcircled{3} \quad (L_1 \circ L_2)^* = L_2^* \circ L_1^*.$$

Beweis. (Satz und Lemma) Eindeutigkeit. Seien M_1 und M_2 zwei lineare Differentialoperatoren n -ter Ordnung, welche der Gleichung (83) genügen. Dann gilt für alle $f \in C^n(U)$ und alle $g \in C_c^n(U)$ aufgrund der Linearität

$$\int_U (M_1 f - M_2 f) g dx = \int_U f (Lg - Lg) dx = 0.$$

Nach Lemma 6.3.3 gilt nun $M_1 f - M_2 f = 0$ und somit $M_1 = M_2$.

Lemma. Exemplarisch sei nur die dritte Rechenregel gezeigt. Für $f \in C^n(U)$ gilt

$$\int_U ((L_2^* \circ L_1^*)(f)) g dx = \int_U (L_1^* f)(L_2 g) dx = \int_U f L_1(L_2(g)) dx.$$

Aufgrund der Eindeutigkeit des adjungierten Operators folgt die Behauptung.

Existenz. Für $L = a \in C^\infty(U)$, einen Differentialoperator 0-ter Ordnung gilt

$$\int_U (Lg)(f) dx = \int_U (ag)(f) dx = \int_U g(af) dx = \int_U (L^* f) g dx$$

und somit $L^* = L$, also ist der Multiplikationsoperator selbstadjungiert.

Ferner gilt für $L = D_i$, einen Differentialoperator erster Ordnung $L^* = -D_i$ nach der Formel der partiellen Integration aus Lemma 6.3.5.

Da sich jeder lineare Differentialoperator durch Addition, skalare Multiplikation und Komposition zusammensetzt, folgt zusammen mit den Rechenregeln aus dem Lemma die Existenz des adjungierten Operators. \square

Beispiel 6.3.15 (Korollar zum obigen Lemma und Beweis) Für glatte Koeffizientenfunktionen $\alpha_1, \dots, \alpha_d \in C^\infty(U)$ und eine Funktion $f \in C^1(U)$ und eine offene Teilmenge $U \subset \mathbb{R}^d$ gilt mit der Produktregel

$$\left(\sum_{i=1}^d \alpha_i D_i \right)^* f \stackrel{\textcircled{2}}{=} \sum_{i=1}^d (\alpha_i D_i)^* f \stackrel{\textcircled{3}}{=} - \sum_{i=1}^d (D_i)(\alpha_i f) = - \left(\sum_{i=1}^d (D\alpha_i) f + \alpha_i (D_i f) \right).$$

Wir haben am Anfang den LAPLACE-Operator in euklidischen Koordinaten definiert, da wir nach den Koordinatenachsen abgeleitet haben.

Wollen wir aber z.B. den LAPLACE-Operator auf eine rotationssymmetrische Funktion anwenden, sparen wir viel Rechenarbeit, wenn ihr den LAPLACE-Operator in dem entsprechenden Koordinaten, also den Polarkoordinaten betrachten.

6.4 Der LAPLACE-Operator in krummlinigen Koordinaten

DEFINITION 6.4.1 (LAPLACE-OPERATOR IN KRUMMLINIGEN KOORDINATEN)

Seien $\Omega_1, \Omega_2 \subset \mathbb{R}^d$ offene Teilmengen und $\Phi = (\Phi_1, \dots, \Phi_d) : \Omega_2 \rightarrow \Omega_1$ eine C^2 -invertierbare Funktion.

Dann heißt für eine $C^2(\Omega_2)$ -Funktion u der Operator

$$\Delta^\Phi u := (\Delta(u \circ \Phi^{-1})) \circ \Phi$$

der LAPLACE-Operator in **krummlinigen Koordinaten**.

Wir wollen nun eine explizite Darstellung in krummlinigen Koordinaten finden.

Erinnerung zur Notation

In der Situation der obigen Definition sei $D\Phi := \left(\frac{\partial \Phi_i}{\partial x_j} \right)_{i,j \in \{1, \dots, d\}}$ die Differenzialmatrix von Φ .

Bezeichnungen Für Zeilenvektoren $a, b \in \mathbb{R}^d$ sei das Skalarprodukt $\langle a, b \rangle := ab^T$ und

$$G(y) := (D\Phi(y))^T (D\Phi(y)) = \sum_{k=1}^d \left(\frac{\partial \Phi_k}{\partial x_i} \frac{\partial \Phi_k}{\partial x_j} \right)_{i,j \in \{1, \dots, d\}}.$$

Bemerkung 6.4.2 Mit diesen Bezeichnungen lässt sich der [Transformationssatz](#) für eine Funktion $u \in C_c(\Omega_1)$ folgendermaßen darstellen:

$$\int_{\Omega_2} u(\Phi(y)) \sqrt{\det(G(y))} dy = \int_{\Omega_1} u(x) dx, \quad (84)$$

da $\sqrt{\det(G(y))} = |\det(\Phi(y))|$ gilt.

Lemma 6.4.3

Für alle Funktionen $u, v \in C^1(\Omega_1)$ gilt

$$\langle \nabla u, \nabla v \rangle \circ \Phi = \langle \nabla(u \circ \Phi) G^{-1}, \nabla(v \circ \Phi) \rangle.$$

Beweis. Aus der **Kettenregel** folgt

$$\nabla(u \circ \Phi) = ((\nabla(u)) \circ \Phi) D\Phi,$$

wobei der Gradient von u , ∇u , als Zeilenvektor aufgefasst wird.

Dann gilt

$$\begin{aligned} \langle \nabla(u \circ \Phi) G^{-1}, \nabla(v \circ \Phi) \rangle &= \langle ((\nabla(u)) \circ \Phi) \underline{D\Phi(D\Phi)^{-T}} (D\Phi)^{-T}, ((\nabla(v)) \circ \Phi) D\Phi \rangle \\ &= ((\nabla(u)) \circ \Phi) (D\Phi)^{-T} (((\nabla(v)) \circ \Phi) D\Phi)^T \\ &= ((\nabla(u)) \circ \Phi) \underline{(D\Phi)^{-T} (D\Phi)^T} ((\nabla(v)) \circ \Phi)^T \\ &= \langle (\nabla u) \circ \Phi, (\nabla v) \circ \Phi \rangle. \end{aligned} \quad \square$$

SATZ 6.4.1: LAPLACE-OPERATOR IN KRUMMLINIGEN KOORDINATEN

Es gilt

$$\begin{aligned}\Delta^\Phi &= \frac{1}{\sqrt{\det(G)}} \sum_{k,\ell=1}^d \frac{\partial}{\partial y_k} \left(G_{k,\ell}^{-1} \sqrt{\det(G)} \frac{\partial}{\partial y_\ell} \right) \\ &= \sum_{k,\ell=1}^d G_{k,\ell}^{-1} \frac{\partial^2}{\partial y_k \partial y_\ell} + \frac{1}{\sqrt{\det(G)}} \sum_{k,\ell=1}^d \frac{\partial (G_{k,\ell}^{-1} \sqrt{\det(G)})}{\partial y_\ell} \frac{\partial}{\partial y_k}.\end{aligned}$$

Beweis. Seien $u \in \mathcal{C}^2(\Omega_1)$ und $v \in \mathcal{C}_c(\Omega_1)$ sowie $\tilde{u} := u \circ \Phi$ und \tilde{v} entsprechend definiert. Dann folgt aus dem Transformationssatz (84)

$$\begin{aligned}\int_{\Omega_1} (\Delta u) v dx &= \int_{\Omega_2} ((\Delta u) \circ \Phi) \circ v \circ \Phi \cdot \sqrt{\det(G(y))} dy \\ &= \int_{\Omega_2} ((\Delta(\tilde{u} \circ \Phi^{-1})) \circ \Phi) \circ \tilde{v} \cdot \sqrt{\det(G(y))} dy \\ &= \int_{\Omega_2} \Delta^\Phi \tilde{u} \circ \tilde{v} \cdot \sqrt{\det(G(y))} dy.\end{aligned}$$

Andererseits gilt nach der Formel der partiellen Integration (\star)

$$\begin{aligned}\int_{\Omega_1} (\Delta u) v dx &\stackrel{(\star)}{=} - \int_{\Omega_1} \langle \nabla u, \nabla v \rangle dx \\ &\stackrel{(84)}{=} - \int_{\Omega_2} \langle \nabla u, \nabla v \rangle \circ \Phi \cdot \sqrt{\det(G)} dy \\ &\stackrel{\text{L 6.4.3}}{=} - \int_{\Omega_2} \langle \nabla(u \circ \Phi) G^{-1}, \nabla(v \circ \Phi) \rangle \cdot \sqrt{\det(G)} dy \\ &= - \int_{\Omega_2} \sum_{k,\ell=1}^d G_{k,\ell}^{-1} \cdot \sqrt{\det(G)} \frac{\partial \tilde{u}}{\partial y_k} \frac{\partial \tilde{v}}{\partial y_\ell} dy \\ &\stackrel{(\star)}{=} \int_{\Omega_2} \sum_{k=1}^d d \frac{\partial}{\partial y_\ell} G_{k,\ell}^{-1} \cdot \sqrt{\det(G)} \frac{\partial \tilde{u}}{\partial y_k} \tilde{v} dy.\end{aligned}$$

Führen wir beide Schritte zusammen, sehen wir

$$\int_{\Omega_2} \sum_{k=1}^d d \frac{\partial}{\partial y_\ell} G_{k,\ell}^{-1} \cdot \sqrt{\det(G)} \frac{\partial \tilde{u}}{\partial y_k} \tilde{v} dy = \int_{\Omega_2} \Delta^\Phi \tilde{u} \circ \tilde{v} \cdot \sqrt{\det(G(y))} dy.$$

Da $\tilde{v} \in \mathcal{C}_c^2(\Omega)$ beliebig ist, folgt aus Lemma 6.3.3

$$\Delta^\Phi \tilde{u} = \sum_{k=1}^d d \frac{\partial}{\partial y_\ell} G_{k,\ell}^{-1} \cdot \sqrt{\det(G)} \frac{\partial \tilde{u}}{\partial y_k}.$$

Da $\tilde{u} \in \mathcal{C}^2(\Omega)$ auch beliebig war, folgt die erste Gleichung, die zweite folgt direkt durch Anwenden der Produktregel. \square

12.11.2018

Bemerkung 6.4.4 Der Beweis geht nur, weil der LAPLACE-Operator selbstadjungiert ist. (WARUM???)

DEFINITION 6.4.5 (HARMONISCHE FUNKTION)

Eine Funktion u heißt **harmonisch**, wenn $\Delta^\Phi u = 0$ gilt.

Beispiel 6.4.6 (Der LAPLACE-Operator in den Polarkoordinaten)

Mithilfe des Beispiels 6.2.18 können wir für $\Omega_1 := \mathbb{R}^2 \setminus S$ und $\Omega_2 := (0, \infty) \times (0, 2\pi)$ und $\Phi : \Omega_2 \rightarrow \Omega_1, (r, y) \mapsto (r \cos(\varphi), r \sin(\varphi))$.

$$G(r, \varphi) = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}, \quad G^{-1}(r, \varphi) = \begin{pmatrix} 1 & 0 \\ 0 & r^{-2} \end{pmatrix}, \quad \sqrt{\det(G(r, \varphi))} = \sqrt{r^2} = r$$

errechnen.

Durch Einsetzen erhält man mit der zweiten Darstellung und $y_1 := r$ und $y_2 := \varphi$

$$\Delta^\Phi = \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r} \frac{\partial}{\partial \varphi} \left(\frac{1}{r} \frac{\partial}{\partial r} \right) = \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \varphi^2}.$$

Beispiel 6.4.7 (Skript) Hier fehlt noch was

Beispiel 6.4.8 (Der LAPLACE-Operator in den Polarkoordinaten)

$$\begin{aligned} \Delta^\Phi &= \frac{1}{r^2 \sin(\theta)} \left[\frac{\partial}{\partial r} \left(r^2 \sin(\theta) \frac{\partial}{\partial r} \right) + \frac{d}{d\theta} \left(\sin(\theta) \frac{\partial}{\partial \theta} \right) + \frac{\partial}{\partial \varphi} \left(\frac{1}{\sin(\theta)} \frac{\partial}{\partial \varphi} \right) \right] \\ &= \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \left(\frac{\partial^2}{\partial \theta^2} + \cot(\theta) \frac{\partial}{\partial \theta} + \frac{1}{\sin^2(\theta)} \frac{\partial^2}{\partial \varphi^2} \right). \end{aligned}$$

Viel ausführlicher!

6.5 Integrale halbstetiger Funktionen

Ziel

Wir wollen für eine Kompaktum $K \subset \mathbb{R}^d$ und eine stetige Funktion $f : K \rightarrow \mathbb{R}$ das Integral

$$\int_K f(x) dx$$

erklären. Am Besten wäre es, wenn wir

$$\int_K f(x) dx := \int_{\mathbb{R}^d} \tilde{f}(x) dx$$

triviale Fortsetzung definieren könnten, wobei \tilde{f} die triviale Fortsetzung der Funktion f ist.

Ist jedoch $f|_{\delta K} \neq 0$, so können wir diese Integrale aufgrund der mangelnden Stetigkeit der Fortsetzung nicht berechnen.

Indikatorfunktion Wir wollen für Teilmengen $A \subset \mathbb{R}^d$ die Integrale über die Indikatorfunktion

$$\mathbb{1}_A : \mathbb{R}^d \rightarrow \mathbb{R}, \quad x \mapsto \begin{cases} 1, & \text{wenn } x \in A, \\ 0, & \text{sonst.} \end{cases}$$

berechnen, weil

$$\text{Vol}(A) = \int_{\mathbb{R}^d} \mathbb{1}_A(x) dx$$

intuitiv gilt.

Idee

Wir approximieren die Indikatorfunktion mit unserer Teilung der Eins mittels Funktionen in $\mathcal{C}_c(\mathbb{R}^d)$, welche gleichmäßig gegen sie konvergieren.

Allgemeiner wollen wir sogar die triviale Fortsetzung auf diese Weise approximieren.

Es genügt, eine Funktionenfolge zu finden, welche *monoton* gegen die Fortsetzung konvergiert. Das folgt aus dem

SATZ 6.5.1: VON DINI

Seien $K \subset \mathbb{R}^d$ eine kompakte Teilmenge und $f, f_0, f_1, \dots : K \rightarrow \mathbb{R}$ stetige Funktionen mit

$$\textcircled{1} \quad f_n \leq f_{n+1} \quad \forall n \in \mathbb{N}_0 \quad \text{und} \quad \textcircled{2} \quad f_n \xrightarrow{n \rightarrow \infty} f. \quad (85)$$

Dann konvergiert die Funktionenfolge $(f_n)_{n \in \mathbb{N}_0}$ gleichmäßig gegen die Funktion f .

Beweis. Sei $\varepsilon > 0$ und $U_n := \{x \in K : f(x) - f_n(x) < \varepsilon\}$ für $n \in \mathbb{N}$. Da die Funktionenfolge $(f_n)_{n \in \mathbb{N}}$ punktweise konvergiert, existiert zu jedem $x \in K$ ein $k \in \mathbb{N}$, sodass $f(x) - f_k(x) < \varepsilon$ ist, also ist die Familie $(U_n)_{n \in \mathbb{N}}$ eine Überdeckung von K , welche auch offen ist, da jedes U_k das Urbild des in \mathbb{R} offenen Intervalls $(-\infty, \varepsilon)$ unter der als Differenz stetiger Abbildungen stetiger Abbildung $f - f_k$ ist. Ferner

gilt, weil die Funktionenfolge $(f_n)_{n \in \mathbb{N}}$ monoton konvergiert (o.B.d.A wächst, sonst benutzt man die Familie $(-f_n)_{n \in \mathbb{N}}$), gilt $U_\ell \subset U_{\ell+1}$ für alle $\ell \in \mathbb{N}$.

Weil K kompakt ist, wird es auch von einer endlichen Teilüberdeckung $(U_k)_{k \in J}$ überdeckt. Sei nun $N := \max(J)$, also gilt $U_i = K$ für alle $i > N$. Somit gilt für alle $i > N$ und $x \in K$

$$|f(x) - f_i(x)| = f(x) - f_i(x) < \varepsilon$$

und somit folgt

$$\limsup_{i \rightarrow \infty} \sup_{x \in K} |f_i(x) - f(x)| = 0,$$

was zu zeigen war. \square

Notation

Im Folgenden schreiben wir für diese Art der Konvergenz (85) $f_n \uparrow f$. Existiert für eine Funktion f eine Folge $(f_n)_{n \in \mathbb{N}_0} \subset \mathcal{C}_c(\mathbb{R}^d)$ mit $f_n \uparrow f$, so sagen wir $f \in \mathcal{H}^\uparrow(\mathbb{R}^d)$ und analog $f_n \downarrow f$ und $\mathcal{H}^\downarrow(\mathbb{R}^d)$.

Korollar 6.5.1

Seien $f, f_0, f_1, \dots \in \mathcal{C}_c(\mathbb{R}^d)$ kompakt getragene Funktionen mit $f_n \uparrow f$. Dann gilt

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} f_n(x) dx = \int_{\mathbb{R}^d} f(x) dx.$$

Beweis. Aufgrund der Monotonie der Funktionenfolge $(f_n)_{n \in \mathbb{N}_0}$ gilt für alle $k \in \mathbb{N}_0$

$$\text{supp}(f_k) \subset \text{supp}(f_0) \cup \text{supp}(f) =: \mathcal{K}$$

als abgeschlossene Teilmenge einer endlichen Vereinigung kompakter Mengen kompakt.

Aus der Anwendung des [Satzes von DINI](#) auf die stetige Funktionenfamilie und das Kompaktum \mathcal{K} folgt nun die *gleichmäßige* Konvergenz $f_n \rightarrow f$.

Aus dem Lemma 6.1.6 und der [Axiomatischen Charakterisierung des RIEMANN-Integrals](#) angewandt auf das Kompaktum K folgt genau die Behauptung. \square

Idee

Für jede Funktion f , für welche eine kompakt getragene Funktionenfolge $(f_n)_{n \in \mathbb{N}_0}$ mit $f_n \uparrow f$ (oder $f_n \downarrow f$) existiert, definieren wir

$$\int_{\mathbb{R}^d} f(x) dx := \lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} f_n(x) dx.$$

Achtung

Hierbei muss f nicht in $\mathcal{C}_c(\mathbb{R}^d)$ sein und auch das Integral von Funktionen der Form $f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{\pm\infty\}$ ist so wohldefiniert.

Beispiel 6.5.2 Wollen wir z.B. das Integral $\int_{\mathbb{R}^d} \mathbf{1}_{\mathbb{R}^d}(x)dx$ berechnen, verwenden wir unsere Teilung der Eins $(h_n)_{n \in \mathbb{N}_0} \subset \mathcal{C}_c(\mathbb{R}^d)$ und definieren $f_n := \sum_{k=0}^n h_n$. Dann gilt $f_n \uparrow \mathbf{1}_{\mathbb{R}^d}$.

Wir können die h_n nun so konstruieren, dass für jedes $n \in \mathbb{N}_0$ gilt

$$n^d \leq \int_{\mathbb{R}^d} f_n(x)dx,$$

indem wir f_n auf dem entsprechenden Würfel mit Volumen n^d konstant eins seien und danach abfallen lassen.

Dann gilt

$$\infty = \lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} f_n(x)dx \stackrel{\text{Def.}}{=} \int_{\mathbb{R}^d} \mathbf{1}_{\mathbb{R}^d}(x)dx.$$

Lemma 6.5.3 (Unabhängigkeit des Integrals von der Folge)

Seien $f \in \mathcal{H}^\uparrow(\mathbb{R}^d)$ und $(f_n)_{n \in \mathbb{N}}, (g_n)_{n \in \mathbb{N}} \subset \mathcal{C}_c(\mathbb{R}^d)$ kompakt getragene Funktionenfolge mit $f_n \uparrow f$ und $g_n \uparrow f$. Dann gilt

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} f_n(x)dx = \lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} g_n(x)dx.$$

Beweis. Aus Symmetriegründen genügt es zu zeigen, dass für $k \in \mathbb{N}_0$ die Abschätzung

$$\int_{\mathbb{R}^d} f_k(x)dx \leq \lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} g_n(x)dx$$

gilt. Dazu betrachten wir die folgende Hilfsfunktion

$$h_n := \min(f_k, g_n) \in \mathcal{C}_c(\mathbb{R}^d).$$

Dann gilt $h_n \uparrow f_k$.

Nach dem Korollar 6.5.1 gilt nun, weil $f_k, h_n \in \mathcal{C}_c(\mathbb{R}^d)$ sind und das Integral monoton ist und der Grenzwert schwache Ungleichungen erhält,

$$\int_{\mathbb{R}^d} f_k(x)dx = \lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} h_n(x)dx \leq \lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} g_n(x)dx. \quad \square$$



Großes Ziel dieser Kapitels

Eine intrinsische Charakterisierung der Mengen \mathcal{H}^\uparrow und \mathcal{H}^\downarrow finden.

Insbesondere zeigen wir $\mathcal{H}^\downarrow \cap \mathcal{H}^\uparrow = \mathcal{C}_c$, d.h., dass die Integral bezüglich monoton wachsender bzw. fallender Approximationsfolgen stimmen Korollar 6.5.1 überein, sofern beide existieren.

Lemma 6.5.4

Sind $\mathcal{K} \subset \mathbb{R}^d$ eine kompakte bzw. $U \subset \mathbb{R}^d$ eine offene Teilmenge und $f : \mathcal{K}$ bzw. $U \rightarrow [0, \infty)$ eine stetige Funktion, so gilt $\tilde{f} \in \mathcal{H}^\downarrow(\mathbb{R}^d)$ bzw. $\mathcal{H}^\uparrow(\mathbb{R}^d)$, wobei wir die triviale Fortsetzung einer Funktion f mit Null mit \tilde{f} bezeichnen.

Beweis. Hausaufgabe. \square

Mit dem obigen Lemma wird folgende Definition sinnvoll:

DEFINITION 6.5.5

$$\int_{\mathcal{K} \text{ bzw. } U} f(x) dx := \int_{\mathbb{R}^d} f(x) dx.$$

DEFINITION 6.5.6 (HALBSTETIGE FUNKTIONEN)

- ① Eine Funktion $f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{\infty\}$ heißt in einem Punkt $x \in \mathbb{R}^d$ **stetig von unten** oder **unterhalbstetig**, wenn gilt unterhalbstetig

$$\forall c \in \mathbb{R} \text{ mit } c < f(x) \exists \delta > 0 : c < f(y) \forall y \in B_\delta^{\|\cdot\|}(x).$$

Ist f unterhalbstetig in jedem $x \in \mathbb{R}^d$, so nennen wir f unterhalbstetig.

- ② Eine Funktion $f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{-\infty\}$ heißt in einem Punkt $x \in \mathbb{R}^d$ **stetig von oben** oder **oberhalbstetig**, wenn die Funktion $-f$ unterhalbstetig in **halbstetig** x ist. Ist f oberhalbstetig in jedem $x \in \mathbb{R}^d$,

Lemma 6.5.7 (Charakterisierungen von Halbstetigkeit)

- ① Die Indikatorfunktion $\mathbb{1}_A$ einer Menge $A \subset \mathbb{R}^d$ ist genau dann unter- bzw. oberhalbstetig, wenn A offen bzw. abgeschlossen ist.
- ② Eine Funktion $f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{\infty\}$ ist genau dann unterhalbstetig, wenn
- ① $\liminf_{y \rightarrow x} f(y) \geq f(x)$ ist.
 - ② $\liminf_{\varepsilon \searrow 0} \inf_{y \in B_\varepsilon(x)} f(y) = f(x)$ ist.
 - ③ die Niveaumenge $\{f \leq s\} := \{x \in \mathbb{R}^d : f(x) \leq s\}$ für jedes $s \in \mathbb{R}$ abgeschlossen sind.
- ③ Jede oberhalbstetige Funktion $f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{-\infty\}$ nimmt auf jedem Kom- paktum $\mathcal{K} \subset \mathbb{R}^d$ ein Maximum an.

Beweis. Hausaufgabe. □

Lemma 6.5.8 (Charakterisierung von \mathcal{H}^\uparrow)

Es gilt

$$\mathcal{H}^\uparrow(\mathbb{R}^d) = \{f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{\infty\} : f \text{ unterhalbstetig und } \{f < 0\} \text{ ist beschränkt.}\}.$$

Beweis. " \subset ": In der großen Übung werden wir sehen, dass das Supremum über eine Familie unterhalbstetiger Funktionen unterhalbstetig ist.

Da $\mathcal{C}_c(\mathbb{R}^d) \subset \mathcal{H}^\uparrow(\mathbb{R}^d)$ ist, ist

$$f(x) := \sup_{n \in \mathbb{N}_0} f_n(x)$$

mit $f_n \uparrow f$ unterhalb stetig.

Da $f_0 \in \mathcal{C}_c(\mathbb{R}^d)$ ist, kann f aufgrund der Monotonie maximal auf $\text{supp}(f_0)$ negativ seien und somit ist $\{f < 0\} \subset \text{supp}(f_0)$ beschränkt.

13.11.2018: Sei $f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{\infty\}$ unterhalb stetig und $\{f < 0\}$ beschränkt.

Wir zeigen nun, dass eine Approximationsfolge $(f_n)_{n \in \mathbb{N}} \subset \mathcal{C}_c(\mathbb{R}^d)$ existiert, sodass $f_n \uparrow f$ gilt, also $f \in \mathcal{H}^\uparrow(\mathbb{R}^d)$.

Zunächst ist $K := \overline{\{f < 0\}}$ kompakt, und, da die Funktion f unterhalbstetig ist, existiert ein $M > 0$, sodass $f(x) > -M$ für alle $x \in \mathbb{R}^d$ gilt.

Das sieht man so: [Aus dem Forster abschreiben]

Sei

$$J := \left\{ (\varepsilon, a, c) \in ((0, \infty) \times \mathbb{R}^d \times [-M, \infty)) \cap \mathbb{Q}^{d+2} : f_{B_\varepsilon(a)}(x) \geq c \right\}.$$

Dann ist J als Teilmenge von \mathbb{Q}^{d+2} abzählbar.

Zu jedem $j \in J$ existiert eine Funktion $g_j \in \mathcal{C}_c(\mathbb{R}^d)$ mit den folgenden Eigenschaften:

$$g_j(x) = \begin{cases} c, & \text{wenn } x \in B_{\frac{\varepsilon}{2}}(a), \\ \leq c, & \text{wenn } x \in B_\varepsilon(a), \\ = -M, & \text{wenn } x \in K \setminus B_\varepsilon(a), \\ \leq 0, & \text{wenn } x \in \mathbb{R}^d \setminus (K \cup B_\varepsilon(a)). \end{cases}$$

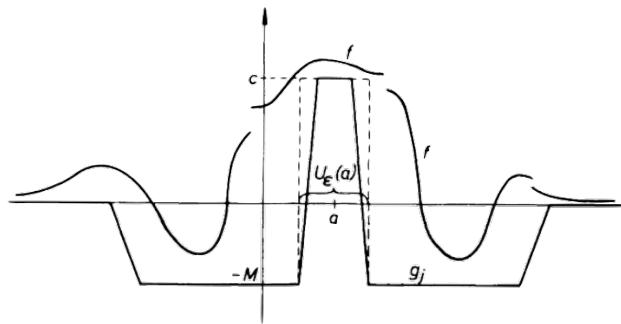


Abb. 39: [Quelle: Forster]

Dann gilt $f \geq g_j$ für alle $j \in J$.

Da zu jedem $x \in \mathbb{R}^d$ eine Folge $(\varepsilon_n, a_n, c_n)_{n \in \mathbb{N}} \subset J$ mit $\varepsilon_n \searrow 0$, $a_n \rightarrow x$ und $c_n \nearrow f(x)$ existiert, folgt

$$f = \sup_{j \in J} g_j.$$

Da die Menge J abzählbar ist, existiert eine Abzählung, sodass $J = \{j_0, j_1, \dots\}$ ist.

Setze nun für jedes $n \in \mathbb{N}_0$

$$f_n := \max(g_{j_0}, \dots, g_{j_n}) \in \mathcal{C}_c(\mathbb{R}^d).$$

Dann gilt $f_n \rightarrow f$, da das Supremum über immer mehr Funktionen gebildet wird. Also folgt $f \in \mathcal{H}^\uparrow(\mathbb{R}^d)$. \square

Korollar 6.5.9 ($\mathcal{H}^\uparrow \cap \mathcal{H}^\downarrow = \mathcal{C}_c$)

Es gilt $\mathcal{H}^\uparrow(\mathbb{R}^d) \cap \mathcal{H}^\downarrow(\mathbb{R}^d) = \mathcal{C}_c(\mathbb{R}^d)$.

Beweis. Für eine Funktion $f : \mathbb{R}^d \rightarrow R \cup \{-\infty\}$ gilt

$$f \in \mathcal{H}^\downarrow(\mathbb{R}^d) \iff -f \in \mathcal{H}^\uparrow(\mathbb{R}^d).$$

(WARUM???) Daraus folgt

$$\mathcal{H}^\uparrow(\mathbb{R}^d) = \{f : \mathbb{R}^d \rightarrow R \cup \{-\infty\} : f \text{ ist unterhalbstetig und } \{f > 0\} \text{ ist beschränkt.}\}$$

Aus dem Tutorium wissen wir, dass eine oberhalb und unterhalbstetige Funktion stetig ist. Ferner folgt daraus, dass $\{f < 0\}$ und $\{f > 0\}$ beschränkt sind, dass f kompakten Träger hat. \square

SATZ 6.5.2: LINEARITÄT DES INTEGRALS AUF $\mathcal{H}^\uparrow(\mathbb{R}^d)$

Für alle Funktionen $f \in \mathcal{H}^\uparrow(\mathbb{R}^d)$ und alle Skalare $\lambda \in [0, \infty)$ gilt $f + \lambda g \in \mathcal{H}^\uparrow(\mathbb{R}^d)$ und

$$\int_{\mathbb{R}^d} f(x) + \lambda g(x) dx = \int_{\mathbb{R}^d} f(x) dx + \lambda \int_{\mathbb{R}^d} g(x) dx$$

Beweis. Übungsaufgabe. (Durch zurückführen auf Approximationen) **TODO** \square

SATZ 6.5.3: TRAFOSATZ FÜR LINEARE TRANSFORMATIONEN

Seien $A \in \mathrm{GL}(d, \mathbb{R})$, $b \in \mathbb{R}^d$ und $f \in \mathcal{H}^\uparrow(\mathbb{R}^d) \cup \mathcal{H}^\downarrow(\mathbb{R}^d)$.

Dann ist die Funktion $x \mapsto f(Ax + b)$ ebenfalls in $\mathcal{H}^\uparrow(\mathbb{R}^d) \cup \mathcal{H}^\downarrow(\mathbb{R}^d)$ und es gilt

$$\int_{\mathbb{R}^d} f(Ax + b) dx = \frac{1}{|\det(A)|} \int_{\mathbb{R}^d} f(x) dx.$$

Beweis. Folgt durch Approximation aus den Sätzen 6.1.2 und 6.2.2. **TODO!** \square

SATZ 6.5.4: VON FUBINI FÜR FUNKTION AUS $\mathcal{H}^\uparrow(\mathbb{R}^d)$

Seien $f \in \mathcal{H}^\uparrow(\mathbb{R}^d)$ und $k \in \{1, \dots, d-1\}$. Dann liegt für jedes $(x_{k+1}, \dots, x_d) \in \mathbb{R}^{d-k}$ die Funktion $\mathbb{R}^k \ni (x_1, \dots, x_k) \mapsto (x_1, \dots, x_k, x_{k+1}, \dots, x_d) \in \mathbb{R}^d$ in $\mathcal{H}^\uparrow(\mathbb{R}^k)$.

Also ist das Integral

$$F(x_{k+1}, \dots, x_d) = \int_{\mathbb{R}^k} f(x) d(x_1, \dots, x_k), \quad x \in (x_1, \dots, x_d)$$

wohldefiniert.

Ferner gilt $F \in \mathcal{H}^\uparrow(\mathbb{R}^{d-k})$ und

$$\int_{\mathbb{R}^{d-k}} F(x_{k+1}, \dots, x_d) d(x_{k+1}, \dots, x_d) = \int_{\mathbb{R}^d} f(x) dx.$$

Beweis. Da $f \in \mathcal{H}^\uparrow(\mathbb{R}^d)$ ist, existiert eine Funktionenfolge $(f_n)_{n \in \infty} \subset \mathcal{C}_c(\mathbb{R}^d)$ mit $f_n \uparrow f$. Für ein festes $\xi := (x_{k+1}, \dots, x_d) \in \mathbb{R}^{d-k}$ definiere

$$f_n^\xi : \mathbb{R}^k \rightarrow \mathbb{R}, (x_1, \dots, x_k) \mapsto f_n^\xi(x_1, \dots, x_k) := f_n(x_1, \dots, x_k, x_{k+1}, \dots, x_d)$$

$$F_n^\xi : \mathbb{R}^k \rightarrow \mathbb{R} \cup \{\infty\}, (x_1, \dots, x_k) \mapsto f_n^\xi(x_1, \dots, x_k) := f(x_1, \dots, x_k, x_{k+1}, \dots, x_d).$$

Offensichtlich ist $(f_n^\xi) \subset \mathcal{C}_c(\mathbb{R}^k)$ und $f_n^\xi \uparrow f^\xi$ und somit gilt $f^\xi \in \mathcal{H}^\uparrow(\mathbb{R}^d)$.

Definiere nun

$$\begin{aligned} F(x_{k+1}, \dots, x_d) &:= \int_{\mathbb{R}^k} f^\xi(x_1, \dots, x_k) d(x_1, \dots, x_k) \\ &= \lim_{n \rightarrow \infty} \int_{\mathbb{R}^k} f_n^\xi(x_1, \dots, x_k) d(x_1, \dots, x_k) \\ &= \lim_{n \rightarrow \infty} \underbrace{\int_{\mathbb{R}^k} f_n(x_1, \dots, x_k, x_{k+1}, \dots, x_d) d(x_1, \dots, x_k)}_{=: F_n(x_{k+1}, \dots, x_d)} \quad (\star) \end{aligned}$$

Da $(F_n)_{n \in \mathbb{N}} \in \mathcal{C}_c(\mathbb{R}^{d-k})$ liegt und $F_n \uparrow F$ gilt, folgt $F \in \mathcal{H}^\uparrow(\mathbb{R}^{d-k})$ und

$$\begin{aligned} &\int_{\mathbb{R}^{d-k}} F(x_{k+1}, \dots, x_d) d(x_{k+1}, \dots, x_d) \\ &\stackrel{F_n \uparrow F}{=} \lim_{n \rightarrow \infty} \int_{\mathbb{R}^{d-k}} F(x_{k+1}, \dots, x_d) d(x_{k+1}, \dots, x_d) \\ &\stackrel{(\star)}{=} \lim_{n \rightarrow \infty} \int_{\mathbb{R}^{d-k}} \left(\int_{\mathbb{R}^k} f_n(x_1, \dots, x_k, x_{k+1}, \dots, x_d) d(x_1, \dots, x_k) \right) d(x_{k+1}, \dots, x_d) \\ &= \lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} f_n(x) dx = \int_{\mathbb{R}^d} f(x) dx. \quad \square \end{aligned}$$

6.6 Berechnung einige Volumina

Ziel

Sei $K \subset \mathbb{R}^d$ eine kompakte Teilmenge. Berechne

$$\text{Vol}_d(K) := \int_K 1 dx = \int_{\mathbb{R}^d} \mathbf{1}_K(x) dx$$

Lemma 6.6.1

Seien $k \in \{1, \dots, d-1\}$ und $K_1 \subset \mathbb{R}^k$ und $K_2 \subset \mathbb{R}^{d-k}$ kompakte Teilmengen.

Dann gilt

$$\text{Vol}_d(K_1 \times K_2) = \text{Vol}_k(K_1) \cdot \text{Vol}_{d-k}(K_2).$$

Beweis. Da die Menge $K_1 \times K_2$ kompakt ist, ist $\text{Vol}_d(K_1 \times K_2)$ wohldefiniert.

Aufgrund von

$$\mathbf{1}_{K_1 \times K_2}(x) = \mathbf{1}_{K_1}(x_1, \dots, x_k) \times \mathbf{1}_{K_2}(x_{k+1}, \dots, x_d)$$

folgt aus dem Satz 6.5.4

$$\begin{aligned} \text{Vol}_d(K_1 \times K_2) &= \int_{\mathbb{R}^{d-k}} \left(\mathbf{1}_{K_2}(x_{k+1}, \dots, x_d) \underbrace{\int_{\mathbb{R}^k} \mathbf{1}_{K_1}(x_1, \dots, x_k)}_{=\text{Vol}_k(K_1)} \right) d(x_{k+1}, \dots, x_d) \\ &= \text{Vol}_k(K_1) \cdot \int_{\mathbb{R}^{d-k}} \mathbf{1}_{K_2}(x_{k+1}, \dots, x_d) d(x_{k+1}, \dots, x_d) \\ &= \text{Vol}_k(K_1) \cdot \text{Vol}_{d-k}(K_2). \end{aligned} \quad \square$$

Beispiel 6.6.2 (1) Für eine Quader ist $K = \times_{i=1}^n [a_i, b_i]$ mit $-\infty < a_i < b_i < \infty$

für alle $i \in \{1, \dots, d\}$. Dann folgt durch d -malige Anwendung des Lemmas

6.6.1

$$\text{Vol}_d(K) = \prod_{i=1}^d \text{Vol}_1([a_i, b_i]) = \prod_{i=1}^d (b_i - a_i).$$

(2) Für einen Zylinder $K = B \times [0, h]$ sei $B \subset \mathbb{R}^{d-1}$ seine Grundfläche und $h \geq 0$ seine Höhe. Dann folgt mit dem Lemma 6.6.1

$$\text{Vol}_d(K) = \text{Vol}_{d-1}(B) \cdot \text{Vol}_1([0, h]) = h \cdot \text{Vol}_{d-1}(B).$$

Lemma 6.6.3 (Volumen unter linearen Transformationen)

Seien $A \in \mathbb{R}^{d \times d}$ eine Matrix, $b \in \mathbb{R}^d$ ein Vektor und $K \subset \mathbb{R}^d$ eine kompakte Teilmenge. Dann gilt

$$\text{Vol}_d(AK + b) := \text{Vol}_d(\{Ak + b : x \in K\}) = |\det(A)| \text{Vol}_d(K).$$

Beweis. Fall 1: $\det(A) \neq 0$. Dann folgt die Aussage aus dem Satz 6.5.3. TODO

Fall 2: $\det(A) = 0$. Da die Menge $AK + b$ in einer Hyperebene enthalten ist, welche sich durch eine orthogonale Transformation (welches das Volumen nicht verändert!) auf die Hyperebene

$$\{x \in \mathbb{R}^d : x_d = 0\}$$

überführen lässt (vgl. Lineare Algebra), folgt, dass die Menge $AK + b$ ein ‘‘Zylinder’’ mit Höhe $h = 0$ ist, also folgt aus dem obigen Beispiel $\text{Vol}_d(AK + b) = 0$ und daraus die Behauptung. \square

SATZ 6.6.1: PRINZIP VON CAVALIERI

Sei $K \subset \mathbb{R}^d$ eine kompakte Teilmenge und für $t \in \mathbb{R}^d$

$$K_t := \{(x_1, \dots, x_{d-1}) \in \mathbb{R}^{d-1} : (x_1, \dots, x_{d-1}, t) \in K\}$$

t-Schnitte der (kompakte) *t*-Schnitte von K . Dann gilt

$$\text{Vol}_d(K) = \int_R \text{Vol}_{d-1}(K_t) dt.$$

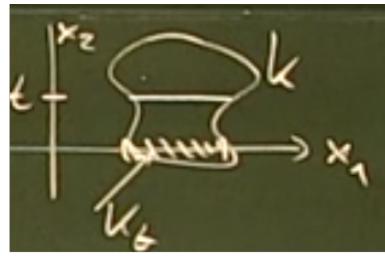


Abb. 40: Ein Beispiel für den obigen Satz für $d = 2$ [Quelle: Vorlesung]

Beweis. Weil

$$\mathbf{1}_{K_t}(x_1, \dots, x_{d-1}) = \mathbf{1}_K(x_1, \dots, x_{d-1}, t)$$

gilt, folgt die Aussage aus dem Satz 6.5.4. \square

Bemerkung 6.6.4 Sei $K \subset \mathbb{R}^d$ eine kompakte Teilmenge und $r > 0$. Dann gilt mithilfe von Lemma 6.6.3

$$\text{Vol}_d(rK) := \text{Vol}_d(\{rk : k \in K\}) = \text{Vol}_d(rI_d K) = r^d \text{Vol}_d(K).$$

Beispiel 6.6.5 (Volumen eines d -dimensionalen Kegels) Seien $B \subset \mathbb{R}^{d-1}$ eine kompakte Menge und $h > 0$ und

$$C_h(B) := \left\{ \begin{pmatrix} (1-\lambda)h \\ \lambda h \end{pmatrix} : \lambda \in [0, 1] \right\}.$$

Dann sind die t -Schnitte des Kegels

$$C_h(B)_t = \begin{cases} (1 - \frac{t}{h})B, & \text{wenn } t \in [0, h], \\ \emptyset, & \text{sonst.} \end{cases}$$

Also gilt mithilfe der obigen Bemerkung

$$\text{Vol}_d(C_h(B)) = \int_0^h \text{Vol}_{d-1} \left(\left(1 - \frac{t}{h} \right) B \right) dt = \text{Vol}_{d-1}(B) \int_0^h \left(1 - \frac{t}{h} \right)^{d-1} dt = \frac{h}{d} \text{Vol}_{d-1}(B).$$

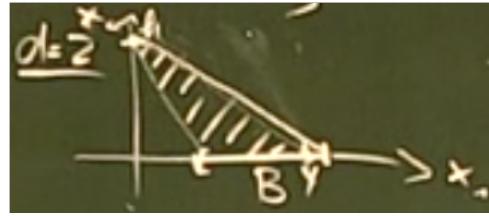


Abb. 41: [Quelle: Vorlesung]

19.11.2018

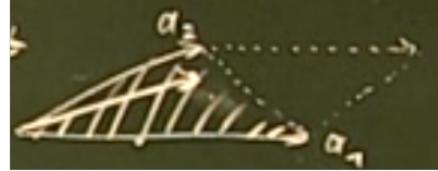


Abb. 42: [Quelle: Vorlesung]

Beispiel 6.6.6 (Volumen eines Simplex) Ein durch die Vektoren a_1, \dots, a_d aufgespannter Simplex ist

$$S(a_1, \dots, a_d) := \left\{ \sum_{k=1}^d \lambda_k a_k, \lambda_1, \dots, \lambda_d \geq 0, \sum_{k=1}^d \lambda_k \leq 1 \right\}.$$

Aus dem Lemma 6.6.3 bzw. Beispiel 6.2.7 folgt nun

$$\begin{aligned} \text{Vol}_d(S(a_1, \dots, a_d)) &= |\det(a_1, \dots, a_d)| \cdot \text{Vol}_d(S(e_1, \dots, e_d)) \\ &= |\det(a_1, \dots, a_d)| \cdot \frac{1}{d!} \text{Vol}_d(\underbrace{P(e_1, \dots, e_d)}_{=[0,1]^d}) = \frac{|\det(a_1, \dots, a_d)|}{d!}, \end{aligned}$$

die zwei Gleichungen wollen wir nun einsehen, wo bei P der Definition 6.2.5 in definierte Spat ist.

Aufgrund von der Darstellung

$$S(e_1, \dots, e_d) = C_1(e_1, \dots, e_{d-1})$$

Nach der Formel aus Beispiel 6.6.5

$$\text{Vol}_d(S(e_1, \dots, e_d)) = \frac{1}{d} \text{Vol}_{d-1}(S(e_1, \dots, e_{d-1})) = \dots = \frac{1}{d!} \text{Vol}_1([0, 1]) = \frac{1}{d!}.$$

Beispiel 6.6.7 (Volumen einer Kugel) Sei $K_d(r) := \{x \in \mathbb{R}^d : \|x\| \leq r\}$ die d -dimensionale Kugel mit Radius $r > 0$.

Wir berechnen nun das Volumen der Einheitskugel $\tau_d := \text{Vol}_d(K_d(1))$ mit dem Prinzip von CAVALIERI.

Die t -Schnitte sind nun gegeben durch

$$(K_d)_t = \begin{cases} \emptyset, & \text{wenn } |t| > 1, \\ K_{d-1}(\sqrt{1-t^2}), & \text{wenn } |t| \leq 1. \end{cases}$$

Nach dem Prinzip von CAVALIERI gilt also

$$\tau_d = \int_{-1}^1 \text{Vol}_{d-1}(K_{d-1}(\sqrt{1-t^2})) dt = \underbrace{\int_{-1}^1 \sqrt{1-t^2})^{d-1} dt}_{=:c_d} \tau_{d-1}.$$

Mit der Substitution $t = \cos(x)$ folgt

$$\int_0^\pi \sin^d(x) dx = 2 \int_0^{\frac{\pi}{2}} \sin^{d-1}(x) \cdot \sin(x) dx$$

Mit der Produktregel findet man eine Rekursionsformel in Termen von \sin^{d-2} und schließlich

$$c_{2d} = \pi \prod_{n=1}^d \frac{2n-1}{2n}, \quad \text{und} \quad c_{2d+1} = 2 \prod_{n=1}^d \frac{2n}{2n+1},$$

und somit

$$\tau_d = \frac{\pi^{\frac{d}{2}}}{\Gamma(1 + \frac{d}{2})}$$

Beispiel 6.6.8 (Volumen eines Ellipsoids) Für positive Zahlen $\alpha_1, \dots, \alpha_d > 0$

Ellip ist in d -dimensionales Ellipsoid

$$E(\alpha, 1, \dots, \alpha_d) := \left\{ (x_1, \dots, x_d) : \sum_{i=1}^d \frac{x_i^2}{\alpha_i^2} \leq 1 \right\}.$$

Dann ist nach Lemma 6.6.3

$$\text{Vol}_d(E(\alpha, 1, \dots, \alpha_d)) = \tau_d \cdot \prod_{i=1}^d \alpha_i.$$

7 Einführung in die Maßtheorie

Ziel

Auf einer beliebigen Menge Ω ein Maß μ definieren, welche die folgenden “schönen Eigenschaften“ besitzt:

- ① $\mu(\emptyset) = 0$
- ② $\mu(A) \geq 0$.
- ③ $\mu(\bigcup_{n \in \mathbb{N}} A_n) \leq \sum_{n \in \mathbb{N}} \mu(A_n)$, wobei die Gleichheit gelten soll, wenn die A_n paarweise disjunkt sind.

Genauer soll μ auf einer geeigneten Teilmenge der Potenzmenge von Ω , also einem **Teilmengensystem** $\mathcal{F} \subset \mathcal{P}(\Omega)$ von Ω definiert sein.

Warum wählen wir nicht $\mathcal{F} := \mathcal{P}(\Omega)$?

Weil es z.B. auf Widersprüche führt, wenn man fordert: $\Omega := \mathbb{R}^2$, $\mathcal{F} := \mathcal{P}(\mathbb{R}^2)$, dass μ jedem Rechteck seinen elementargeometrischen Inhalt zuweist zusammen mit den obigen Rechenregeln (Gegenbeispiel: VITALI-Menge).

7.1 Motivation und Einführung: Das LEBESGUE-Maß auf \mathbb{R}^2

Definition	1.1.1	metrischer Raum, Metrik	4
counter.1.1.1 Definition	1.1.6	Spurmetrik	6
counter.1.1.6 Definition	1.1.8	Beschränktheit	6
counter.1.1.8 Definition	1.1.9	Durchmesser	6
counter.1.1.9 Lemma	1.1.11		7
counter.1.1.11 Definition	1.1.12	Menge der beschränkten Abbildung	7
counter.1.1.12 Definition	1.2.1	Topologischer Raum, Umgebung	18
counter.1.2.1 Definition	1.2.5	Umgebung, offen, abgeschlossen	19
counter.1.2.5 Lemma	1.2.6	Äquivalenz von Ana-I-Offenheit und neuer Definition	19
counter.1.2.6 Definition	1.2.9	Hausdorff'sch	21
counter.1.2.9 Lemma	1.2.14	DEMORGANSchen Regeln	23
counter.1.2.14 Definition	1.2.15	Spur- bzw. Unterraumtopologie	23
counter.1.2.15 Definition	1.2.18	Inneres, abgeschlossene Hülle, Rand	26
counter.1.2.18 Definition	1.3.1	Konvergente Folge, Grenzwert, divergent	34
counter.1.3.1 Lemma	1.3.2	Äquivalente Metriken	34
counter.1.3.2 Definition	1.3.5	Gleichmäßige Konvergenz	35
counter.1.3.5 Definition	1.3.8	d -CAUCHY-Folge	37
counter.1.3.8 Definition	1.3.9	Vollständigkeit	37
counter.1.3.9 Lemma	1.3.12	Vollständigkeit	38
counter.1.3.12 Lemma	1.3.15	CAUCHY-Folgenkriterium	39
counter.1.3.15 Definition	1.3.17	Kontraktion(smodul)	39
counter.1.3.17 Definition	1.3.20	gewöhnliche Differentialgleichung, AWP	41
counter.1.3.20 Definition	1.4.1	offene Überdeckung	45
counter.1.4.1 Definition	1.4.2	(Überdeckungs)Kompaktheit	45
counter.1.4.2 Lemma	1.4.4		45
counter.1.4.4 Definition	1.4.7	Folgenkompaktheit	48
counter.1.4.7 Definition	1.5.1	Zusammenhang	56
counter.1.5.1 Definition	1.6.1	Stetigkeit	61
counter.1.6.1 Definition	1.6.3	Partielle Stetigkeit	61
counter.1.6.3 Definition	1.7.4		70
counter.1.7.4 Lemma	1.7.5	Existenz der LEBESGUEN Zahl	71
counter.1.7.5 Definition	1.8.1	normierter Vektorraum	74
counter.1.8.1 Lemma	1.8.6	umgekehrte $\Delta \neq$	74
counter.1.8.6 Definition	1.8.8	Vektorraum der linearen Abbildungen	75
counter.1.8.8 Definition	1.8.9	Vektorraum-Dimension, Isomorphismus	75
counter.1.8.9 Definition	1.8.15		79
counter.1.8.15 Definition	2.1.1		84
counter.2.1.1 Lemma	2.1.3	Eindeutigkeit des Differentials D_p	84
counter.2.1.3 Lemma	2.1.9	Komponentenweise Differentiation	87
counter.2.1.9 Definition	2.1.10	Funktionalmatrix	88
counter.2.1.10 Lemma	2.2.1		92
counter.2.2.1 Definition	2.3.1	partielle- und Richtungsableitung	99
counter.2.3.1 Definition	2.4.2	Zweite Ableitung	103
counter.2.4.2 Definition	2.4.4	Höhere Ableitungen	104
counter.2.4.4 Lemma	2.4.5		104
counter.2.4.5 Lemma	2.4.9	Ableiten ist "assoziativ"	106
counter.2.4.9 Definition	2.4.10	C^k -Funktionen	107
counter.2.4.10 Lemma	2.4.11		107
counter.2.4.11 Definition	2.6.1		116
counter.2.6.1 Lemma	2.6.2		116
counter.2.6.2 Definition	2.6.3	EINFÜHRUNG IN DIE MASSTHEORIE	117
counter.2.6.3 Lemma	2.6.4	Hauptminorenkriterium	118
counter.2.6.4 Definition	2.7.1		120

Abbildungsverzeichnis

1	[Quelle]	3
2	Der Durchmesser der Menge A	7
3	Veranschaulichung der Metrik d^{sup} . [Quelle: Ferus]	8
4	Man sich einen Fisch vorstellen, der in einem Flusses entlang der x -Achse schwimmt, welcher unendlich viele Nebenflüsse hat, die parallel zu der y -Achse verlaufen. [Quelle]	10
5	Beispiele und Gegenbeispiele zu Topologien - die sechs Abbildungen stellen Teilmengen der Potenzmenge von $\{1, 2, 3\}$ dar (der kleine Kreis links oben ist jeweils die leere Menge). Die ersten vier sind Topologien: im Beispiel unten links fehlt $\{2, 3\}$, unten rechts $\{2\}$ zur Topologie-Eigenschaft. [Quelle]	18
6	Die ε -Kugeln Umgebungen für die Standardmetrik und für d^∞ auf \mathbb{R}^2 . [Quelle]	19
7	Veranschaulichung, dass offen/abg. ε -Kugeln offen/abg. sind. [Quelle]	20
8	Die DEMORGAN'schen Gesetze für zwei Mengen $A, B \subset u$. [Quelle]	23
9	Das Distributivgesetz $\bigcup_{i \in I} (X_i \cap Y) = Y \cap \bigcup_{i \in I} X_i$. [Quelle: WA]	24
10	Hier ist X die x - y -Ebene, A die x -Achse und b das offene Intervall $(0.5, 1)$. Die in X offene Menge Y ist die offene Kreisscheibe	25
11	(a) Eine Teilmenge $M \subset \mathbb{R}^2$, (b) ihr Inneres, (c) ihr Rand und (d) ihr Abschluss. [Quelle]	26
12	Die Ebene E und die Kreisscheibe U . [Quelle: Ferus]	27
13	Die Funktionen x, x^2, x^3, x^4 auf dem Intervall $[0, 1]$. [Quelle: Ferus]	36
14	Die Funktionen g_k und f_k für $k \in \{1, \dots, 4\}$ sowie die Grenzfunktion f . [Quelle: GeoGebra]	37
15	Eine Abbildung $f : \mathbb{R} \rightarrow \mathbb{R}$, welche in einer Umgebung um einen Fixpunkt x^* kontrahierend ist. [Quelle: Vorlesung]	40
16	Veranschaulichung der 2^n -Teilung des Quaders in $n = 2$. [Quelle: Vorlesung]	46
17	Visualisierung für die Argumentation der Abgeschlossenheit	47
18	Die Funktionenfolge $(f_n : [0, 1] \rightarrow X, x \mapsto 1 - x^n)_{n \in \mathbb{N}}$ für $n \in \{1, 2, 3, 4, 5, 10, 100\}$	55
19	A ist zusammenhängend, B, A^C und B^C sind es nicht. [Quelle]	56
20	Quelle: Ferus	62
21	Eine Verdeutlichung der Rückrichtung von (i)	65
22	Die Funktion $f : (0, \infty) \rightarrow \mathbb{R}, x \mapsto \frac{\ln x}{x}$ [Quelle: GeoGebra6]	69
23	Geometrische Interpretation [Quelle: Ferus]	86
24	Der Graph der Funktion $f(x, y) := xy$ und die Tangentialebene an dem Punkt $E = (2.2, 1.7)$, welche durch $f(a, b) + (x - a)b + (y - b)a$ gegeben ist [Quelle: GeoGebra6]	87
25	Die im obigen Beispiel angeführte Funktion [Quelle: Ferus]	100
26	[Quelle: GeoGebra5]	101
27	RIEMANN-Integral mit Unter- und Obersummen	165
28	Wirkung des Shift-Operators τ_a auf eine Funktion f . [Quelle: Forster]	167
29	Beispiel für so eine Funktion φ . [Quelle: Forster]	169
30	f_j für $j \in \{1, \dots, 6\}$. [Quelle: GeoGebra6]	169

31	Die Zacken- oder Hütchenfunktion ψ sowie $\Psi(x)$ (schwarz) und $\Psi_3(x)$ (blau) für $d = 2$. [Quelle: Forster und GeoGebra6]	170
32	[Quelle: Forster]	170
33	[Quelle: Forster]	171
34	g (blau), G (lila) und h (grün) für $k \in [-33, 33]$ anstatt ganz \mathbb{Z} . [Quelle: GeoGebra6]	182
35	$\alpha_{1,0,7}$ für $d = 2$. [Quelle: GeoGebra6]	183
36	[Quelle: Vorlesung]	183
37	[Quelle: Vorlesung]	183
38	[Quelle: Forster]	195
39	Ein Beispiel für den obigen Satz für $d = 2$ [Quelle: Vorlesung]	199
40	[Quelle: Vorlesung]	200
41	[Quelle: Vorlesung]	200