

Analysis – Weiterführende Konzepte

Aktualisiert am 19. Oktober 2020

Stefan Neukamm ¹

**Dieser Text ist nur für die Hörer der Vorlesung
Analysis – Weiterführende Konzepte an der TU Dresden,
Sommersemester 2020 bestimmt.² Er darf nicht weitergegeben werden.**

**Für Anmerkungen und Korrekturen an stefan.neukamm@tu-dresden.de bin ich
dankbar.**

¹Professur für Angewandte Analysis, TU Dresden,
<https://tu-dresden.de/mn/math/wir/neukamm>

²https://youtu.be/T5_w0Xq7-LY

Inhaltsverzeichnis

1	Normierte Räume	3
1.1	Wiederholung - Struktur von $\mathbb{R}^2$	3
1.2	Vektorraum	3
1.3	Normierte Räume	11
1.4	Konvexe Mengen und Funktionen	17
1.5	Endlichdimensionale normierte Räume	27
1.6	Hilberträume	37
1.7	Konvergenz von Funktionenfolgen	46
2	Differentialrechnung im Mehrdimensionalen	52
2.1	Kurven im $\mathbb{R}^d$	52
2.2	Totale Ableitung	73
2.3	Partielle Ableitung	80
2.4	Skalarfeld, Niveaumenge, Gradient	91
2.5	Taylorentwicklung	94
2.6	Lokale Extrema von Skalarfeldern	102
2.7	Vektorfelder, Divergenz und Rotation	108
3	Banachscher Fixpunktsatz und Anwendungen	116
3.1	Diffeomorphismen	119
3.2	Satz über implizite Funktionen	128
4	Das Riemannintegral	132
4.1	Motivation	132
4.2	Grundlegende Eigenschaften	136
4.3	Das eindimensionale Riemannintegral	154
4.4	Das n -dimensionale Volumen	167
4.5	Transformationssatz	173
5	Lokale Extrema unter Nebenbedingungen	174
6	Kurvenintegral	176
7	Topologie und Kompaktheit in metrischen Räumen	182
7.1	Topologie	182
7.2	Kompaktheit	184
A	Schwerpunkte der Vorlesung	189

1 Normierte Räume

1.1 Wiederholung - Struktur von $\mathbb{R}^2$

- *Menge.* $\mathbb{R}^2 = \{x = (x_1, x_2) : x_1, x_2 \in \mathbb{R}\} = \mathbb{R} \times \mathbb{R}$ (Kartesisches Produkt von Mengen).
- *Vektorraum.* Vektoren $x, y \in \mathbb{R}^2$, Skalar $\lambda \in \mathbb{R}$, Linearkombination: $z = \lambda x + y$,
 $0 := (0, 0)$ (Nullvektor = neutrales Element der Vektoraddition)
- *Metrischer Raum.* $d_2(x, y) := \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$ (Euklidischer Abstand)
- *Normierter Raum.* $|x|_2 := d_2(x, 0)$ (Euklidische Norm)
- *Hilbertraum.* Winkel: $\angle(x, y)$, Orthogonalität: $x \perp y$, Skalarprodukt: $x \cdot y = x_1 y_1 + x_2 y_2$

1.2 Vektorraum

Ziele:

- Konzept: Vektorraum
- Beispiel: $\mathbb{R}^d$ in mathematischen Modellen
- Beispiel: Funktionenräume

In dieser Vorlesung steht $\mathbb{K}$ stellvertretend für $\mathbb{R}$ oder $\mathbb{C}$. Aus der LA kennen wir:

Definition (Vektorraum). Ein Tripel $(X, +, \cdot)$ bestehend aus einer Menge X und zweier Abbildungen

$$\begin{array}{lll} + : X \times X \rightarrow X, & (x, y) \mapsto x + y & \text{(Vektorraum)addition,} \\ \cdot : \mathbb{K} \times X \rightarrow X & (\lambda, x) \mapsto \lambda x & \text{Skalarmultiplikation,} \end{array}$$

heißt *Vektorraum über $\mathbb{K}$* (kürzer: *$\mathbb{K}$ -Vektorraum* oder einfach *Vektorraum*), falls gilt:

- (a) $(X, +)$ ist eine abelsche Gruppe mit Nullelement 0 .
- (b) Für alle *Skalare* $\lambda, \mu \in \mathbb{K}$ und *Vektoren* $x, y \in X$ gilt

$$\begin{aligned} 1x &= x, \\ (\lambda\mu)x &= \lambda(\mu x), \\ (\lambda + \mu)x &= \lambda x + \mu x, \\ \lambda(x + y) &= \lambda x + \lambda y. \end{aligned}$$

Bemerkung. In dieser Analysis-Vorlesung darf man gerne immer $\mathbb{K} = \mathbb{R}$ setzen. Wir werden nur in sehr wenigen Situationen speziell Vektorräume über $\mathbb{C}$ betrachten.

Wichtige Konzepte für Vektorräume (bekannt aus LA):

- *Lineare Abbildung, Isomorphismus, $X \simeq Y$* :

Seien X, Y Vektorräume.

- Eine Abbildung $T : X \rightarrow Y$ heißt *linear*, falls für alle $\lambda \in \mathbb{K}$ und $x, y \in X$ gilt

$$T(\lambda x + y) = \lambda T(x) + T(y)$$

(Hinweis: Bei linearen Abbildungen schreibt man gerne einfach Tx statt $T(x)$. Wir werden dies in dieser Vorlesung aber nicht tun).

- Eine lineare bijektive Abbildung $T : X \rightarrow Y$ heißt Isomorphismus. Man nennt X zu Y isomorph ($X \simeq Y$), falls ein Isomorphismus von X nach Y existiert.

- *Untervektorraum, Linearkombination, Lineare Hülle* :

Seien X Vektorraum, $U \subseteq X$ Menge.

- U heißt Untervektorraum von X , wenn U zusammen mit der Addition und Skalarmultiplikation von X ein Vektorraum ist.³
- Seien $x_1, \dots, x_m \in U$, $\lambda_1, \dots, \lambda_m \in \mathbb{K}$, $m \in \mathbb{N}$, dann heißt $x = \sum_{i=1}^m \lambda_i x_i$ Linearkombination der Vektoren $x_1, \dots, x_m$.⁴ Die Menge

$$\text{span}(U) = \left\{ \sum_{i=1}^m \lambda_i x_i : m \in \mathbb{N}, x_1, \dots, x_m \in U, \lambda_1, \dots, \lambda_m \in \mathbb{K} \right\}$$

aller Linearkombinationen von Vektoren aus U heißt *lineare Hülle*.⁵

- *Linear unabhängig, Basis, Koordinaten(vektor), Dimension* :

Sei X Vektorraum, $U \subseteq X$.

- Eine (endliche) Menge von Vektoren $x_1, \dots, x_m \in X$ heißt linear unabhängig, falls für alle $\lambda_1, \dots, \lambda_m \in \mathbb{K}$ gilt

$$\lambda_1 x_1 + \dots + \lambda_m x_m = 0 \quad \Rightarrow \quad \lambda_1 = \dots = \lambda_m = 0.$$

- Eine (möglicherweise unendliche Menge) $U \subseteq X$ heißt linear unabhängig, falls jede beliebige endliche Menge von Vektoren $x_1, \dots, x_m \in U$ linear unabhängig ist.
- Eine Menge $B \subseteq X$ heißt Basis von X , falls B linear unabhängig ist und $\text{span}(B) = X$ gilt.
- Mit dem Lemma von Zorn kann man beweisen, dass jeder Vektorraum X eine Basis B besitzt. Die Kardinalität von B heißt *Dimension* von X und wird mit $\dim(X)$ bezeichnet.

³Beispiel: $X = \mathbb{R}^2$, $U := \{(x_1, x_2) \in \mathbb{R}^2 : 3x_1 + 2x_2 = 0\}$ ist Untervektorraum von X . Andererseits sind die Mengen $\{(x_1, 1) \in \mathbb{R}^2 : x_1 \in \mathbb{R}\}$ und $\{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 = x_2\}$ keine Untervektorräume von X .

⁴Linearkombinationen sind immer Summen endlich vieler Vektoren.

⁵Beispiel: Für $X = \mathbb{R}^2$, $U = \{(1, 1), (0, 1)\}$ gilt $\text{span}(U) = \mathbb{R}^2$

- Sei $\dim(X) = d \in \mathbb{N}_+$, d.h. X ist ein endlicher Vektorraum, und ist $b_1, \dots, b_d$ eine Basis, so besitzt jeder Vektor $x \in X$ eine eindeutige Darstellung der Form $x = \sum_{i=1}^d \lambda_i b_i$. Man nennt den Vektor

$$\lambda = (\lambda_1, \dots, \lambda_d) \in \mathbb{K}^d$$

den Koordinatenvektor von x bezüglich der Basis $b_1, \dots, b_d$. Die Abbildung $X \rightarrow \mathbb{K}^d$, $x \mapsto \lambda$ ist ein Isomorphismus.

Wichtigstes Beispiel für einen Vektorraum:

Beispiel 1.1 (Der Vektorraum $\mathbb{K}^d$). Für $d \in \mathbb{N}_+$ (die Dimension) bezeichnet

$$\mathbb{K}^d := \underbrace{\mathbb{K} \times \dots \times \mathbb{K}}_{d\text{-mal}}$$

die Menge aller d -Tupel $x = (x_1, \dots, x_d)$ mit Koordinaten $x_\alpha \in \mathbb{K}$ für $\alpha \in \{1, \dots, d\}$.

Addition und Skalarmultiplikation für $x, y \in \mathbb{K}^d$, $\lambda \in \mathbb{K}$:

$$\lambda x + y := \lambda(x_1, \dots, x_d) + (y_1, \dots, y_d) = (\lambda x_1 + y_1, \dots, \lambda x_d + y_d) \quad (\text{"komponentenweise"}).$$

Die Menge der d -Tupel

$$e_\alpha := (0, \dots, 0, \underbrace{1}_{\alpha\text{te Koordinate}}, 0, \dots, 0) \in \mathbb{K}^d, \quad \alpha = 1, \dots, d$$

heißt Standardbasis von $\mathbb{K}^d$.

Bemerkung 1.2 (Lineare Abbildungen von $\mathbb{K}^d$ nach $\mathbb{K}^n$). Jeder linearen Abbildung $T: \mathbb{K}^d \rightarrow \mathbb{K}^n$ kann eindeutig eine Matrix $A \in \mathbb{K}^{n \times d}$ zugeordnet werden, so dass

$$y = T(x) \quad \Leftrightarrow \quad y_i = \sum_{k=1}^d A_{ik} x_k \quad \text{für alle } i = 1, \dots, n,$$

für alle $x \in \mathbb{K}^d$ und $y \in \mathbb{K}^n$ gilt.

Bemerkung 1.3 (Notation).

- 3 Schreibweisen: d -Tupel, Spaltenvektor, Zeilenvektor

$$x = (x_1, \dots, x_d) \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix} \quad x = (x_1 \ \dots \ x_d)$$

- Matrix $\times$ Vektor - Produkt
- Zeilenvektor $= 1 \times d$ Matrix = Lineare Abbildung von $\mathbb{K}^d$ nach $\mathbb{K}$.

Alle endlichdimensionalen $\mathbb{K}$ -Vektorräume sehen wie $\mathbb{K}^d$ aus:

Satz. Sei X ein $\mathbb{K}$ -Vektorraum mit endlicher Dimension d . Dann gilt $X \simeq \mathbb{K}^d$.

Beweis siehe LA. Beweisidee. Sei $B = (b_1, \dots, b_d)$ eine Basis von X . Jeder Vektor $x \in X$ besitzt eine eindeutige Darstellung bzgl. B ,

$$(1.1) \quad x = \sum_{i=1}^d \lambda_i b_i \quad \text{mit } (\lambda_1, \dots, \lambda_d) \in \mathbb{K}^d.$$

Die Abbildung

$$T : X \rightarrow \mathbb{K}^d, \quad x \mapsto \lambda \quad (\text{mit } x, \lambda \text{ gemäß (1.1)})$$

ist ein Isomorphismus. □

Unterräume sind häufig als Kern und Bild linearer Abbildungen gegeben.

Lemma 1.4. *Seien X, Y $\mathbb{K}$ -Vektorräume und $T : X \rightarrow Y$ eine lineare Abbildung. Die Mengen*

$$\begin{aligned} \text{Kern}(T) &:= \{x \in X : T(x) = 0\} \subseteq X \\ \text{Bild}(T) &:= \{T(x) : x \in X\} \subseteq Y \end{aligned}$$

heißen Kern und Bild von T und sind Untervektorräume von X , bzw. Y .

Beweis.

1. Seien $x_1, x_2 \in \text{Kern}(T)$, $\lambda \in \mathbb{K}$. Dann gilt

$$T(\lambda x_1 + x_2) = \lambda T(x_1) + T(x_2) = 0,$$

also $\lambda x_1 + x_2 \in \text{Kern}(T)$.

2. Seien $y_1, y_2 \in \text{Bild}(T)$, $\lambda \in \mathbb{K}$. Dann existieren $x_1, x_2 \in X$ mit $T(x_1) = y_1$ und $T(x_2) = y_2$. Es gilt

$$\lambda y_1 + y_2 = \lambda T(x_1) + T(x_2) = T(\lambda x_1 + x_2),$$

also $\lambda y_1 + y_2 \in \text{Bild}(T)$. □

Kontrollfrage 1.1. *Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ und es bezeichne $U := \{x \in \mathbb{R}^2 : f(x) = 0\}$ die Nullniveau-Menge von f . Wann ist U ein Unterraum von $\mathbb{R}^2$?*

(A) $f(x) = x_1 + x_2 + 1$.

$$\times : 0 \notin U$$

(B) $f(x) = x_1^2 + x_2$.

$$\times : U \text{ nicht abgeschlossen bezüglich Vektorraumaddition}$$

(C) f ist linear.

$$\checkmark$$

Kontrollfrage 1.2. *Es bezeichne $e_1, \dots, e_d \in \mathbb{R}^d$ die Standardbasis von $\mathbb{R}^d$. Für $i = 1, \dots, d$ setze $b_i := \sum_{k=1}^i e_k$. Dann gilt:*

(A) Die lineare Hülle von $\{b_1, \dots, b_d\}$ ist echte Teilmenge von $\mathbb{R}^d$

$$\times: \text{span}\{b_1, \dots, b_d\} = \mathbb{R}^d!$$

(B) Die Vektoren $b_1, \dots, b_d$ sind linear unabhängig.

✓: Folgt aus $e_1, \dots, e_d \in \mathbb{R}^d$ Basis.

(C) $\{b_1, \dots, b_d\}$ ist Basis von $\mathbb{R}^d$.

✓: linear unabhängig und Dimensionsargument.

5.4.2020

#2

<https://youtu.be/ICAbDknGFGY>

5.4.2020

Anwendungen von $\mathbb{R}^d$ in mathematischen Modellen.

Beispiel 1.5 (Position eines Objektes). Die Position eines Objektes im (d -dimensionalen, affinen) Raum kann durch $x \in \mathbb{R}^d$ beschrieben werden:

- Bezugssystem =: Punkte im Raum $O, P_1, \dots, P_d$ mit $\overrightarrow{OP_1}, \dots, \overrightarrow{OP_d}$ linear unabhängig.
- Jeder Punkt X im Raum besitzt eindeutige Koordinatendarstellung $\overrightarrow{OX} = \sum_{i=1}^d x_i \overrightarrow{OP_i}$.

Beispiel: Position eines Fadenpendels der Länge ℓ im $d = 3$ -dimensionalen Raum. Wähle Bezugssystem mit Ursprung O = Aufhängepunkt und $\overrightarrow{OP_1}$ in Lotrichtung.

$$(\ell \cos(\alpha), \ell \sin(\alpha), 0) \in \mathbb{R}^3 \quad \text{mit } \alpha \in \mathbb{R} \text{ Auslenkungswinkel.}$$

Implizite Annahme: Pendel schwingt in der $\overrightarrow{OP_1} - \overrightarrow{OP_2}$ -Ebene.

Vereinfachte Beschreibung $x = \ell(\cos(\alpha), \sin(\alpha)) \in \mathbb{R}^2$.

Beispiel 1.6 (Bewegung eines Objektes). Die Trajektorie eines Objektes im d -dimensionalen Raum kann durch eine Funktion

$$x: I \rightarrow \mathbb{R}^d, \quad t \mapsto x(t), \quad \text{mit } I \subseteq \mathbb{R} \text{ Intervall,}$$

beschrieben werden. Hierbei beschreibt der Vektor $x(t)$ die Position des Objektes zum Zeitpunkt $t \in I$.

→ Funktionen dieser Art heißen **Kurven**.

Beispiel Fadenpendel (harmonische Approximation):

$$x(t) = \ell(\cos(\alpha(t)), \sin(\alpha(t))), \quad \alpha(t) := \alpha_0 \cos(\omega_0 t),$$

mit Parametern α_0 (maximaler Auslenkungswinkel), ω_0 (Winkelgeschwindigkeit).

Wir erinnern an die Notation: $\mathbb{R}_{\geq 0} := [0, \infty)$.

Beispiel 1.7 (Räumliche Verteilung in Epidemiemodellen). Epidemiemodelle zur räumlichen Ausbreitung von Viren beinhalten Funktionen der Form

$$u: \mathbb{R}_{\geq} \times \Omega \rightarrow \mathbb{R}_{\geq}, \quad (t, x) \mapsto u(t, x).$$

mit

$$\begin{array}{ll} t \in \mathbb{R}_{\geq} & \text{Zeit} \\ x = (x_1, x_2) \in \Omega \subset \mathbb{R}^2 & \text{räumliches Gebiet} \\ u(t, x) & \text{Infektionsdichte im Raumpunkt } x \text{ zum Zeitpunkt } t \end{array}$$

→ Funktionen auf (Teilmenge von) $\mathbb{R}^d$ mit Werten in $\mathbb{R}$ heißen **Skalarfelder** (oder skalare Funktionen).

Beispiel 1.8 (Strömungsfelder). Die Dynamik einer Flüssigkeit wird durch Funktionen der Form

$$v : \mathbb{R} \times \Omega \rightarrow \mathbb{R}^3, \quad (t, x) \mapsto v(t, x) = (v_1(t, x), \dots, v_3(t, x))$$

beschrieben, wobei $v(t, x)$ den Geschwindigkeitsvektor zum Zeitpunkt t im Raumpunkt x bezeichnet. → Funktionen auf (Teilmenge von) $\mathbb{R}^d$ mit Werten in einem Vektorraum heißen **Vektorfelder** (oder vektorwertige Funktionen).

Beispiel 1.9 (Bildbearbeitung). Graustufenbilder können mathematisch durch Skalarfelder der Form

$$f : (0, 1)^2 \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}_{\geq}, \quad x \mapsto f(x)$$

beschrieben werden, wobei $f(x)$ für den Grauwert des Bildpunktes an der Stelle x steht. RGB-Bilder (Farbbilder) werden durch Vektorfelder der Form

$$f : (0, 1)^2 \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}_{\geq}^3, \quad x \mapsto f(x) = (f_1(x), \dots, f_3(x))$$

beschrieben, wobei die drei Komponenten von $f(x)$ für die Farbkanäle rot, grün und blau stehen.

5.4.2020

#3 <https://youtu.be/4TVmlVenc14>

5.4.2020

In der Analysis spielen unendliche-dimensionale Vektorräume eine wichtige Rolle. Wichtigstes Beispiel sind Funktionenräume.

Definition 1.10 (Raum der (vektorraumwertigen) Abbildungen). Sei D eine nicht-leere Menge, Y ein $\mathbb{K}$ -Vektorraum. Die Menge aller Abbildungen von D nach Y ,

$$\text{Abb}(D; Y) := \{ f \mid f : D \rightarrow Y \},$$

wird mit der Addition

$$+ : \text{Abb}(D; Y) \times \text{Abb}(D; Y) \rightarrow \text{Abb}(D; Y), \quad (f + g)(x) := f(x) + g(x)$$

und Skalarmultiplikation

$$\cdot : \mathbb{K} \times \text{Abb}(D; Y) \rightarrow \text{Abb}(D; Y), \quad (\lambda f)(x) := \lambda f(x)$$

zu einem $\mathbb{K}$ -Vektorraum. Im Fall $Y := \mathbb{R}$ schreiben wir $\text{Abb}(D) := \text{Abb}(D; \mathbb{R})$.

Kontrollfrage 1.3. Zwei Funktionen $f, g \in \text{Abb}(\mathbb{R}; \mathbb{R})$ sind genau dann linear abhängig, falls

(A) $\exists \lambda, x \in \mathbb{R} : f(x) = \lambda g(x)$.

×

(B) $\exists \lambda \in \mathbb{R}, \forall x \in \mathbb{R} : f(x) = \lambda g(x)$.

✓

Kontrollfrage 1.4. Sei D eine nichtleere Menge. In welche Fällen ist $\text{Abb}(D; \mathbb{R})$ endlichdimensional?

(A) $D = \{1, 2\}$.

✓: isomorph zu $\mathbb{R}^2$

(B) $D = \mathbb{N}$.

×: $e_1, e_2, \dots$ sind linear unabhängig

(C) $D = (0, 1) \subseteq \mathbb{R}$.

×: z.B. sind $f(x) := x^\alpha, \alpha > 0$, linear unabhängig

Beispiele 1.11. (a) **Raum der linearen Abbildungen.** Seien X, Y $\mathbb{K}$ -Vektorräume. Die Menge der linearen Abbildungen

$$\text{Lin}(X; Y) := \{ f \in \text{Abb}(X; Y) : f \text{ ist linear} \}$$

ist ein Unterraum von $\text{Abb}(X; Y)$.

(b) **Raum der beschränkten Funktionen.** Sei D eine Menge. Der Raum $B(D; \mathbb{K})$ aller beschränkten Funktionen von D nach $\mathbb{K}$ ist ein Unterraum von $\text{Abb}(D; \mathbb{K})$.⁶

(c) **Raum der stetigen Funktionen.** Sei M ein metrischer Raum. Der Raum $C(M; \mathbb{K})$ aller stetigen Abbildungen von M nach $\mathbb{K}$ ist ein Unterraum von $\text{Abb}(M; \mathbb{K})$.

(d) **Raum der stetig differenzierbaren Funktionen.** Sei $D \subseteq \mathbb{R}$ offen und nichtleer, $k \in \mathbb{N}$. Der Raum $C^k(D; \mathbb{R})$ ist ein Unterraum von $\text{Abb}(D; \mathbb{R})$.

(e) **Folgenräume.** $\text{Abb}(\mathbb{N}; \mathbb{R})$ ist der Vektorraum aller Folgen in $\mathbb{R}$. Die Räume

$$c := \{ (x_i)_{i \in \mathbb{N}} \subseteq \mathbb{R} \mid \lim_{i \rightarrow \infty} x_i \text{ existiert} \} \quad (\text{Raum der konvergenten Folgen}),$$

$$c_0 := \{ (x_i)_{i \in \mathbb{N}} \subseteq \mathbb{R} \mid \lim_{i \rightarrow \infty} x_i = 0 \} \quad (\text{Raum der Nullfolgen}).$$

sind Unterräume von $\text{Abb}(\mathbb{N}; \mathbb{R})$.

Kontrollfrage 1.5. Wann ist folgende Aussage richtig? Für den Raum $X = \text{Lin}(\mathbb{R}^d; \mathbb{R}^m)$ gilt

(A) $\dim(X) = d + m$.

×

⁶Eine $\mathbb{K}$ -wertige Funktion heißt beschränkt, wenn $\text{Bild}(f)$ eine beschränkte Teilmenge von $\mathbb{K}$ ist.

(B) $\dim(X) = dm$.

✓: X isomorph zu $\mathbb{R}^{m \times d}$

(C) X ist unendlichdimensional.

×

Kontrollfrage 1.6. Wann ist folgende Aussage richtig? Sei $y \in \mathbb{R}$. Die Menge

$$X := \{(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{R} \mid \lim_{n \rightarrow \infty} x_n = y\}$$

ist genau dann ein Unterraum von $\text{Abb}(\mathbb{N}; \mathbb{R})$, wenn

(A) $y = 0$.

✓

(B) $y \neq 0$.

×: $0 \notin X$

Kontrollfrage 1.7. Sei $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) := \sin(x)$. Sind folgende Aussagen wahr?

(A) $f \in B(\mathbb{R}; \mathbb{R})$.

✓: $\text{Bild}(\sin) = [-1, 1]$.

(B) $f \in \text{Lin}(\mathbb{R}; \mathbb{R})$.

×: $\sin(2\pi) \neq 2 \sin(\pi)$.

(C) $f \in C^3(\mathbb{R}; \mathbb{R})$.

✓: $\sin$ ist beliebig oft differenzierbar.

Kontrollfrage 1.8. Wir betrachten die Folgenräume c_0 und c . Sei $x = (x_n)_{n \in \mathbb{N}} \in c$ mit $\lim_{n \rightarrow \infty} x_n = 1$ und sei $y = (y_n)_{n \in \mathbb{N}} \in c_0$. Betrachte $z = 5x + y$.

(A) Es gilt $z \in \text{Abb}(\mathbb{N}; \mathbb{R})$.

✓: $x, y \in \text{Abb}(\mathbb{N}; \mathbb{R})$ und $\text{Abb}(\mathbb{N}; \mathbb{R})$ ist Vektorraum

(B) Es gilt $z \in c_0$.

×

(C) Es gilt $z \in c$ mit $\lim_{n \rightarrow \infty} z_n = 5$.

✓: Satz 3.20 aus Analysis 1

1.3 Normierte Räume

Ziele:

- Konzept: Normierter Raum
- Normierte Räume als metrische Räume
- Beispiele: Maximums- und Supremumsnorm
- Stetige lineare Abbildungen

Eigenschaften des Betrages $|\cdot| : \mathbb{R} \rightarrow \mathbb{R}_{\geq}$ auf $\mathbb{R}$ (siehe AN10, Lemma 2.26): Für alle $x, y, \lambda \in \mathbb{R}$ gilt:

- (Definitheit). $|x| = 0 \Rightarrow x = 0$,
- (absolute Homogenität). $|\lambda x| = |\lambda||x|$,
- (Dreiecksungleichung). $|x + y| \leq |x| + |y|$.

Wir verallgemeinern diese Struktur auf allgemeine Vektorräume.

Definition 1.12 (Norm, normierter Raum). *Sei X ein $\mathbb{K}$ -Vektorraum. Ein Abbildung $\|\cdot\| : X \rightarrow \mathbb{R}_{\geq}$ heißt Norm auf X , falls für alle $x, y \in X$ und $\lambda \in \mathbb{K}$ gilt:*

- (Definitheit). $\|x\| = 0 \Rightarrow x = 0$,
- (absolute Homogenität). $\|\lambda x\| = |\lambda|\|x\|$
- (Dreiecksungleichung). $\|x + y\| \leq \|x\| + \|y\|$

Ist $\|\cdot\|$ eine Norm auf X , so heißt $(X, \|\cdot\|)$ normierter Raum.

Kontrollfrage 2.1. Welche der nachfolgenden Abbildungen $\varphi : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$ definiert eine Norm auf $\mathbb{R}$?

(A) $\varphi(x) := 10|x|$

✓

(B) $\varphi(x) := \begin{cases} |x| & \text{falls } x \geq 0 \\ 2|x| & \text{falls } x < 0 \end{cases}$

$\times : \varphi(-x) \neq \varphi(x)$

(C) $\varphi(x) := 1 + \sqrt{x^2}$

$\times : \varphi(0) \neq 0$

Beispiele 1.13. $(\mathbb{R}, |\cdot|)$ und $(\mathbb{C}, |\cdot|)$ sind normierte Räume.

Definition 1.14 (Maximumsnorm). Wir definieren die Maximumsnorm von $x = (x_1, \dots, x_d) \in \mathbb{K}^d$ durch $\|x\|_{\infty} := \max_{1 \leq i \leq d} |x_i|$.

Lemma 1.15. $(\mathbb{K}^d, \|\cdot\|_\infty)$ ist ein normierter Raum.

Beweis.

1. $\|x\|_\infty = 0 \Rightarrow \max_{1 \leq i \leq d} |x_i| = 0 \Rightarrow x_i = 0$ für alle $1 \leq i \leq d$.

2. Für $\lambda \in \mathbb{K}$ und $x \in \mathbb{K}^d$ gilt:

$$\|\lambda x\|_\infty = \max_{1 \leq i \leq d} |\lambda x_i| = |\lambda| \max_{1 \leq i \leq d} |x_i| = |\lambda| \|x\|_\infty.$$

3. Für $x, y \in \mathbb{K}^d$ gilt

$$\|x + y\|_\infty = \max_{1 \leq i \leq d} |x_i + y_i| \leq \max_{1 \leq i \leq d} (|x_i| + |y_i|) \leq \|x\|_\infty + \|y\|_\infty.$$

□

Kontrollfrage 2.2. Sei X ein endlichdimensionaler Vektorraum mit Dimension d und $T : X \rightarrow \mathbb{R}^d$ eine lineare Abbildung. Wir betrachten die Abbildung

$$\|\cdot\| : X \rightarrow \mathbb{R}_{\geq}, \quad \|x\| := \|T(x)\|_\infty.$$

Welche der nachfolgenden Aussagen sind richtig?

(A) $\|\cdot\|$ erfüllt die Dreiecksungleichung.

✓: Linearität von T , Dreiecksungleichung von $\|\cdot\|_\infty$

(B) $\|\cdot\|$ ist absolut homogen.

✓: Linearität von T , Homogenität von $\|\cdot\|_\infty$

(C) $\|\cdot\|$ ist positiv definit.

✗: Im Allgemeinen gilt nicht $T(x) = 0 \Rightarrow x = 0$

(D) $\|\cdot\|$ ist eine Norm auf X genau dann, wenn T injektiv ist.

✓

Definition 1.16 (Supremumsnorm). Sei D eine Menge. Für $f \in B(D; \mathbb{K})$ definieren wir die Supremumsnorm von f durch

$$\|f\|_\infty := \sup_{x \in D} |f(x)|.$$

Lemma 1.17. Sei D eine Menge. Dann ist $(B(D; \mathbb{K}), \|\cdot\|_\infty)$ ein normierter Raum.

Beweis. $B(D; \mathbb{K})$ ist als Unterraum von $\text{Abb}(D; \mathbb{K})$ ein Vektorraum. Eine Funktion $f \in B(D; \mathbb{K})$ ist beschränkt und damit gilt $\|f\|_\infty \in [0, \infty)$. Für $f, g \in B(D; \mathbb{K})$ und $\lambda \in \mathbb{K}$ gilt:

1. $f \neq 0 \Leftrightarrow \exists x \in D : f(x) \neq 0 \Leftrightarrow \|f\|_\infty > 0$.

2. $\|\lambda f\|_\infty = \sup_{x \in D} |\lambda f(x)| = \sup_{x \in D} |\lambda| |f(x)| = |\lambda| \|f\|_\infty$.

$$3. \|f + g\|_\infty = \sup_{x \in D} |f(x) + g(x)| \leq \sup_{x \in D} |f(x)| + \sup_{x \in D} |g(x)| = \|f\|_\infty + \|g\|_\infty.$$

□

Die Supremumsnorm ist die Standardnorm auf dem Raum der beschränkten Funktionen — wenn nichts anderes gesagt wird, betrachten wir immer die Supremumsnorm auf $B(D; \mathbb{K})$.

Kontrollfrage 2.3. Welche der nachfolgenden Aussagen sind richtig?

(A) In $(B((0, \frac{\pi}{2}); \mathbb{R}); \|\cdot\|_\infty)$ gilt $\|\sin(\cdot) - \cos(\cdot)\|_\infty = 1$.

$$\checkmark : \lim_{x \rightarrow 0} \sin(x) - \cos(x) = 1$$

(B) In $(\mathbb{R}^3, \|\cdot\|_\infty)$ gilt $\|(1, -\sqrt{6}, 3)\|_\infty = 4$

$$\times : 3 \neq 4$$

12.4.2020

#5

<https://youtu.be/8vAJ4cwfE>

12.4.2020

Wir haben auf $\mathbb{R}$ mit Hilfe des Betrages $|\cdot|$ die Metrik (den Abstand) $d_{\mathbb{R}}(x, y) := |x - y|$ definiert. Das ist auch für allgemeine normierte Räume möglich:

Lemma 1.18 (Normierte Räume sind metrische Räume). Sei $(X, \|\cdot\|)$ ein normierter Raum. Dann wird durch

$$(1.2) \quad d : X \times X \rightarrow \mathbb{R}_{\geq}, \quad d(x, y) := \|x - y\|$$

eine Metrik auf X definiert.

Beweis. Der Beweis für $(\mathbb{R}; d_{\mathbb{R}})$ mit $d_{\mathbb{R}}(x, y) := |x - y|$ überträgt sich “Eins zu eins” auf die hier betrachtete Situation. □

Definition 1.19 (Induzierte Metrik). Sei $(X, \|\cdot\|)$ normierter Raum. Die in (1.2) definierte Metrik heißt die “(durch $\|\cdot\|$) induzierte Metrik auf X ”.

Mit Hilfe der induzierten Metrik sind nun alle Konzepte für metrische Räume auch auf normierte Räume anwendbar. Beispiele:

Konzept	metrischer Raum (M, d)	normierter Raum $(X, \ \cdot\)$
Offene Kugel $B(x; r)$	$\{y \in M : d(x, y) < r\}$	$\{y \in X : \ x - y\ < r\}$ $= x + B(0; r)$.
Beschränkte Teilmenge A	$\exists x \in A, R \in \mathbb{R} : A \subseteq B(x; R)$	$\exists x \in A, R \in \mathbb{R} : A \subseteq B(x; R)$ $\Leftrightarrow \exists R \in \mathbb{R} : A \subseteq B(0; R)$
Konvergente Folge $\lim_{n \rightarrow \infty} x_n = x$	$\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n \geq n_0 : d(x_n, x) < \varepsilon$	$\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n \geq n_0 : \ x_n - x\ < \varepsilon$ Schreibweise: $\lim_{n \rightarrow \infty} x_n = x$ in $(X, \ \cdot\)$
Cauchy-Folge	$\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n, m \geq n_0 : d(x_n, x_m) < \varepsilon$	$\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n, m \geq n_0 : \ x_n - x_m\ < \varepsilon$
$f : D \rightarrow \mathbb{R}$ ist stetig bei x	$\forall \varepsilon > 0 \exists \delta > 0 \forall y \in D : d(x, y) < \delta \Rightarrow f(x) - f(y) < \varepsilon$	$\forall \varepsilon > 0 \exists \delta > 0 \forall y \in D : \ x - y\ < \delta \Rightarrow f(x) - f(y) < \varepsilon$
$f : D \rightarrow \mathbb{R}$ ist Lipschitz stetig	$\exists L \in \mathbb{R} \forall x, y \in D : f(x) - f(y) \leq L d(x, y)$	$\exists L \in \mathbb{R} \forall x, y \in D : f(x) - f(y) \leq L \ x - y\ $

Kontrollfrage 2.4. Wir betrachten $(\mathbb{R}^2, \|\cdot\|_\infty)$. Welche der nachfolgenden Aussagen ist richtig?

(A) Die Menge $\mathbb{R} \times \{0\}$ ist beschränkt.

$\times$: enthält $(n, 0)$ für alle $n \in \mathbb{N}$.

(B) Die Menge $\mathbb{R} \times [0, 1]$ ist abgeschlossen.

$\checkmark$: Produkt abgeschlossener Mengen ist abgeschlossen

(C) Seien $O_1, O_2 \subseteq \mathbb{R}$ offen. Dann ist $O_1 \times O_2$ offen.

$\checkmark$: Produkt offener Mengen ist offen

(D) Die Menge $\{x = (x_1, x_2) \in \mathbb{R}^2 : 2x_1 + x_2 = 0\}$ ist offen.

$\times$: Menge ist abgeschlossen, da $(x_1, x_2) \mapsto 2x_1 + x_2$ stetig

Kontrollfrage 2.5. Sei $I := (0, 1)$. Wir betrachten die Funktionenfolge $(f_k)_{k \in \mathbb{N}_+} \subseteq B(I; \mathbb{R})$ mit

$$f_k(x) := \begin{cases} 1 & \text{falls } x < \frac{1}{k} \\ 0 & \text{sonst} \end{cases}$$

Welche der folgenden Aussagen sind richtig?

(A) (f_k) konvergiert in $(B(I; \mathbb{R}), \|\cdot\|_\infty)$.

$\times$: Betrachte Gebiet um 0.

(B) (f_k) ist eine Cauchyfolge in $(B(I; \mathbb{R}), \|\cdot\|_\infty)$.

$\times$: divergente Folgen sind keine Cauchyfolgen

(C) (f_k) ist eine beschränkte Folge in $(B(I; \mathbb{R}), \|\cdot\|_\infty)$.

$\checkmark$: $\|f_k\|_\infty = 1$

(D) (f_k) ist in $(B(I; \mathbb{R}), \|\cdot\|_\infty)$ nicht konvergent.

$\checkmark$

Kontrollfrage 2.6. Sei $I := (0, 2\pi)$. Wir betrachten $(B(I; \mathbb{R}), \|\cdot\|_\infty)$. Welche der folgenden Aussagen sind richtig?

(A) $\sin(\cdot)$ liegt in der offenen Kugel $B(0; 1)$.

$\times$: $\|\sin\|_\infty = 1$

(B) $\cos(\cdot)$ liegt in der abgeschlossenen Kugel $K(0; 1)$.

$\checkmark$: $\|\cos\|_\infty = 1$.

Definition 1.20 (Banachraum). Ein vollständiger normierter Raum $(X, \|\cdot\|)$ heißt Banachraum.

Mit dem Cauchy-Kriterium (AN10, Satz 3.34) folgt, dass $(\mathbb{R}, |\cdot|)$ (sowie $(\mathbb{C}, |\cdot|)$) ein Banachraum ist.

Satz 1.21. $(\mathbb{K}^d, \|\cdot\|_\infty)$ ist ein Banachraum.

Wir stellen zunächst fest, dass in $\mathbb{K}^d$ Konvergenz bzgl. der Maximumsnorm und die Konvergenz der Komponenten äquivalent sind:

Lemma 1.22. Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in $\mathbb{K}^d$, $a \in \mathbb{K}^d$. Dann sind äquivalent:

- (a) $(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{K}^d$ beschränkt $\Leftrightarrow \forall i \in \{1, \dots, d\} : (x_{n,i})_{n \in \mathbb{N}} \subseteq \mathbb{K}$ beschränkt.
- (b) $(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{K}^d$ Cauchy-Folge $\Leftrightarrow \forall i \in \{1, \dots, d\} : (x_{n,i})_{n \in \mathbb{N}} \subseteq \mathbb{K}$ Cauchy-Folge.
- (c) $\lim_{n \rightarrow \infty} x_n = a$ in $(\mathbb{K}^d, \|\cdot\|_\infty)$ $\Leftrightarrow \forall i \in \{1, \dots, d\} : \lim_{n \rightarrow \infty} x_{n,i} = a_i$ in $\mathbb{K}$.

Konvention: Wir schreiben $x_{n,i}$ für die i te Komponente von $x_n \in \mathbb{K}^d$.

Beweis. Die Behauptungen folgen aus der Abschätzung:

$$(1.3) \quad |x_i| \leq \|x\|_\infty \leq \sum_{j=1}^d |x_j| \quad \text{für alle } x \in \mathbb{K}^d \text{ und } i = 1, \dots, d.$$

Als Beispiel beweisen wir (c).

Argument für $\Rightarrow$: Es gelte $x_n \rightarrow a$ in $\mathbb{K}^d$. Sei $\varepsilon > 0$. Dann existiert $n_0 \in \mathbb{N}$, so dass

$$\|x_n - a\|_\infty \leq \varepsilon \quad \text{für alle } n \geq n_0,$$

Aus (1.3) folgt nun für alle $i \in \{1, \dots, d\}$: $|x_{n,i} - a_i| \leq \varepsilon$ für alle $n \geq n_0$, also $x_{n,i} \rightarrow a_i$.

Argument für $\Leftarrow$: Es gelte $x_{n,i} \rightarrow a_i$ für alle $i \in \{1, \dots, d\}$. Sei $\varepsilon > 0$ und wähle $n_0 \in \mathbb{N}$, so dass für alle $i \in \{1, \dots, d\}$,

$$|x_{n,i} - a_i| \leq \frac{\varepsilon}{d} \quad \text{für alle } n \geq n_0.$$

Aus (1.3) folgt nun $\|x_n - a\|_\infty \leq \sum_{i=1}^d |x_{n,i} - a_i| \leq \varepsilon$ für alle $n \geq n_0$, also $x_n \rightarrow a$ in $\mathbb{K}^d$. □

Beweis von Satz 1.21. Sei $(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{K}^d$ eine Cauchy-Folge. Dann ist $(x_{n,i})_{n \in \mathbb{N}}$ Cauchy-Folge für alle $1 \leq i \leq d$ wegen Lemma 1.22 (b). $(\mathbb{K}, |\cdot|)$ ist vollständig und damit ist $(x_{n,i})_{n \in \mathbb{N}}$ konvergent in $\mathbb{K}$. Die Behauptung folgt nun aus Lemma 1.22 (c). □

Bemerkung 1.23. Später zeigen wir: Jeder endlich-dimensionale normierte Raum ist ein Banachraum. Aber: Normierte Räume mit unendlicher Dimension sind meistens nicht vollständig. Beispiele hierfür folgen später.

Lemma 1.24. Sei $(X, \|\cdot\|)$ ein Banachraum und $Y \subseteq X$ ein abgeschlossener Unterraum. Dann ist $(Y, \|\cdot\|)$ ein Banachraum.

Beweis. Nur zu zeigen: $(Y, \|\cdot\|)$ ist vollständig. Argument: Sei $(x_n)_{n \in \mathbb{N}}$ Cauchy-Folge in Y . Da X vollständig ist, gilt $x_n \rightarrow x$ für ein $x \in X$. Da Y abgeschlossen ist, folgt $x \in Y$. □

Satz 1.25 (Stetigkeit linearer Abbildungen). *Seien $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ normierte Räume und $T \in \text{Lin}(X; Y)$. Dann sind äquivalent:*

(a) $T(B(0; 1)) \subseteq Y$ ist beschränkt.

(b) T ist stetig bei 0.

(c) T ist stetig.

(d) T ist Lipschitzstetig.

Beweis. Offensichtlich gilt (d) $\Rightarrow$ (c) (siehe AN10, Satz 6.14) sowie (c) $\Rightarrow$ (b).

1. Argument für (b) $\Rightarrow$ (a): T ist stetig bei 0; also existiert $\delta > 0$ mit

$$\forall x \in X : \|x\|_X = \|x - 0\|_X < \delta \Rightarrow \|T(x)\|_Y = \|T(x) - T(0)\|_Y < 1.$$

Für $x \in B(0; 1)$ folgt nun

$$\|T(x)\|_Y = \|T(\frac{\delta}{\delta}x)\|_Y = \frac{1}{\delta} \|T(\delta x)\|_Y < \frac{1}{\delta},$$

also $T(B(0; 1)) \subseteq B(0; \frac{1}{\delta})$.

2. Argument für (a) $\Rightarrow$ (d): Da $T(B(0; 1))$ beschränkt, existiert $L \in \mathbb{R}$ mit $T(B(0; 1)) \subseteq B(0; L)$, also

$$\forall x \in X : \|x\|_X < 1 \Rightarrow \|T(x)\|_Y < L.$$

Für beliebiges $x \in X \setminus \{0\}$ folgt nun

$$\|T(x)\|_Y = 2\|x\|_X \|T(\frac{x}{2\|x\|_X})\|_Y \leq 2L\|x\|_X.$$

Die Abschätzung gilt auch für $x = 0$, denn dann sind beide Seiten gleich 0. Mit der Linearität folgt nun

$$\forall x, y \in X : \|T(x) - T(y)\|_Y = \|T(x - y)\|_Y \leq 2L\|x - y\|_X,$$

also die Lipschitzstetigkeit. □

Beispiel 1.26. Wir betrachten $\mathbb{K}^d$ und $\mathbb{K}^n$ mit der Maximumsnorm. Sei $T : \mathbb{K}^d \rightarrow \mathbb{K}^n$ eine lineare Abbildung der Form $T(x) = Ax$ mit $A \in \mathbb{K}^{n \times d}$.

Es gilt für alle $x \in \mathbb{K}^d$,

$$\|T(x)\|_\infty = \max_{1 \leq i \leq n} \left| \sum_{j=1}^d A_{ij} x_j \right| \leq \max_{1 \leq i \leq n} \sum_{j=1}^d |A_{ij}| |x_j| \leq \left(\max_{1 \leq i \leq n} \sum_{j=1}^d |A_{ij}| \right) \|x\|_\infty.$$

Damit ist $T(B(0; 1)) \subseteq \mathbb{K}^n$ beschränkt und T (Lipschitz)stetig auf ganz $\mathbb{K}^d$ (Satz 1.25). Wir werden später sehen, dass jede lineare Abbildung auf einem endlich dimensional normierten Raum stetig ist.

1.4 Konvexe Mengen und Funktionen

Ziele:

- Konzept: Konvexität
- konvexe Funktionen auf $\mathbb{R}$
- Anwendung: 3 fundamentale Ungleichungen (Young, Hölder, Minkowski)

Definition 1.27 (Verbindungsstrecke, konvexe Mengen). *Sei X ein Vektorraum.*

(i) Für $x, y \in X$ heißt die Menge

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

Verbindungsstrecke zwischen x und y .

(ii) Eine Menge $D \subseteq X$ heißt konvex, falls

$$[[x, y]] \subseteq D \text{ für alle } x, y \in D.$$

(iii) Eine Funktion $f : D \subseteq X \rightarrow \mathbb{R}$ heißt konvex, falls D konvex ist und

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \quad \text{für alle } x, y \in D \text{ und } \lambda \in [0, 1].$$

Eine Funktion $f : D \subseteq X \rightarrow \mathbb{R}$ heißt konkav, wenn $-f$ konvex ist.

Beispiele 1.28.

(a) Lineare Funktionen $T \in \text{Lin}(X; \mathbb{R})$ sind konvex und konkav.

(b) $f(x) := x^2$ ist konvex auf $\mathbb{R}$. (Beweis später)

(c) Die Kreisscheibe $K := \{x = (x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}$ ist konvex. (Beweis später) Die Kreislinie ∂K ist nicht konvex.

(d) $f(x) := x^3$ ist nicht konvex auf $\mathbb{R}$.

f ist aber konvex auf $\mathbb{R}_{\geq 0}$ und konkav auf $\mathbb{R}_{\leq 0}$. (Beweis später)

(e) $f : D \rightarrow \mathbb{R}$, $f(x) := x$, $D := (-\infty, -1) \cup (1, \infty)$ ist nicht konvex, da D nicht konvex ist.

(f) Die Betragsfunktion $|\cdot|$ ist konvex auf $\mathbb{R}$, denn aus der Dreiecksungleichung und der absoluten Homogenität folgt

$$|\lambda x + (1 - \lambda)y| \leq |\lambda x| + |(1 - \lambda)y| = \lambda|x| + (1 - \lambda)|y|$$

Lemma 1.29. Sei X ein Vektorraum, $\|\cdot\|$ eine Norm auf X und $x_0 \in X$. Dann ist die durch $f : X \rightarrow \mathbb{R}$, $f(x) := \|x - x_0\|$ definierte Funktion konvex.

Beweis. Für $x, y \in X$, $\lambda \in [0, 1]$ gilt

$$\begin{aligned} f(\lambda x + (1 - \lambda)y) &= \|\lambda x + (1 - \lambda)y - x_0\| = \|\lambda(x - x_0) + (1 - \lambda)(y - x_0)\| \\ &\leq \|\lambda(x - x_0)\| + \|(1 - \lambda)(y - x_0)\| = \lambda\|x - x_0\| + (1 - \lambda)\|y - x_0\| \\ &= \lambda f(x) + (1 - \lambda)f(y). \end{aligned}$$

□

Lemma 1.30. Sei X ein Vektorraum, $D \subseteq X$ eine konvexe Teilmenge, $f : D \rightarrow \mathbb{R}$ eine konvexe Funktion. Dann ist die Subniveaumenge

$$D_\alpha := \{x \in D : f(x) \leq \alpha\}$$

konvex für alle $\alpha \in \mathbb{R}$.

Beweis. Sei $x, y \in D_\alpha$ und $\lambda \in (0, 1)$. Dann gilt

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \leq \lambda\alpha + (1 - \lambda)\alpha = \alpha.$$

□

Beispiel 1.31.

(a) Sei $(X, \|\cdot\|)$ ein normierter Raum. Die (abgeschlossene) Kugel $K(x_0; r) := \{x \in X : \|x - x_0\| \leq r\} \subseteq X$ ist konvex, denn $K(x_0; r)$ ist Subniveaumenge der konvexen Funktion $f(x) := \|x - x_0\|$.

(b) Die Menge $\{x = (x_1, x_2) \in \mathbb{R}^2 : \|x\|_\infty \leq 1\}$ ist konvex.

Lemma 1.32. Sei X ein Vektorraum, $f : D \subseteq X \rightarrow \mathbb{R}$ konvex. Seien $x_i \in D$, $\lambda_i \in [0, 1]$ für $i \in \{1, \dots, d\}$ mit

$$(1.4) \quad \sum_{i=1}^d \lambda_i = 1.$$

Dann gilt

$$f\left(\sum_{i=1}^d \lambda_i x_i\right) \leq \sum_{i=1}^d \lambda_i f(x_i).$$

Beweis. Siehe Übung. □

Kontrollfrage 2.7. Welche der folgenden Aussagen sind für jede konvexe Menge $K \subseteq \mathbb{R}^2$ richtig?

(A) $\{(1, 0), (0, 1)\} \subseteq K \Rightarrow (\frac{1}{3}, \frac{1}{3}) \in K$.

✗: Gegenbeispiel: $K =$ Gerade durch $(1, 0), (0, 1)$

(B) $\{(1, 0), (0, 1)\} \subseteq K \Rightarrow (\frac{1}{3}, \frac{2}{3}) \in K$.

✓: $\frac{1}{3}(1, 0) + \frac{2}{3}(0, 1)$ ist Konvexkombination

(C) $\{0, (1, 0), (0, 1)\} \subseteq K \Rightarrow (1, 1) \in K$.

$\times$: Gegenbeispiel: $K =$ Das den Punkten zugeordnete Dreieck.

$$(D) \{0, (1, 0), (0, 1)\} \subseteq K \Rightarrow (\frac{1}{3}, \frac{1}{3}) \in K.$$

✓: Ist Konvexkombination der drei Punkte

Bemerkung 1.33. Eine Linearkombinationen $\sum_{i=1}^d \lambda_i x_i$ mit Koeffizienten der Form

$$\lambda_i \in [0, 1] \text{ für alle } i \in \{1, \dots, d\} \quad \text{und} \quad (1.4)$$

heißt Konvexkombination der Vektoren $x_1, \dots, x_d$. Im Fall $d = 2$ gilt

$$\lambda_1 x_1 + (1 - \lambda_1) x_2 = \lambda_1 x_1 + \lambda_2 x_2,$$

d.h. die Linearkombination in Definition 1.27 ist eine Konvexkombination.

Satz 1.34. Sei $I \subseteq \mathbb{R}$ ein offenes Intervall.

(a) Sei $f : I \rightarrow \mathbb{R}$ differenzierbar, Dann gilt

$$f \text{ monoton wachsend} \Leftrightarrow \forall x \in I : f'(x) \geq 0.$$

(b) Sei $f : I \rightarrow \mathbb{R}$ zweimal differenzierbar. Dann gilt

$$f \text{ konvex} \Leftrightarrow \forall x \in I : f''(x) \geq 0.$$

Beweis. Für (a) siehe AN10, Korollar 7.17. Wir beweisen (b).

1. Argument für $\Leftarrow$: Sei $f''(x) \geq 0$ für alle $x \in I$. Nach (a) ist dann $f' : I \rightarrow \mathbb{R}$ monoton wachsend. Seien $x_1, x_2 \in I$ und $\lambda \in (0, 1)$ mit $x_1 < x_2$ und setze $x := \lambda x_1 + (1 - \lambda)x_2$. Dann gibt es (Mittelwertsatz) $\xi_1 \in (x_1, x)$ und $\xi_2 \in (x, x_2)$ mit

$$\frac{f(x) - f(x_1)}{x - x_1} = f'(\xi_1) \leq f'(\xi_2) = \frac{f(x_2) - f(x)}{x_2 - x}.$$

Aus $x - x_1 = (1 - \lambda)(x_2 - x_1)$ und $x_2 - x = \lambda(x_2 - x_1)$ folgt

$$\frac{f(x) - f(x_1)}{1 - \lambda} \leq \frac{f(x_2) - f(x)}{\lambda}.$$

Multiplikation mit $\lambda(1 - \lambda)$ führt auf

$$\lambda(f(x) - f(x_1)) \leq (1 - \lambda)(f(x_2) - f(x)).$$

Umsortieren ergibt

$$f(x) \leq \lambda f(x_1) + (1 - \lambda)f(x_2).$$

2. Argument für $\Rightarrow$: Seien $x, y \in I$ mit $y > x$ und $\lambda \in (0, 1)$ beliebig. Dann gilt

$$f(\lambda y + (1 - \lambda)x) \leq \lambda f(y) + (1 - \lambda)f(x).$$

Aus $\lambda y + (1 - \lambda)x = x + \lambda(y - x)$ folgt

$$f(x + \lambda(y - x)) - f(x) \leq \lambda(f(y) - f(x))$$

Division mit $\lambda(y - x)$ führt auf

$$\frac{f(x + \lambda(y - x)) - f(x)}{\lambda(y - x)} \leq \frac{f(y) - f(x)}{y - x}.$$

Grenzübergang $\lambda \downarrow 0$ liefert

$$f'(x) \leq \frac{f(y) - f(x)}{y - x} \quad \text{für alle } y > x.$$

Sei nun (y_n) eine Folge in I mit $y_n > x$ und $y_n \rightarrow x$. Dann existieren (Mittelwertsatz) $\xi_n \in (x, y_n)$ mit

$$f'(\xi_n) = \frac{f(y_n) - f(x)}{y_n - x} \geq f'(x),$$

also

$$\frac{f'(\xi_n) - f'(x)}{\xi_n - x} \geq 0.$$

Grenzübergang $n \rightarrow \infty$ liefert $\xi_n \rightarrow x$ und $\frac{f'(\xi_n) - f'(x)}{\xi_n - x} \rightarrow f''(x)$, also $f''(x) \geq 0$.

□

Beispiele 1.35.

(a) $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) := x^2$ ist konvex, denn $f''(x) = 2 \geq 0$ für alle $x \in \mathbb{R}$.

(b) $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) := x^3$ ist nicht konvex, denn $f''(x) = 6x$ hat einen Vorzeichenwechsel bei 0. f ist konvex auf $\mathbb{R}_{\geq}$, denn $f'' \geq 0$ auf $\mathbb{R}_{\geq}$.

(c) $\exp : \mathbb{R} \rightarrow \mathbb{R}$ ist konvex, denn $\exp'(x) = \exp(x) \geq 0$.

Kontrollfrage 2.8. Welche der folgenden Aussagen sind richtig?

(A) $\sin : \mathbb{R} \rightarrow \mathbb{R}$ ist konvex.

$$\times : \sin''(\frac{\pi}{2}) < 0$$

(B) $\sin : [0, \pi] \rightarrow \mathbb{R}$ ist konkav.

$$\checkmark : \sin'' \geq 0 \text{ auf } [0, \pi].$$

Beispiel 1.36. Für $\ln : (0, \infty) \rightarrow \mathbb{R}$ gilt

$$\ln'(x) = \frac{1}{x}, \quad \ln''(x) = -\frac{1}{x^2},$$

also ist $f := -\ln$ eine konvex Funktion. Zusammen mit Lemma 1.32 folgt

$$(1.5) \quad -\ln\left(\sum_{i=1}^n \lambda_i x_i\right) \leq -\sum_{i=1}^n \lambda_i \ln(x_i)$$

für beliebige Konvexkombinationen $\sum_{i=1}^n \lambda_i x_i$ mit $x_1, \dots, x_n > 0$.

Wir erinnern an die Ungleichung vom arithmetischen und geometrischen Mittel:

$$\forall x_1, \dots, x_n \in \mathbb{R}_{\geq} : \sqrt[n]{x_1 x_2 \dots x_n} \leq \sum_{i=1}^n \frac{1}{n} x_i.$$

Setzen wir $\lambda_1 = \dots = \lambda_n := \frac{1}{n}$, so können wir die Abschätzung in die Form

$$\prod_{i=1}^n x_i^{\lambda_i} \leq \sum_{i=1}^n \lambda_i x_i$$

bringen. Die rechte Seite ist eine Konvexkombination. Die Ungleichung bleibt für allgemeine Konvexkombination gültig.

Satz 1.37 (Ungleichung vom gewichteten arithmetischen und geometrischen Mittel). *Seien $x_i > 0$, $\lambda_i \in [0, 1]$ für $i = 1, \dots, n$ mit $\sum_{i=1}^n \lambda_i = 1$. Dann gilt*

$$\prod_{i=1}^n x_i^{\lambda_i} \leq \sum_{i=1}^n \lambda_i x_i.$$

Beweis. Multiplikation von (1.5) mit -1 führt auf

$$\ln\left(\sum_{i=1}^n \lambda_i x_i\right) \geq \sum_{i=1}^n \lambda_i \ln(x_i)$$

Wir wenden $\exp(\cdot)$ (die Umkehrfunktion von $\ln$) auf die Abschätzung an. Wegen (a) der Monotonie, (b) der Funktionalgleichung und (c) der Potenzregel von $\exp(\cdot)$ folgt

$$\sum_{i=1}^n \lambda_i x_i \stackrel{(a)}{\geq} \exp\left(\sum_{i=1}^n \lambda_i \ln(x_i)\right) \stackrel{(b)}{=} \prod_{i=1}^n \exp(\lambda_i \ln(x_i)) \stackrel{(c)}{=} \prod_{i=1}^n \exp(\ln(x_i))^{\lambda_i}.$$

Mit $x_i = \exp(\ln(x_i))$ folgt nun die Behauptung. □

Satz 1.38 (Youngsche Ungleichung). *Seien $p, q \in (1, \infty)$ mit*

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Dann gilt für alle $x, y \geq 0$,

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}.$$

Beweis. Wir wenden Satz 1.37 mit

$$\lambda_1 = \frac{1}{p}, \quad \lambda_2 = \frac{1}{q}, \quad x_1 = x^p, \quad x_2 = y^q.$$

□

Bemerkung 1.39. Exponenten $1 \leq p, q \leq \infty$ mit

$$(1.6) \quad \frac{1}{p} + \frac{1}{q} = 1 \quad (\text{Konvention: } \frac{1}{\infty} = 0)$$

nennt man duale Exponenten. Dies ist genau dann der Fall, wenn $q = \frac{p}{p-1}$ (Konvention: $\frac{1}{0} = \infty$).

Definition 1.40 (p -Norm). Sei $1 \leq p < \infty$. Für $x \in \mathbb{K}^d$ bezeichnet

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

die p -Norm von x .

Wir werden mit Hilfe der nachfolgenden Sätze feststellen, dass $(\mathbb{K}^d, \|\cdot\|_p)$ ein Banachraum ist. Offensichtlich ist $\|\cdot\|_p : \mathbb{K}^d \rightarrow \mathbb{R}_{\geq}$

- positiv definit, denn für alle $x \in \mathbb{K}^d$ gilt

$$\|x\|_p = 0 \quad \Rightarrow \quad |x_i| = 0 \text{ für alle } i = 1, \dots, d \quad \Rightarrow \quad x_i = 0 \text{ für alle } i = 1, \dots, d.$$

- absolut homogen, denn für alle $x \in \mathbb{K}^d$ und $\lambda \in \mathbb{K}$ gilt

$$\|\lambda x\|_p^p = \sum_{i=1}^d |\lambda x_i|^p = \sum_{i=1}^d (|\lambda| |x_i|)^p = |\lambda|^p \|x\|_p^p.$$

Wir müssen noch die Dreiecksungleichung zeigen.

Satz 1.41 (Höldersche Ungleichung). Seien $p, q \in [1, \infty]$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Dann gilt

$$\sum_{i=1}^d |x_i y_i| \leq \|x\|_p \|y\|_q \quad \text{für alle } x, y \in \mathbb{K}^d.$$

Beweis. 1. Der Fall $p = 1, q = \infty$:

$$\sum_{i=1}^d |x_i y_i| \leq \sum_{i=1}^d (|x_i| \max_{1 \leq j \leq d} |y_j|) = \left(\sum_{i=1}^d |x_i| \right) \left(\max_{1 \leq j \leq d} |y_j| \right)$$

2. Der Spezialfall $1 < p, q < \infty$ mit

$$(1.7) \quad \|x\|_p = \|y\|_q = 1.$$

Aus der Youngschen Ungleichung folgt

$$|x_i y_i| \leq \frac{|x_i|^p}{p} + \frac{|y_i|^q}{q} \quad \text{für alle } i = 1, \dots, d.$$

Summation über $i = 1, \dots, d$ führt auf

$$\sum_{i=1}^d |x_i y_i| \leq \frac{1}{p} \|x\|_p^p + \frac{1}{q} \|y\|_q^q \stackrel{(1.7)}{=} \frac{1}{p} + \frac{1}{q} = 1 = \|x\|_p \|y\|_q.$$

3. Der allgemeine Fall $1 < p, q < \infty$. Wir nehmen an, dass $\|x\|_p \neq 0$ und $\|y\|_q \neq 0$ (andernfalls ist die Ungleichung trivial). Setze $\tilde{x} = \frac{x}{\|x\|_p}$ und $\tilde{y} = \frac{y}{\|y\|_q}$. Dann gilt mit Schritt 2.:

$$1 \geq \sum_{i=1}^d |\tilde{x}_i \tilde{y}_i| = \frac{1}{\|x\|_p \|y\|_q} \sum_{i=1}^d |x_i y_i|.$$

Die Höldersche Ungleichung folgt durch Multiplikation mit $\|x\|_p \|y\|_q$. □

Satz 1.42 (Minkowskische Ungleichung). *Sei $1 \leq p \leq \infty$. Dann gilt*

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p \quad \text{für alle } x, y \in \mathbb{C}^d.$$

Beweis.

1. Der Fall $p = \infty$ entspricht der Dreiecksungleichung für die Maximumsnorm und ist bereits bewiesen.

2. Im Fall $p = 1$ gilt

$$\|x + y\|_1 = \sum_{i=1}^d |x_i + y_i| \leq \sum_{i=1}^d (|x_i| + |y_i|) = \sum_{i=1}^d |x_i| + \sum_{i=1}^d |y_i|.$$

3. Der Fall $1 < p < \infty$. Es bezeichne $q := \frac{p}{p-1}$ den zu p dualen Exponenten. Mit (a) der Dreiecksungleichung für $|\cdot|$ und (b) der Hölderschen Ungleichung folgt

$$\begin{aligned} \sum_{i=1}^d |x_i + y_i|^p &= \sum_{i=1}^d |x_i + y_i| \underbrace{|x_i + y_i|^{p-1}}_{=: z_i} \stackrel{(a)}{\leq} \sum_{i=1}^d |x_i z_i| + \sum_{i=1}^d |y_i z_i| \\ &\stackrel{(b)}{\leq} (\|x\|_p + \|y\|_p) \|z\|_q. \end{aligned}$$

Weiter gilt wegen $(p-1)q = (p-1)\frac{p}{p-1} = p$

$$\|z\|_q = \left(\sum_{i=1}^d |x_i + y_i|^{(p-1)q} \right)^{\frac{1}{q}} = \left(\sum_{i=1}^d |x_i + y_i|^p \right)^{\frac{p-1}{p}} = \|x + y\|_p^{p-1}.$$

Es folgt

$$\|x + y\|_p^p \leq (\|x\|_p + \|y\|_p)\|x + y\|_p^{p-1},$$

und damit die Behauptung. □

Korollar 1.43. Sei $1 \leq p < \infty$. Mit

$$\|\cdot\|_p : \mathbb{K}^d \rightarrow \mathbb{R}_{\geq}, \quad \|x\|_p := \left(\sum_{i=1}^d |x_i|^p \right)^{\frac{1}{p}}$$

sind $(\mathbb{R}^d, \|\cdot\|_p)$ und $(\mathbb{C}^d, \|\cdot\|_p)$ normierte Räume.

Bemerkung 1.44 (Wann gilt Gleichheit?).

(a) Eine konvexe Funktion $f : D \subseteq X \rightarrow \mathbb{R}$ heißt strikt konvex, wenn sie konvex ist und für alle $x, y \in D$ und $\lambda \in (0, 1)$ zusätzlich gilt:

$$f(\lambda x + (1 - \lambda)y) = \lambda f(x) + (1 - \lambda)f(y) \Rightarrow x = y.$$

(b) Die Exponentialfunktion ist strikt konvex. Argument durch Widerspruch: Annahme es existieren $\lambda \in (0, 1)$ und $x < y$ mit

$$\exp(\lambda x + (1 - \lambda)y) = \lambda \exp(x) + (1 - \lambda) \exp(y).$$

Wir setzen $z := \lambda x + (1 - \lambda)y$. Umsortieren führt auf

$$\lambda(\exp(z) - \exp(x)) = (1 - \lambda)(\exp(y) - \exp(z))$$

Division mit $\lambda(1 - \lambda)(y - x) > 0$ führt auf

$$\frac{\exp(z) - \exp(x)}{z - x} = \frac{\exp(y) - \exp(z)}{y - z},$$

wobei wir $(1 - \lambda)(y - x) = z - x$ und $\lambda(y - x) = y - z$ benutzt haben. Aus dem Mittelwertsatz folgt die Existenz von Zwischenstellen

$$x < \xi_1 < z < \xi_2 < y$$

mit

$$\exp'(\xi_1) = \frac{\exp(z) - \exp(x)}{z - x} = \frac{\exp(y) - \exp(z)}{y - z} = \exp'(\xi_2).$$

Also gilt $\exp(\xi_1) = \exp(\xi_2)$. Da $\exp$ streng monoton ist, folgt $\xi_1 = \xi_2$ im Widerspruch zu $\xi_1 < \xi_2$.

(c) Youngsche Ungleichung: Für $x, y \geq 0$ und duale Exponenten $1 < p, q < \infty$ gilt

$$(1.8) \quad xy = \frac{x^p}{p} + \frac{y^q}{q} \quad \Leftrightarrow \quad x^p = y^q.$$

Argument für " $\Rightarrow$ ": Es gilt

$$\begin{aligned} xy &= \exp(\ln(xy)) = \exp(\ln x + \ln y) = \exp\left(\frac{1}{p} \ln(x^p) + \frac{1}{q} \ln(y^q)\right), \\ \frac{x^p}{p} + \frac{y^q}{q} &= \frac{1}{p} \exp(\ln(x^p)) + \frac{1}{q} \exp(\ln(y^q)). \end{aligned}$$

Aus der Gleichheit der linken Seiten folgt

$$\exp\left(\frac{1}{p} \ln(x^p) + \frac{1}{q} \ln(y^q)\right) = \frac{1}{p} \exp(\ln(x^p)) + \frac{1}{q} \exp(\ln(y^q)).$$

Da $\exp(\cdot)$ strikt konvex ist, folgt $\ln(x^p) = \ln(y^q)$ und damit $x^p = y^q$.

(d) Höldersche Ungleichung: Für $x, y \in \mathbb{C}^d$ und duale Exponenten $p, q \in (1, \infty)$ gilt $\sum_{i=1}^d |x_i y_i| = \|x\|_p \|y\|_q$ genau dann, wenn die Vektoren $(|x_1|^p, \dots, |x_d|^p)$ und $(|y_1|^q, \dots, |y_d|^q)$ linear abhängig sind. (Beweis siehe Übung).

(e) Minkowskische Ungleichung: Für $x, y \in \mathbb{C}^d$ und $p \in (1, \infty)$ gilt $\|x + y\|_p = \|x\|_p + \|y\|_p$ genau dann, wenn die Vektoren x, y positiv linear abhängig sind, d.h. $x = \lambda y$ mit $\lambda > 0$. (Beweis siehe Übung).

Kontrollfrage 3.1. Sei $1 < p < \infty$. Welche der folgenden Aussagen sind für alle $x, y \in \mathbb{R}^d$ wahr?

(A) $\left(\sum_{i=1}^d |x_i y_i|^2\right)^{\frac{1}{2}} \leq \|x\|_{2p} \|y\|_{\frac{2p}{p-1}}$

✓: Höldersche Ungleichung mit $(p, \frac{p}{p-1})$

(B) $\sum_{i=1}^d |x_i y_i| \leq \|x\|_p \|y\|_{\frac{p}{p-1}}$

✓: Höldersche Ungleichung mit $(p, \frac{p}{p-1})$

(C) $\sum_{i=1}^d |x_i y_i| \leq \|x\|_2 \|y\|_3$

✗: Gegenbeispiel: $x = (1, 1)$, $y = (2, 2)$. Man beachte $\|\cdot\|_3 \not\leq \|\cdot\|_2$.

Kontrollfrage 3.2. Sei $1 < p < \infty$. Welche der folgenden Aussagen sind für alle $x, y \in \mathbb{R}^d$ wahr?

(A) $\|x + y\|_p \geq \|x\|_p + \|y\|_p \quad \Rightarrow \quad x = y$

✗: Gegenbeispiel $x = 2y$

(B) $|\|x\|_p - \|y\|_p| \leq \|x + y\|_p$

✓: inverse Dreiecksungleichung

Kontrollfrage 3.3. Wir betrachten $f : [1, \infty] \rightarrow \mathbb{R}$, $f(p) = q$, wobei q den zu p dualen Exponenten bezeichnet. Welche der folgenden Aussagen sind wahr?

(A) f ist streng monoton wachsend.

$\times :$

(B) f ist streng monoton fallend.

$\checkmark : (q = p/(p-1) \text{ nach } p \text{ ableiten.})$

(C) f ist stetig.

$\checkmark$

(D) f besitzt einen Fixpunkt (d.h. $\exists p \in [1, \infty] : f(p) = p$).

$\checkmark : p = q = 2$

1.5 Endlichdimensionale normierte Räume

Ziele:

- Topologische Eigenschaften von $(\mathbb{K}^d, \|\cdot\|_p)$: Offene Mengen, Konvergenz von Folgen, Stetigkeit, Kompakte Mengen.
- Äquivalente Normen, und Äquivalenz der Normen in endlichdimensionalen Räumen.
- Erkenntnis: Die topologische Struktur endlich dimensionaler Räume ist mit der von $\mathbb{K}^d$ identisch und erbt viele Eigenschaften von $(\mathbb{K}, |\cdot|)$.

Bemerkung 1.45. • Auf $\mathbb{K}^d$ sind durch die p -Normen unendlich viele Normen definiert.

- Die Einheitskugel $\{x \in \mathbb{K}^d : \|x\|_p \leq 1\}$ in $(\mathbb{K}^d, \|\cdot\|_p)$ hat eine von p abhängige Geometrie, siehe Abbildung 1.
- Die "Länge" $\|x\|_p$ eines Vektors $x \in \mathbb{K}^d \setminus \{0\}$ ist von der Wahl des Exponenten p abhängig. Es gilt: Die Abbildung $p \mapsto \|x\|_p$ ist stetig differenzierbar in $\mathbb{R}_+$ und monoton fallend in $[1, \infty]$. Insbesondere,

$$\|x\|_1 = \lim_{p \downarrow 1} \|x\|_p, \quad \|x\|_\infty = \lim_{p \uparrow \infty} \|x\|_p.$$

Beweis siehe Übung. (Für welche $x \in \mathbb{K}^d$ ist $p \mapsto \|x\|_p$ streng monoton fallend?).

- Wir werden aber sehen, dass die Normen $\|\cdot\|_p$ auf $\mathbb{K}^d$ die gleichen offenen Mengen, konvergenten Folgen und stetigen Funktionen definieren.
- Wir werden sehen, dass dies für beliebige endlichdimensionale normierte Räume zutrifft.
- Wir werden feststellen, dass dies für unendlichdimensionale normierte Räume nicht richtig sein muss.

Definition 1.46 (Äquivalenz von Normen). Seien X ein Vektorraum. Zwei Normen $\|\cdot\|_1, \|\cdot\|_2$ auf X heißen äquivalent, falls $C > 0$ existiert mit

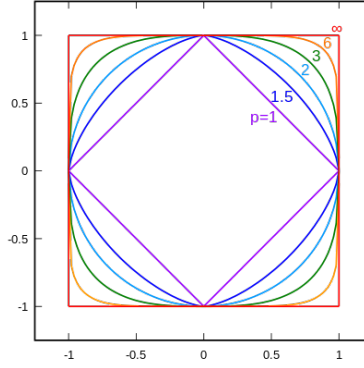
$$(1.9) \quad \frac{1}{C} \|x\|_1 \leq \|x\|_2 \leq C \|x\|_1 \quad \text{für alle } x \in X.$$

Bemerkung 1.47. Definition 1.46 ist symmetrisch:

$$(1.9) \quad \Leftrightarrow \quad \frac{1}{C} \|x\|_2 \leq \|x\|_1 \leq C \|x\|_2 \quad \text{für alle } x \in X.$$

Bemerkung 1.48. Sei X ein Vektorraum mit Normen $\|\cdot\|_1, \|\cdot\|_2$. $\|\cdot\|_1$ heißt schwächer (oder auch größer) als $\|\cdot\|_2$, falls $C > 0$ existiert mit $\|\cdot\|_1 \leq C \|\cdot\|_2$. In diesem Fall gilt:

- (a) (schwächer). Ist eine Folge $(x_n)_{n \in \mathbb{N}}$ bzgl. $\|\cdot\|_2$ beschränkt / Cauchy-Folge / konvergent, so auch bzgl. $\|\cdot\|_1$.



(from wikimedia by Quartl)

Abbildung 1: Einheitskugeln in $(\mathbb{R}^2, \|\cdot\|_p)$

$(X, \ \cdot\ _1)$	$(X, \ \cdot\ _2)$
schwächer	stärker
gröber	feiner
mehr konvergente Folgen	weniger konvergente Folgen
weniger offene Mengen	mehr offene Menge
weniger stetige Funktionen $f : X \rightarrow \mathbb{R}$	mehr stetige Funktionen $f : X \rightarrow \mathbb{R}$

Tabelle 1: Vergleich von Normen

(b) (gröber). Ist eine Menge $O \subseteq X$ offen bzgl. $\|\cdot\|_1$, so ist sie auch bzgl. $\|\cdot\|_2$ offen.

(c) Ist eine Funktion $f : X \rightarrow \mathbb{R}$ bzgl. $\|\cdot\|_1$ stetig, so auch bzgl. $\|\cdot\|_2$.

Äquivalenz der Normen bedeutet, dass $\|\cdot\|_1$ sowohl schwächer als auch stärker als $\|\cdot\|_2$ ist. In diesem Fall definieren beide Normen die gleichen offenen Mengen, Konvergenz- und Stetigkeitsbegriffe.

Lemma 1.49 (Äquivalenz der p -Normen). Seien $1 \leq p < q \leq \infty$. Dann gilt für alle $x \in \mathbb{K}^d$:

$$(1.10) \quad \|x\|_p \leq d^{\frac{1}{p} - \frac{1}{q}} \|x\|_q, \quad \|x\|_q \leq \|x\|_p.$$

Beweis. **1.** Im Fall $q = \infty$ ist die erste Abschätzung klar. Andernfalls folgt sie aus der Hölderschen Ungleichung mit Exponenten $(\frac{q}{p}, \frac{q}{q-p})$:

$$\sum_{i=1}^d |x_i|^p \leq d^{\frac{q-p}{q}} \left(\sum_{i=1}^d |x_i|^q \right)^{\frac{p}{q}}.$$

2. Für die zweite Abschätzung zeigen wir zunächst:

$$(1.11) \quad \|x\|_p = 1 \quad \Rightarrow \quad \|x\|_q \leq 1.$$

Argument: Aus $1 = \|x\|_p^p = \sum_{i=1}^d |x_i|^p$ folgt $|x_i| \leq 1$ für alle $i = 1, \dots, d$ und damit

$$|x_i|^q = \underbrace{|x_i|^{q-p}}_{\leq 1} |x_i|^p \leq |x_i|^p.$$

Es folgt

$$\left(\sum_{i=1}^d |x_i|^q\right)^{\frac{1}{q}} \leq \left(\sum_{i=1}^d |x_i|^p\right)^{\frac{1}{p}} = \left(\sum_{i=1}^d |x_i|^p\right)^{\frac{1}{p}} \underbrace{\left(\sum_{i=1}^d |x_i|^p\right)^{\frac{1}{q} - \frac{1}{p}}}_{=1} = 1.$$

- 3.** Wir beweisen nun die $\|\cdot\|_q$ - $\|\cdot\|_p$ -Abschätzung. Im Fall $x = 0$ ist nichts zu zeigen. Sei $x \neq 0$ also $\|x\|_p > 0$. Dann folgt mit (1.11)

$$\|x\|_q = \|x\|_p \left\| \frac{x}{\|x\|_p} \right\|_q \leq \|x\|_p.$$

□

Korollar 1.50. Für alle $1 \leq p \leq \infty$ ist $(\mathbb{K}^d, \|\cdot\|_p)$ ein Banachraum.

Beweis. Wir wissen bereits, dass $(\mathbb{K}^d, \|\cdot\|_\infty)$ ein Banachraum ist. Die Behauptung folgt nun aus der Äquivalenz der Normen $\|\cdot\|_p$ und $\|\cdot\|_\infty$. □

Bemerkung 1.51. In $\mathbb{K}^d$ sind Konvergenz bzgl. der p -Norm und komponentenweise Konvergenz äquivalent.

Kontrollfrage 3.4. Welche der nachfolgenden Funktionen sind nicht strikt konvex?

(A) $f: \mathbb{R}^d \rightarrow \mathbb{R}, f(x) := \|x\|_\infty.$

✓: (Linearkombinationen der Eckpunkte der Einheitskugeln betrachten!)

(B) $f: \mathbb{R}^d \rightarrow \mathbb{R}, f(x) := \|x\|_1.$

✓: (Ebenso wie in (A) nur konvex.)

(C) $f: \mathbb{R} \rightarrow \mathbb{R}, f(x) = \exp(x).$

×: (Zweite Ableitung positiv.)

(D) $f: \mathbb{R}^3 \rightarrow \mathbb{R}, f(x) := 3x_1 + 2x_2 + x_3.$

✓: (linear $\Rightarrow$ konkav und konvex)

Hinweis: Betrachten Sie Abbildung 1.

Kontrollfrage 3.5. Sei $(X, \|\cdot\|)$ ein normierter Raum. Sei $\|\cdot\|_*$ eine Norm auf X , die schwächer als $\|\cdot\|$ ist. Wir bezeichnen die Kugeln um 0 mit

$$B_{\|\cdot\|}(0; r) := \{x \in X : \|x\| < r\}, \quad B_{\|\cdot\|_*}(0; r) := \{x \in X : \|x\|_* < r\}.$$

Welche der folgenden Aussagen sind dann wahr?

(A) Für alle $r > 0$ gilt $B_{\|\cdot\|_*}(0; r) \subseteq B_{\|\cdot\|}(0; r).$

×: (Konstante C aus Bem. 1.48 kann beliebig groß oder klein sein!)

(B) Es existiert $\delta > 0$, so dass $B_{\|\cdot\|_*}(0; \delta r) \subseteq B_{\|\cdot\|}(0; r)$ für alle $r > 0$. -Richtig (siehe Def. von „schwächer“.)

(C) Es existiert $C > 0$, so dass $B_{\|\cdot\|}(0; r) \subseteq B_{\|\cdot\|_*}(0; Cr)$ für alle $r > 0$.

×:

(D) $(X, \|\cdot\|)$ vollständig $\Rightarrow (X, \|\cdot\|_*)$ vollständig.

×: (Bzgl. $\|\cdot\|_*$ gibt es i.A. mehr Cauchy-Folgen, die bei Überprüfung der Vollständigkeit betrachtet werden müssen.)

Kontrollfrage 3.6. Sei $(X, \|\cdot\|)$ normierter Raum und $f : X \rightarrow \mathbb{R}$. Welche der folgenden Aussagen sind äquivalent zur Aussage “ f ist stetig”?

(A) $\forall \varepsilon > 0 \exists \delta > 0 \forall x, y \in X : (\|x - y\| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon)$.

×: (Das ist sogar gleichmäßige Stetigkeit.)

(B) $f(x_n) \rightarrow f(x)$ für jede konvergente Folge $(x_n) \subseteq X$ mit Grenzwert x .

✓: (Def. 6.6, AN10)

(C) $f(O)$ ist offen für jede offene Menge $O \subseteq X$.

×: (Gegenbeispiel: konstante Fkt.)

(D) $f^{-1}(O)$ ist offen für jede offene Menge $O \subseteq \mathbb{R}$.

✓: (siehe Bem. 1.61.)

Kontrollfrage 3.7. Sei $(x_n) \subseteq \mathbb{R}^d$ eine Folge. Für $i = 1, \dots, d$ seien die Komponentenfolgen $(x_{n,i})_{n \in \mathbb{N}} \subseteq \mathbb{R}$ Cauchy-Folgen in $\mathbb{R}$. Welche der folgenden Aussagen sind richtig?

(A) Die Folge konvergiert in $(\mathbb{R}^d, \|\cdot\|_\infty)$.

✓: (Lemma 1.22 und Vollständigkeit.)

(B) Die Folge konvergiert in $(\mathbb{R}^d, \|\cdot\|_{\sqrt{2}})$.

✓: (Bem. 1.51.)

(C) $\lim_{n \rightarrow \infty} \|x_n\|_{\sqrt{2}} \leq \lim_{n \rightarrow \infty} \|x_n\|_\infty$

×: (Lemma 1.49.)

Definition 1.52 (Kompakte Teilmenge). Sei (M, d) ein metrischer Raum. Eine Teilmenge $K \subseteq M$ heißt kompakt, falls jede Folge $(x_n)_{n \in \mathbb{N}} \subseteq K$ eine in K konvergente Teilfolge besitzt.

Bemerkung 1.53.

- (a) Steng genommen wird in Definition 1.52 der Begriff der Folgenkompaktheit definiert.
- (b) Wir werden am Ende der Vorlesung die Struktur des metrischen Raumes verallgemeinern und topologische Räume einführen. Diese basieren auf dem Begriff der offenen Menge. Jeder metrische Raum (und damit auch jeder normierte Raum) ist ein topologischer Raum. In topologischen Räumen wird Kompaktheit mit Hilfe der sogenannten Überdeckungseigenschaft eingeführt. Es stellt sich heraus, dass sich in allgemeinen topologischen Räumen diese “Überdeckungs”-Kompaktheit von der in Definition 1.52 eingeführten Folgenkompaktheit unterscheiden können. In metrischen Räumen sind jedoch beide Begriffe äquivalent.
- (c) In AN10 haben wir Mengen $D \subseteq \mathbb{R}$, die abgeschlossen und beschränkt sind, kompakt genannt. Das nachfolgende Lemma 1.55 beinhaltet diese Situation als Spezialfall und rechtfertigt die Begriffseinführung aus AN10. Obacht: In unendlichdimensionalen normierten Räumen existieren abgeschlossene und beschränkte Mengen, die nicht kompakt sind!

Lemma 1.54. Sei K eine kompakte Teilmenge in $(X, \|\cdot\|)$. Dann ist K abgeschlossen und beschränkt.

Beweis. Beweis durch Kontraposition.

1. Annahme K ist nicht beschränkt. Dann existiert kein $n \in \mathbb{N}_+$ mit $K \subseteq B(0; n)$, d.h.

$$\forall n \in \mathbb{N}_+ \exists x_n \in K : \|x_n\| \geq n.$$

Wir folgern: (x_n) und jede ihrer Teilfolgen sind unbeschränkt. Damit kann sie keine konvergente Teilfolge besitzen.

2. Annahme K ist nicht abgeschlossen. Dann existiert eine konvergente Folge $(x_n) \subseteq K$ mit Grenzwert $x \in X \setminus K$. Diese Folge kann keine in K konvergente Teilfolge besitzen.

□

Lemma 1.55. Eine Teilmenge $K \subseteq \mathbb{K}^d$ ist genau dann kompakt in $(\mathbb{K}^d, \|\cdot\|_\infty)$, wenn sie abgeschlossen und beschränkt ist.

Beweis. Die Richtung $\Rightarrow$ folgt aus Lemma 1.54. Wir zeigen $\Leftarrow$. Sei $K \subseteq \mathbb{K}^d$ abgeschlossen und beschränkt und $(x_n) \subseteq K$ eine Folge. Wegen $|x_{n,i}| \leq \|x_n\|_\infty \leq \sup_{k \in \mathbb{N}} \|x_k\|_\infty < \infty$ sind die Komponentenfolgen $(x_{n,i})_{n \in \mathbb{N}}$ (und deren Teilfolgen) beschränkt in $\mathbb{K}$. Nach dem Satz von Bolzano-Weierstrass besitzt die Folge der ersten Komponenten $(x_{n,1})_{n \in \mathbb{N}}$ eine in $\mathbb{K}$ konvergente Teilfolge $(x_{n_k,1})_{k \in \mathbb{N}}$. Durch nochmaliges Anwenden des Satzes von Bolzano-Weierstrass auf die Teilfolge der zweiten Komponenten $(x_{n_k,2})_{k \in \mathbb{N}}$ können wir zu einer weiteren Teilfolge $(x_{n_{k_\ell},2})_{\ell \in \mathbb{N}}$ übergehen, die in $\mathbb{K}$ konvergiert. Entlang dieser zweiten Teilfolge sind sowohl die erste als auch die zweite

Komponente in $\mathbb{K}$ konvergent. Durch Iteration des Argumentes folgern wir, dass $y \in \mathbb{K}^d$ und eine Teilfolge $(n_m)_{m \in \mathbb{N}}$ existieren mit $\lim_{m \rightarrow \infty} x_{n_m, i} = y_i$ für alle $i = 1, \dots, d$. Aus Lemma 1.22 folgt nun

$$\lim_{m \rightarrow \infty} x_{n_m} = y \quad \text{in } (\mathbb{K}^d, \|\cdot\|_\infty).$$

Da K abgeschlossen ist, folgt $y \in K$. □

Wir werden im Laufe dieses Abschnittes sehen, dass diese Aussage für beliebige normierte Räume mit endlicher Dimension gilt.

Beispiel 1.56. Die $\|\cdot\|_\infty$ -Einheitssphäre $S := \{x \in \mathbb{K}^d : \|x\|_\infty = 1\}$ ist kompakt in $(\mathbb{K}^d, \|\cdot\|_\infty)$. Argument: Wegen Lemma 1.55 genügt es zu zeigen, dass S abgeschlossen und beschränkt ist. S ist offensichtlich beschränkt. Ausserdem ist S abgeschlossen, da die Funktion $\|\cdot\|_\infty : \mathbb{K}^d \rightarrow \mathbb{R}$ stetig ist. (Warum ist die Funktion stetig?)

Satz 1.57 (Satz vom Minimum und Maximum). Sei $(X, \|\cdot\|)$ ein normierter Raum, $K \subseteq X$ nichtleer und kompakt und $f : K \rightarrow \mathbb{R}$ stetig. Dann existieren $x_*, x^* \in K$ mit

$$f(x_*) = \inf_{x \in K} f(x) \quad \text{und} \quad f(x^*) = \sup_{x \in K} f(x);$$

Darüber hinaus ist $\text{Bild}(f) \subseteq \mathbb{R}$ beschränkt.

Beweis. Der Beweis ist analog zum Beweis von AN10, Satz 6.21. Wir diskutieren ihn in der Übung. □

Beispiel 1.58. Wir betrachten $(\mathbb{K}^d, \|\cdot\|_\infty)$. Es bezeichne S die $\|\cdot\|_\infty$ -Einheitssphäre aus Beispiel 1.56. Sei $f : S \subseteq \mathbb{K}^d \rightarrow \mathbb{R}$ stetig und mit $f(x) > 0$ für alle $x \in S$. Dann gilt

$$\inf_{x \in S} f(x) > 0.$$

Argument: Wir wissen bereits, dass S kompakt ist. Wegen Satz 1.57 existiert eine Minimumsstelle $x_* \in S$ mit $f(x_*) = \inf_{x \in S} f(x)$. Da f strikt positiv auf S ist, gilt $f(x_*) > 0$ und die Behauptung folgt.

Satz 1.59 (Äquivalenz der Normen auf endlichdimensionalen Räumen). *Auf einem Vektorraum X mit endlicher Dimension sind alle Normen äquivalent.*

Beweis. 1. Der Fall $X = \mathbb{K}^d$. Es genügt zu zeigen, dass jede Norm $\|\cdot\|$ auf $\mathbb{K}^d$ äquivalent zu $\|\cdot\|_\infty$ ist. Zunächst gilt für alle $x \in \mathbb{K}^d$,

$$(1.12) \quad \|x\| \leq \sum_{i=1}^d |x_i| \|e_i\| \leq \left(\sum_{i=1}^d \|e_i\| \right) \|x\|_\infty.$$

Es verbleibt zu zeigen, dass $\|\cdot\|_\infty \leq C_2 \|\cdot\|$ für ein $C_2 > 0$. Zunächst folgt aus (1.12), dass $\|\cdot\|$ (Lipschitz) stetig auf $(\mathbb{K}^d, \|\cdot\|_\infty)$ ist. Mit der Aussage aus Beispiel 1.58 folgt nun, dass

$$c := \inf_{x \in S} \|x\| > 0, \quad S := \{x \in \mathbb{K}^d : \|x\|_\infty = 1\}.$$

Also gilt für alle $x \in \mathbb{K}^d \setminus \{0\}$

$$\|x\| = \|x\|_\infty \left\| \frac{x}{\|x\|_\infty} \right\| \geq c \|x\|_\infty.$$

Da $c > 0$ folgt hieraus $\|x\|_\infty \leq \frac{1}{c} \|x\|$ und damit die Äquivalenz der beiden Normen.

- 2.** Der Fall eines $\mathbb{K}$ -Vektorraumes X mit endlicher Dimension d . Wir führen diesen auf $\mathbb{K}^d$ zurück. Seien $\|\cdot\|_1$ und $\|\cdot\|_2$ Normen auf X . Es bezeichne $T : \mathbb{K}^d \rightarrow X$ eine lineare Bijektion. Wir setzen

$$\|x\|_{i,T} := \|T(x)\|_i \quad \text{für } i = 1, 2.$$

Man prüft nach, dass hierdurch Normen $\|\cdot\|_{i,T}$, $i = 1, 2$, auf $\mathbb{K}^d$ definiert werden. Wegen Schritt 1 existiert $C > 0$ mit

$$\frac{1}{C} \|x\|_{1,T} \leq \|x\|_{2,T} \leq C \|x\|_{1,T} \quad \text{für alle } x \in \mathbb{K}^d.$$

Da T surjektiv ist, folgt

$$\frac{1}{C} \|y\|_1 \leq \|y\|_2 \leq C \|y\|_1 \quad \text{für alle } y \in \mathbb{K}^d.$$

□

Definition 1.60 (Isomorphismus). *Seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ normierte Räume. Eine Abbildung $T : X \rightarrow Y$ heißt*

- (a) Isomorphismus, falls T eine lineare Bijektion ist und sowohl T als auch T^{-1} stetig sind. In diesem Fall nennt man die Räume isomorph.
- (b) isometrischer Isomorphismus, falls T ein Isomorphismus ist und $\|T(x)\|_Y = \|x\|_X$ für alle $x \in X$ gilt.

Eine stetige Bijektion zwischen zwei metrischen Räumen, deren Umkehrabbildung auch stetig ist heißt Homöomorphismus. Man nennt in diesem Fall die metrischen Räume homöomorph.

Bemerkung 1.61. Seien $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ normierte Räume und $T : X \rightarrow Y$ eine Abbildung.

- (a) Es gilt: *isometrischer Isomorphismus* $\Rightarrow$ *Isomorphismus* $\Rightarrow$ *Homöomorphismus*. Also sind isomorphe Räume insbesondere homöomorph.
- (b) Aus AN10 wissen wir bereits, dass Funktionen (zwischen metrischen Räumen) genau dann stetig sind, wenn die Urbilder offener Mengen wieder offen sind. Ist also T ein Homöomorphismus, so gilt

$$\{O \subseteq Y : O \text{ ist offen}\} = \{T(Q) : Q \subseteq X, Q \text{ ist offen}\}.$$

Das gleiche gilt für abgeschlossene Mengen.

- (c) Wir erinnern an folgende Charakterisierung der Folgenkonvergenz im metrischen Raum (M, d) : Seien $(x_n) \subseteq M$ und $x \in M$. Dann gilt $x_n \rightarrow x$ genau dann, wenn für jede offene Menge $O \subseteq M$ gilt: $x \in O \Rightarrow x_n \in O$ für alle bis auf endliche viele $n \in \mathbb{N}$.
- (d) Als Konsequenz von (b) und (c) gilt: Seien $x, x_n \in X$ und $y_n := T(x_n)$, $y := T(x)$ für $n \in \mathbb{N}$. Ist T ein Homöomorphismus, dann gilt

$$x_n \rightarrow x \text{ in } X \quad \Leftrightarrow \quad y_n \rightarrow y \text{ in } Y$$

- (e) Aus der Folgencharakterisierung der Stetigkeit folgt nun: Sei (M, d) ein metrischer Raum und T ein Homöomorphismus, dann

$$f : X \rightarrow M \text{ ist stetig} \Leftrightarrow (f \circ T^{-1}) : Y \rightarrow M \text{ ist stetig.}$$

Homöomorphe Räume besitzen die selbe Struktur offener Mengen und den selben Konvergenz- und Stetigkeitsbegriff. Sie können jedoch sehr verschiedene geometrische Formen besitzen. Isometrisch isomorphe Räume besitzen hingegen auch die selbe Geometrie.

Beispiele 1.62. Wir betrachten $\mathbb{R}^d$ mit beliebiger Norm $\|\cdot\|$. Die $\|\cdot\|$ -Kugel $B(0; 1) \subseteq \mathbb{R}^d$ (mit der von $(\mathbb{R}^d, \|\cdot\|)$ geerbten Metrik) ist homöomorph zum $\mathbb{R}^d$. Ein Homöomorphismus ist gegeben durch $f : B(0; 1) \rightarrow \mathbb{R}^d$,

$$f(x) := \frac{x}{1 - \|x\|}, \quad f^{-1}(x) := \frac{x}{1 + \|x\|}.$$

(Zeigen Sie: Die Einheitskugeln in $\mathbb{R}^d$ zu beliebigen Normen sind untereinander homöomorph.)

Beispiele 1.63 (Isomorphismen).

- (a) Sei $T \in \text{Lin}(\mathbb{K}^d; \mathbb{K}^d)$ bijektiv. Beispiel 1.26 zeigt, dass T ein Isomorphismus von $(\mathbb{K}^d, \|\cdot\|_\infty)$ nach $(\mathbb{K}^d, \|\cdot\|_\infty)$ ist. Wegen der Äquivalenz aller Normen ist T auch dann ein Isomorphismus, wenn wir im Bild- und Urbildraum die Supremumsnorm durch beliebige Normen auf $\mathbb{K}^d$ austauschen, z.B. ist $T : (\mathbb{K}^d, \|\cdot\|_1) \rightarrow (\mathbb{K}^d, \|\cdot\|_\infty)$ ein Isomorphismus.
- (b) Sei $\sigma : \{1, \dots, d\} \rightarrow \{1, \dots, d\}$ bijektiv (d.h. eine Permutation). Wir bezeichnen mit $T_\sigma \in \text{Lin}(\mathbb{K}^d; \mathbb{K}^d)$ die eindeutige lineare Abbildung mit $T(e_i) = e_{\sigma(i)}$ für alle $i = 1, \dots, d$. T ist eine lineare Bijektion. Nach (a) ist T ein Isomorphismus (unabhängig von der Wahl der Normen auf $\mathbb{K}^d$). Im Falle von p -Normen gilt sogar:

$$T : (\mathbb{K}^d, \|\cdot\|_p) \rightarrow (\mathbb{K}^d, \|\cdot\|_p) \quad \text{isometrischer Isomorphismus}$$

(c) Wie im Beweis von Satz 1.59 sei $T \in \text{Lin}(\mathbb{K}^d; X)$ eine Bijektion, $(X, \|\cdot\|)$ normierter Raum und durch $\|x\|_T := \|T(x)\|$ eine Norm auf $\mathbb{K}^d$ definiert. Dann ist $T : (\mathbb{K}^d, \|\cdot\|_T) \rightarrow (X, \|\cdot\|)$ ein isometrischer Isomorphismus.

(d) Auf dem Vektorraum X seien zwei Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ gegeben. Es bezeichne $I : X \rightarrow X$, $I(x) := x$, die Identitätsabbildung. Dann gilt:

$$I : (X, \|\cdot\|_1) \rightarrow (X, \|\cdot\|_2) \text{ ist Isomorphismus} \quad \Leftrightarrow \quad \|\cdot\|_1, \|\cdot\|_2 \text{ sind äquivalent.}$$

Satz 1.64 (Isomorphie normierter Vektorräume endlicher Dimension). *Normierte Vektorräume mit endlicher Dimension sind genau dann isomorph, wenn sie die selbe Dimension besitzen.*

Beweis. Wir müssen nur die Richtung $\Leftarrow$ beweisen. Seien $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ Räume mit gleicher Dimension $d \in \mathbb{N}_+$. Dann existiert eine lineare Bijektion $T : X \rightarrow Y$. Durch $\|x\|_T := \|T(x)\|_Y$ wird eine Norm auf X definiert und es gilt

$$\|T(x)\|_Y = \|x\|_T, \quad \|T^{-1}(y)\|_T = \|T(T^{-1}(y))\|_Y = \|y\|_Y \quad \text{für alle } x \in X \text{ und } y \in Y.$$

Mit Satz 1.59 sind die Normen $\|\cdot\|_X$ und $\|\cdot\|_T$ äquivalent. Also gibt es $C > 0$ mit

$$\|T(x)\|_Y \leq C\|x\|_X, \quad \|T^{-1}(y)\|_X \leq C\|y\|_Y \quad \text{für alle } x \in X \text{ und } y \in Y.$$

Aus Satz 1.25 folgt nun, dass sowohl $T : (X, \|\cdot\|_X) \rightarrow (Y, \|\cdot\|_Y)$ als auch die Umkehrabbildung T^{-1} stetig sind. Damit ist T ein Isomorphismus. $\square$

Satz 1.65 (Vollständigkeit und Kompaktheit in endlichdimensionalen Räumen). *Seien $(X, \|\cdot\|)$ ein normierter $\mathbb{K}$ -Vektorraum mit endlicher Dimension $d \in \mathbb{N}_+$. Dann gilt:*

(a) Für alle $1 \leq p \leq \infty$ ist $(X, \|\cdot\|)$ zu $(\mathbb{K}^d, \|\cdot\|_p)$ isomorph.

(b) Sei Y ein weiterer normierter $\mathbb{K}$ -Vektorraum. Jede lineare Abbildung $T : X \rightarrow Y$ ist stetig.

(c) $(X, \|\cdot\|)$ ist ein Banachraum.

(d) $K \subseteq X$ ist genau dann kompakt, wenn K abgeschlossen und beschränkt ist.

Beweis. Alle Aussagen sind im Spezialfall $X = \mathbb{K}^d$ und $Y = \mathbb{K}^m$ bereits bewiesen. Sie übertragen sich durch Anwendung von Satz 1.64 auf den Fall allgemeiner normierter Räume mit endlicher Dimension. Siehe Übung. $\square$

Die Aussagen gelten allgemein nicht im unendlichdimensionalen Fall:

Satz 1.66 (Kompaktheit der abgeschlossenen Einheitskugel). *Sei $(X, \|\cdot\|)$ ein normierter Raum. Dann gilt:*

$$K(0; 1) := \{x \in X : \|x\| \leq 1\} \text{ ist kompakt} \quad \Leftrightarrow \quad \dim(X) < \infty.$$

Beweis. Die Richtung $\Leftarrow$ folgt aus Satz 1.65 (d). Die interessantere Richtung $\Rightarrow$ beweisen wir nicht in dieser Vorlesung. $\square$

Kontrollfrage 3.8. Sei $(X, \|\cdot\|)$ ein normierter Raum und $K \subseteq X$ eine Teilmenge. Welche der folgenden Aussagen sind wahr?

(A) K ist genau dann kompakt, wenn K abgeschlossen und beschränkt ist.

×: (Bem. 1.53 c.)

(B) Ist K kompakt, dann ist K abgeschlossen und beschränkt.

✓: (Lemma 1.54.)

(C) Die lineare Hülle von K sei endlichdimensional. Dann gilt: K ist genau dann kompakt, wenn K abgeschlossen und beschränkt ist.

✓: (Bemerkung 1.53 c.)

Kontrollfrage 4.1. Sei X ein Vektorraum, $T : X \rightarrow \mathbb{K}^d$ eine surjektive, lineare Abbildung. Für $x \in X$ setzen wir

$$\|x\|_T := \|T(x)\|_2.$$

Welche der folgenden Aussagen sind wahr?

(A) $\|\cdot\|_T$ definiert eine Norm auf X .

×: Im Allgemeinen gilt nicht $T(x) = 0 \Rightarrow x = 0$

(B) $\|\cdot\|_T$ erfüllt die Dreiecksungleichung.

✓: Linearität von T , Dreiecksungleichung von $\|\cdot\|_2$

(C) $\|\cdot\|_T$ ist genau dann eine Norm auf X , wenn $\dim(X) = d$.

✓: „ $\Leftarrow$ “: Aus Dimensionsgleichheit, Linearität und Surjektivität folgt Bijektivität von T und damit dann positive Definitheit von $\|\cdot\|_T$. „ $\Rightarrow$ “: Aus positiver Definitheit von $\|\cdot\|_T$ und Surjektivität folgt Bijektivität von T und mit Linearität dann Dimensionsgleichheit.

Kontrollfrage 4.2. Seien X ein Vektorraum und $\|\cdot\|, \|\cdot\|_*$ Normen auf X . Welche der folgenden Aussagen sind äquivalent zur Aussage „ $\|\cdot\|$ ist äquivalent zu $\|\cdot\|_*$ “?

(A) Es existieren $c, C > 0$, so dass $\|x\| \leq C\|x\|_*$ und $\|x\|_* \leq c\|x\|$ für alle $x \in X$.

✓: Vgl. Definition 1.46

(B) Die Abbildung $T : (X, \|\cdot\|) \rightarrow (X, \|\cdot\|_*)$, $T(x) := x$ ist stetig.

×: Impliziert nur, dass $\|\cdot\|_*$ schwächer als $\|\cdot\|$ ist.

(C) Die Abbildung $T : (X, \|\cdot\|) \rightarrow (X, \|\cdot\|_*)$, $T(x) := x$ ist ein isometrischer Isomorphismus.

×: Nicht notwendig; z.B. $\|\cdot\|_* = 3\|\cdot\|$.

(D) Die Abbildung $T : (X, \|\cdot\|) \rightarrow (X, \|\cdot\|_*)$, $T(x) := x$ ist ein Isomorphismus.

✓: Vgl. Beispiel 1.63 (d)

1.6 Hilberträume

Ziele:

- Die Geometrie des Euklidischen Raumes $(\mathbb{R}^d, \|\cdot\|_2)$.
- Das Standardskalarprodukt auf $\mathbb{R}^d$
- $\mathbb{R}$ -Hilberträume

In diesem Kapitel betrachten wir nur $\mathbb{R}$ -Vektorräume.

Wir erinnern an folgende Definitionen aus der Linearen Algebra:

- Eine Matrix $Q \in \mathbb{R}^{d \times d}$ heißt orthogonal, falls $Q^{-1} = Q^\top$. Hierbei bezeichnet $Q^\top$ die zu Q transponierte Matrix. Wir bezeichnen mit $O(d)$ die Menge aller orthogonalen Matrizen in $\mathbb{R}^{d \times d}$. $O(d)$ bildet mit dem Matrixprodukt eine Gruppe.
- Eine Abbildung der Form

$$T : \mathbb{R}^d \rightarrow \mathbb{R}^d, \quad T(x) := Qx + c$$

mit $Q \in O(d)$ und $c \in \mathbb{R}^d$ heißt Euklidische Transformation.

Euklidische Transformationen sind Kompositionen aus Translationen, Spiegelungen und Rotation*. Sie erhalten Längen und Winkel und bilden Geraden auf Geraden ab.

Lemma 1.67. $\|\cdot\|_2$ ist invariant unter Euklidischen Transformationen auf $\mathbb{R}^d$, d.h. für jede Euklidische Transformation $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$ gilt:

$$\|T(x) - T(y)\|_2 = \|x - y\|_2 \quad \text{für alle } x, y \in \mathbb{R}^d.$$

Beweis. Sei $Q \in O(d)$ und $z \in \mathbb{R}^d$. Dann gilt

$$\begin{aligned} \|Qz\|_2^2 &= \sum_{i=1}^d |(Qz)_i|^2 = \sum_{i=1}^d \left(\sum_{j=1}^d Q_{ij} z_j \right) \left(\sum_{k=1}^d Q_{ik} z_k \right) \\ &= \sum_{j,k=1}^d \underbrace{\left(\sum_{i=1}^d Q_{ij} Q_{ik} \right)}_{=(Q^\top Q)_{jk}} z_j z_k = \sum_{j=1}^d z_j z_j = \|z\|_2^2. \end{aligned}$$

$= (Q^\top Q)_{jk} \begin{cases} 1 & \text{falls } j = k, \\ 0 & \text{sonst.} \end{cases}$

Sei nun $T(z) = Qz + c$ eine euklidische Transformation. Dann gilt für $x, y \in \mathbb{R}^d$,

$$\|T(x) - T(y)\|_2 = \|Q(x - y)\|_2 = \|x - y\|_2.$$

□

Lemma 1.68. Sei $\|\cdot\|$ eine Norm auf $\mathbb{R}^d$, die invariant unter Euklidischen Transformationen ist. Dann existiert $\lambda > 0$, sodass $\|x\| = \lambda\|x\|_2$ für alle $x \in \mathbb{R}^d$.

Beweis. Für den Beweis benutzen wir folgenden Fakt aus der linearen Algebra: Für jeden Vektor $e \in \mathbb{R}^d$ mit $\|e\|_2 = 1$ existiert eine orthogonale Matrix $Q \in O(d)$ mit $e = Qe_1$. Hieraus folgern wir, dass für alle $x \in \mathbb{R}^d$ eine orthogonale Matrix $Q \in O(d)$ existiert mit $x = \|x\|_2 Qe_1$. Aus der Invarianz von $\|\cdot\|$ folgt nun

$$\|x\| = \|(\|x\|_2 Qe_1)\| = \|x\|_2 \|Qe_1\| = \|x\|_2 \|e_1\| = \lambda\|x\|_2,$$

mit $\lambda := \|e_1\| > 0$. □

Aufgrund dieses Alleinstellungsmerkmals wird die $\|\cdot\|_2$ -Norm auf $\mathbb{R}^d$ auch als die Euklidische Norm bezeichnet und $(\mathbb{R}^d, \|\cdot\|_2)$ wird auch der Euklidische Vektorraum $\mathbb{R}^d$ genannt.

Lemma 1.69 (Standardskalarprodukt). Die Abbildung

$$\cdot : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}, \quad (x, y) \mapsto x \cdot y := \sum_{i=1}^d x_i y_i$$

heißt Standardskalarprodukt und erfüllt folgende Eigenschaften:

- (a) (Bilinearität). $(x, y) \mapsto x \cdot y$ ist bilinear, d.h. für alle $x \in \mathbb{R}^d$ sind die Abbildungen $y \mapsto x \cdot y$ und $y \mapsto y \cdot x$ lineare Abbildungen von $\mathbb{R}^d$ nach $\mathbb{R}$.
- (b) (Symmetrie). Es gilt $x \cdot y = y \cdot x$ für alle $x, y \in \mathbb{R}^d$.
- (c) (Positive Definitheit). Für alle $x \in \mathbb{R}^d$ gilt $x \cdot x \geq 0$ sowie $x \cdot x = 0 \Leftrightarrow x = 0$.

Darüberhinaus gelten

$$\|x\|_2 = \sqrt{x \cdot x} \quad \text{für alle } x \in \mathbb{R}^d,$$

sowie die Cauchy-Schwarz Ungleichung

$$|x \cdot y| \leq \|x\|_2 \|y\|_2 \quad \text{für alle } x, y \in \mathbb{R}^d.$$

Beweis. Der Beweis von (a) - (c) ist offensichtlich. Die Cauchy-Schwarz Ungleichung folgt aus der Dreiecks- und Hölderschen Ungleichung (mit Exponenten (2, 2)):

$$|x \cdot y| = \left| \sum_{i=1}^d x_i y_i \right| \leq \sum_{i=1}^d |x_i y_i| \leq \|x\|_2 \|y\|_2.$$

□

Lemma 1.70 (Polarisationsformel). Für alle $x, y \in \mathbb{R}^d$ gilt

$$\begin{aligned} \|x + y\|_2^2 &= \|x\|_2^2 + 2x \cdot y + \|y\|_2^2, \\ \|x - y\|_2^2 &= \|x\|_2^2 - 2x \cdot y + \|y\|_2^2, \\ \frac{1}{4}(\|x + y\|_2^2 - \|x - y\|_2^2) &= x \cdot y \quad (\text{Polarisationsformel}) \end{aligned}$$

Beweis. Siehe Übung. □

Lemma 1.71 (Parallelogrammgleichung). Für alle $x, y \in \mathbb{R}^d$ gilt

$$\|x + y\|_2^2 + \|x - y\|_2^2 = 2(\|x\|_2^2 + \|y\|_2^2).$$

Beweis. Siehe Übung □

26.4.2020

#13

26.4.2020

Definition 1.72 (Skalarprodukt, Prähilbertraum). Sei H ein Vektorraum. Eine Abbildung $\langle \cdot, \cdot \rangle : H \times H \rightarrow \mathbb{R}$ heißt Skalarprodukt, falls:

(a) (Bilinearität). Für alle $x \in H$ sind die Abbildungen $y \mapsto \langle x, y \rangle$ und $y \mapsto \langle y, x \rangle$ lineare Abbildungen von H nach $\mathbb{R}$.

(b) (Symmetrie). Es gilt $\langle x, y \rangle = \langle y, x \rangle$ für alle $x, y \in H$

(c) (Positive Definitheit). Für alle $x \in H$ gilt $\langle x, x \rangle \geq 0$ sowie $\langle x, x \rangle = 0 \Leftrightarrow x = 0$.

Das Paar $(H, \langle \cdot, \cdot \rangle)$ heißt Prähilbertraum.

Beispiel 1.73. Sei $A \in \mathbb{R}^{d \times d}$. Wir betrachten die Abbildung $\langle \cdot, \cdot \rangle_A : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$, $\langle x, y \rangle_A := x \cdot Ay$. Dann ist $\langle \cdot, \cdot \rangle_A$

(a) bilinear.

(b) symmetrisch $\Leftrightarrow A^\top = A$,

(c) positiv definit $\Leftrightarrow A$ ist positiv definit (d.h. $\exists \lambda > 0 : x \cdot Ax \geq \lambda \|x\|_2^2$).

Wir folgern: $(\mathbb{R}^d, \langle \cdot, \cdot \rangle_A)$ ist genau dann ein Prähilbertraum, wenn A symmetrisch und positiv definit ist.

Satz 1.74. Auf jedem Prähilbertraum wird durch

$$\|\cdot\| : H \rightarrow \mathbb{R}, x \mapsto \|x\| := \sqrt{\langle x, x \rangle}$$

eine Norm definiert. Für alle $x, y \in H$ gelten:

$$(1.13) \quad |\langle x, y \rangle| \leq \|x\| \|y\|, \quad (\text{Cauchy-Schwarz Ungleichung}),$$

$$(1.14) \quad \|x + y\|^2 = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2,$$

$$(1.15) \quad \|x - y\|^2 = \|x\|^2 - 2\langle x, y \rangle + \|y\|^2,$$

$$(1.16) \quad \langle x, y \rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2), \quad (\text{Polarisationsformel}),$$

$$(1.17) \quad \|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2), \quad (\text{Parallelogrammgleichung}).$$

Beweis. 1. Die Identitäten (1.14) - (1.17) folgen durch Nachrechnen ganz genauso wie im Falle des Standardskalarproduktes.

2. Aus der Bilinearität und positiven Definitheit von $\langle \cdot, \cdot \rangle$ folgt die absolute Homogenität und positive Definitheit von $\| \cdot \|$.
3. Cauchy-Schwarzsche Ungleichung (CS): Zunächst betrachten wir den Fall $\|x\| = \|y\| = 1$. Es gilt

$$0 \leq \|x \pm y\|^2 = \|x\|^2 \pm 2\langle x, y \rangle + \|y\|^2 = 2(1 \pm \langle x, y \rangle).$$

Also,

$$|\langle x, y \rangle| \leq 1.$$

Aus der Bilinearität folgt nun für $x, y \in H \setminus \{0\}$,

$$|\langle x, y \rangle| = \|x\| \|y\| \left| \left\langle \frac{x}{\|x\|}, \frac{y}{\|y\|} \right\rangle \right| \leq \|x\| \|y\|.$$

4. Dreiecksungleichung:

$$\|x + y\|^2 = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2 \stackrel{\text{CS}}{\leq} \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$

□

Definition 1.75 (Hilbertraum). *Ein vollständiger Prähilbertraum wird Hilbertraum genannt.*

⁷Jeder Prähilbertraum ist ein normierter Raum. Da endlichdimensionale normierte Räume automatisch vollständig sind, sind endlichdimensionale Prähilberträume automatisch Hilberträume. Wir werden daher im endlichdimensionalen Fall immer nur vom Hilbertraum sprechen.

Beispiel 1.76. Sei $A \in \mathbb{R}^{d \times d}$ symmetrisch, positiv definit und $\langle \cdot, \cdot \rangle_A$ wie in Beispiel (1.73) definiert. Wir setzen $\| \cdot \|_A := \sqrt{\langle x, x \rangle_A}$. Dann ist $(\mathbb{R}^d, \| \cdot \|_A)$ ein Hilbertraum.

Wann ist ein normierter Raum ein Prähilbertraum?

Lemma 1.77. Sei $(X, \| \cdot \|)$ ein normierter Raum. Es gelte die Parallelogrammgleichung (1.17). Dann wird durch

$$\langle x, y \rangle := \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)$$

ein Skalarprodukt auf X definiert und es gilt $\|x\| = \sqrt{\langle x, x \rangle}$. D.h.: Ein normierter Raum ist genau dann ein Prähilbertraum, wenn die Parallelogrammgleichung gilt.

Beweis. Siehe Übung. □

Beispiel 1.78. Unter den p -Normen auf $\mathbb{R}^d$ erfüllt nur die 2-Norm die Parallelogrammgleichung. D.h. $(\mathbb{R}^d, \| \cdot \|_p)$ sind für $p \neq 2$ keine Hilberträume.

Argument: Es gilt

$$\|e_1 + e_2\|_p^2 + \|e_1 - e_2\|_p^2 = 2^{1+\frac{2}{p}}, \quad 2(\|e_1\|_p^2 + \|e_2\|_p^2) = 4.$$

Beide Terme sind genau dann identisch, wenn $p = 2$ gilt.

⁷Kommentiert am 28.4.20.

Definition 1.79 (Orthogonalität und Winkel). Sei $(H, \langle \cdot, \cdot \rangle)$ ein (Prä)hilbertraum.

(a) Wir definieren den Winkel $\alpha := \angle(x, y)$ zwischen zwei Vektoren $x, y \in H \setminus \{0\}$ durch

$$\alpha := \arccos \frac{\langle x, y \rangle}{\|x\| \|y\|}.$$

(b) $x, y \in H$ werden orthogonal (zueinander) genannt, falls $\langle x, y \rangle = 0$. Wir schreiben in diesem Fall $x \perp y$.

(c) Für $U \subseteq H$ bezeichnen wir mit $U^\perp := \{x \in H : \langle x, y \rangle = 0 \text{ für alle } y \in U\}$ das orthogonale Komplement von U .

Beispiele 1.80. Wir betrachten den Euklidischen Raum $\mathbb{R}^d$.

(a) Für die Einheitsvektoren $e_1, \dots, e_d$ gilt

$$e_i \cdot e_j = \delta_{ij} := \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{sonst.} \end{cases}$$

D.h. die Einheitsvektoren sind paarweise orthogonal.

(b) Es gilt $\{e_1\}^\perp = \{x = (0, x_2, \dots, x_d) : x_2, \dots, x_d \in \mathbb{R}\} = \text{span}\{e_2, \dots, e_d\}$. Das ist ein $d - 1$ -dimensionaler Unterraum von $\mathbb{R}^d$.

(c) Sei nun $A \in \mathbb{R}^{d \times d}$ eine symmetrische, positiv definite Matrix und $\langle \cdot, \cdot \rangle_A$ das Skalarprodukt aus Beispiel 1.76. Dann sind im Allgemeinen die Einheitsvektoren nicht orthogonal in $(\mathbb{R}^d, \langle \cdot, \cdot \rangle_A)$.
Beispiel: Für $d = 2$ betrachten wir die Abbildung

$$A := \begin{pmatrix} a & c \\ c & b \end{pmatrix}.$$

Dann gilt $e_1 \perp e_2$ in $(\mathbb{R}^2, \langle \cdot, \cdot \rangle_A)$ genau dann, wenn $c = 0$; in diesem Fall ist A eine anisotrope Streckung/Verzerrung der x_1 - und x_2 -Achse.

Lemma 1.81 (Pythagoras). Sei $(H, \langle \cdot, \cdot \rangle)$ ein Prähilbertraum. Dann gilt für alle $x, y \in H$,

$$x \perp y \quad \Leftrightarrow \quad \|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

Beweis. Das folgt direkt aus der Identität (1.14). □

Definition 1.82 (Orthonormalbasis). Sei $(H, \langle \cdot, \cdot \rangle)$ ein Hilbertraum mit Dimension $d \in \mathbb{N}_+$. Eine Basis $q_1, \dots, q_d \in H$ wird Orthonormalbasis genannt, falls

$$\langle q_i, q_j \rangle = \delta_{ij} \quad \text{für alle } i, j = 1, \dots, d.$$

Beispiel 1.83. (a) Die Standardbasis $e_1, \dots, e_d$ ist eine Orthonormalbasis für den Euklidischen Vektorraum $\mathbb{R}^d$.

Sei $Q \in O(d)$ (eine orthogonale Matrix). Dann bilden die Vektoren $q_i := Qe_i$, $i = 1, \dots, d$, (d.h. die Spaltenvektoren von Q) eine Orthonormalbasis des Euklidischen Vektorraums $\mathbb{R}^d$.

(b) Sei $A \in \mathbb{R}^{d \times d}$ symmetrisch und positiv definit, $\langle x, y \rangle_A := x \cdot Ay$. Dann gilt

$$\langle e_i, e_j \rangle_A = e_i \cdot Ae_j = A_{ij}.$$

Also ist $e_1, \dots, e_d$ genau dann eine Orthonormalbasis für $(\mathbb{R}^d, \langle \cdot, \cdot \rangle_A)$, wenn $A = I$.

(c) Mit Hilfe des Spektralsatzes aus der Linearen Algebra zeigt man, dass jeder endlichdimensionale Hilbertraum eine Orthonormalbasis besitzt. Der Begriff der Orthonormalbasis kann auch für unendlichdimensionale Räume eingeführt werden. Damit beschäftigt man sich in der Funktionalanalysis.

Beispiel 1.84 (Koordinatendarstellung). (a) Für $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ gilt

$$x = \sum_{i=1}^d \underbrace{(x \cdot e_i)}_{=x_i} e_i.$$

Die Abbildungen

$$\mathbb{R}^d \rightarrow \mathbb{R}, \quad x \mapsto x \cdot e_i$$

sind linear und stetig und heißen Koordinatenfunktionen (bzgl. der Standardbasis).

(b) Für $T \in \text{Lin}(\mathbb{R}^d, \mathbb{R}^d)$ gilt

$$T(x) = \sum_{i,j=1}^d A_{ij} (x \cdot e_j) e_i \text{ mit } A_{ij} := e_i \cdot T(e_j)$$

Wir fassen die A_{ij} zur sogenannten Darstellungsmatrix $A \in \mathbb{R}^{d \times d}$ zusammen.

(c) Sei $(H, \langle \cdot, \cdot \rangle)$ ein Hilbertraum mit Dimension $d \in \mathbb{N}_+$ und $q_1, \dots, q_d$ eine Orthonormalbasis. Für $x \in H$ gilt

$$x = \sum_{i=1}^d \langle x, q_i \rangle q_i$$

Die Abbildungen

$$H \rightarrow \mathbb{R}, \quad x \mapsto \langle x, q_i \rangle$$

sind linear und stetig und heißen Koordinatenfunktionen (bzgl. der Basis $q_1, \dots, q_d$).

Sei $T \in \text{Lin}(H; H)$. Dann gilt für alle $x \in H$,

$$T(x) = \sum_{i,j=1}^d A_{ij} \langle x, q_j \rangle q_i \text{ mit Darstellungsmatrix } A \in \mathbb{R}^{d \times d}, A_{ij} := \langle q_i, T(q_j) \rangle.$$

Beispiel 1.85. ⁸ Sei $(H, \langle \cdot, \cdot \rangle)$ ein endlichdimensionaler Hilbertraum mit Orthonormalbasis $q_1, \dots, q_d$. Betrachte die affine Transformation

$$T : H \rightarrow H, \quad T(x) := \left(\sum_{i,j=1}^d Q_{ij} \langle x, q_j \rangle q_i \right) + c \quad \text{mit } Q \in O(d), c \in H.$$

(Im Spezialfall $H = \mathbb{R}^d$ mit Standardskalarprodukt und Standardbasis ist T eine Euklidische Transformation). T ist Längen- und Winkelerhaltend: Für alle $x, y \in H$ gilt

$$\|T(x) - T(y)\| = \|x - y\|, \quad \angle(x, y) = \angle(Tx, Ty).$$

Das sieht man genau so, wie im Spezialfall $H = \mathbb{R}^d$ mit Standardskalarprodukt und Standardbasis (siehe Übung).

Lemma 1.86. Sei $(H, \langle \cdot, \cdot \rangle)$ ein Prähilbertraum, $a \in H$. Dann ist die Abbildung

$$T : H \rightarrow \mathbb{R}, \quad T(x) := \langle x, a \rangle$$

linear und stetig.

Beweis. Die Linearität folgt direkt aus der Bilinearität des Skalarproduktes. Aus der CS-Ungleichung folgt

$$|T(x)| = |\langle x, a \rangle| \leq \|x\| \|a\| \quad \text{für alle } x \in H.$$

Insbesondere folgt hieraus $T(B(0; 1)) \subseteq B(0; \|a\|)$. Nach Satz 1.25 ist T stetig. $\square$

Beispiel 1.87. Sei $(H, \langle \cdot, \cdot \rangle)$ ein Prähilbertraum und (x_n) eine Folge in H . Dann gilt für alle $a \in H$ die Implikation

$$x_n \rightarrow x \quad \Rightarrow \quad \langle x_n, a \rangle \rightarrow \langle x, a \rangle.$$

Falls H endlichdimensional ist gilt auch die Rückrichtung: Sei $q_1, \dots, q_d$ eine Orthonormalbasis von H und es gelte $\langle x_n, q_i \rangle \rightarrow \langle x, q_i \rangle$ für alle $i = 1, \dots, d$. Dann folgt

$$\|x_n - x\| = \left\| \sum_{i=1}^d \langle x_n - x, q_i \rangle q_i \right\| \leq \sum_{i=1}^d |\langle x_n - x, q_i \rangle| \rightarrow 0.$$

Die Richtung " $\Leftarrow$ " ist für unendlichdimensionale Hilberträume übrigens im Allgemeinen nicht wahr.

Beispiel 1.88. Sei $(H, \langle \cdot, \cdot \rangle)$ ein Prähilbertraum und $U \subseteq H$. Sei (x_n) eine Folge in $U^\perp$ mit $x_n \rightarrow x$ in H . Dann gilt für alle $a \in U$,

$$0 = \langle x_n, a \rangle \rightarrow \langle x, a \rangle.$$

Also $x \in U^\perp$. Wir haben gezeigt, dass das orthogonale Komplement einer beliebigen Menge abgeschlossen ist. Man kann übrigens zeigen, dass $(U^\perp)^\perp = \overline{\text{span}(U)}$.

Kontrollfrage 4.3. Wir betrachten die p -Norm auf $\mathbb{R}^d$. Es gilt $\|Qx\|_p = \|x\|_p$ für alle $x \in \mathbb{R}^d$ und alle $Q \in O(d)$,

(A) für alle $1 < p < \infty$.

⁸Korrektur am 30.4.20

$\times$: Vgl. 1.67

(B) genau dann, wenn $p = 2$.

✓: Vgl. Lemma 1.68

Kontrollfrage 4.4. Wir betrachten den Euklidischen Raum $\mathbb{R}^d$. Welche der folgenden Aussagen sind wahr?

(A) Es existiert $A \in \mathbb{R}^{d \times d} \setminus O(d)$, so dass: Für alle $x, y \in \mathbb{R}^d$ sind die Winkel zwischen x, y und zwischen Ax, Ay identisch.

✓: $A = 3I$

(B) Falls $A \in \mathbb{R}^{d \times d}$ orthogonal ist, dann gilt: Für alle $x, y \in \mathbb{R}^d$ sind die Winkel zwischen x, y und zwischen Ax, Ay identisch.

✓: Vgl. Beispiel 1.85

Kontrollfrage 4.5. Für welche der folgenden Matrizen A ist die Abbildung $(x, y) \mapsto x \cdot Ay$ ein Skalarprodukt auf $\mathbb{R}^2$?

(A) $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$

$\times$: nicht symmetrisch

(B) $A = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$

✓: Vgl. Bemerkung 1.73.

(C) $A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

$\times$: nicht positiv definit

(D) $A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$

$\times$: nicht positiv

Kontrollfrage 4.6. Wir betrachten $\mathbb{R}^d$. Welche der folgenden Aussagen sind wahr.

(A) Auf $\mathbb{R}^d$ gibt es unendlich viele Skalarprodukte.

✓: Bemerkung 1.73

(B) Auf $\mathbb{R}^d$ gibt es genau ein Skalarprodukt mit der Eigenschaft, dass die Vektoren $e_1, \dots, e_d$ paarweise orthogonal sind.

$\times$: Gegenbeispiel: $\langle x, y \rangle := 42x \cdot y$

(C) Sei $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf $\mathbb{R}^d$. Dann bilden die Vektoren $e_1, \dots, e_d$ genau dann eine Orthonormalbasis in $(\mathbb{R}^d, \langle \cdot, \cdot \rangle)$, wenn das Skalarprodukt das Standardskalarprodukt ist.

✓: $\langle x, y \rangle = \langle (\sum_i x_i e_i), (\sum_j y_j e_j) \rangle$. Dann Ausmultiplizieren und sehen, dass der Wert eindeutig definiert ist.

Kontrollfrage 4.7. Sei $(H, \langle \cdot, \cdot \rangle)$ ein Prähilbertraum. Dann ist das orthogonale Komplement der Menge H bzgl. des Skalarproduktes

(A) $\emptyset$

✗: Es gilt immer $0 \cdot x = 0$.

(B) $\{0\}$

✓: $(\forall x \in H : \langle 0, x \rangle = 0 \text{ und für } x \in H \setminus \{0\} \text{ gilt } \langle x, x \rangle \neq 0.)$

(C) H

✗: Aus $H \perp H$ folgt $\|x\|^2 = \langle x, x \rangle = 0$ für alle $x \in H$, also $H = \{0\}$.

Kontrollfrage 4.8. In einem Prähilbertraum ist das orthogonale Komplement einer Menge stets abgeschlossen.

(A) Das ist wahr.

✓: Folgt mit Lemma 1.86

(B) Das ist falsch.

✗

Zuletzt skizzieren wir ein Beispiel für einen Hilbertraum mit unendlicher Dimension: Wir bezeichnen den Raum der quadratsummierbaren Folgen mit

$$\ell^2 := \{(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{R} : \sum_{n=0}^{\infty} |x_n|^2 < \infty\}.$$

Auf ℓ^2 werden ein Skalarprodukt und die induzierte Norm durch

$$\langle x, y \rangle_{\ell^2} := \sum_{n=0}^{\infty} x_n y_n, \quad \|x\|_{\ell^2} := \left(\sum_{n=0}^{\infty} x_n^2 \right)^{\frac{1}{2}} \quad \text{definiert.}$$

Man kann zeigen, dass ℓ^2 vollständig ist und damit ein Hilbertraum. Man kann sich ℓ^2 als “Grenzwert” von $(\mathbb{R}^d, \|\cdot\|_2)$ für $d \rightarrow \infty$ vorstellen: Der Raum der endlichen Folgen ist eine Teilmenge von ℓ^2 deren Abschluss ganz ℓ^2 ergibt.

Die Folgen

$$e_i := (0, \dots, 0, \underbrace{1}_{\text{ites Folgenglied}}, 0, 0, \dots)$$

sind orthogonal in ℓ^2 und es gilt $\|e_i\|_{\ell^2} = 1$. Damit ist die Dimension von ℓ^2 unendlich.

In der Funktionalanalysis zeigt man: Jedes $x \in H$ lässt sich beliebig genau durch eine (endliche) Linearkombination der Folgen $\{e_i\}_{i \in \mathbb{N}}$ approximieren.

26.4.2020

#15

3.5.2020

1.7 Konvergenz von Funktionenfolgen

Ziele:

- Funktionenfolgen: gleichmäßige versus punktweise Konvergenz
- Der normierte Raum $C(K)$ der stetigen Funktionen auf Kompakta K

Definition 1.89. Sei D eine nichtleere Menge, $(f_n)_{n \in \mathbb{N}}$ eine Folge in $\text{Abb}(D; \mathbb{K}^d)$.

(a) (f_n) heißt punktweise konvergent gegen $f \in \text{Abb}(D; \mathbb{K}^d)$, falls

$$\lim_{n \rightarrow \infty} f_n(x) = f(x) \quad \text{in } \mathbb{K}^d \text{ für alle } x \in D.$$

(b) (f_n) heißt gleichmäßig konvergent gegen $f \in \text{Abb}(D; \mathbb{K}^d)$, falls für alle $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert mit

$$\|f_n(x) - f(x)\| < \varepsilon \quad \text{für alle } x \in D \text{ und } n \geq n_0.$$

(Hierbei bezeichnet $\|\cdot\|$ eine beliebige Norm auf $\mathbb{K}^d$.)

Bemerkung 1.90.

- (a) Da alle Normen auf $\mathbb{K}^d$ äquivalent sind, ist es egal welche man bei der Definition der gleichmäßigen Konvergenz betrachtet.
- (b) Offensichtlich impliziert gleichmäßige Konvergenz die punktweise Konvergenz gegen die selbe Grenzfunktion.

Beispiele 1.91. Wir betrachten die Funktionenfolge $f_n : (0, 1) \rightarrow \mathbb{R}$, $n \in \mathbb{N}_+$,

$$f_n(x) := \begin{cases} n & \text{falls } x < \frac{1}{n}, \\ \frac{1}{x} & \text{sonst.} \end{cases}$$

Offensichtlich gilt

$$\lim_{n \rightarrow \infty} f_n(x) = f(x) := \frac{1}{x} \quad \text{für alle } x \in (0, 1),$$

und andererseits

$$\left| f_n\left(\frac{1}{2n}\right) - f\left(\frac{1}{2n}\right) \right| = |n - 2n| = n \quad \text{für alle } n \in \mathbb{N}_+.$$

Damit ist die Folge punktweise gegen f konvergent, aber nicht gleichmäßig. Weiter beobachten wir, dass $(f_n)_{n \in \mathbb{N}}$ eine Folge beschränkter Funktionen ist, aber die Grenzfunktion unbeschränkt ist.

Beispiel 1.92. Wir betrachten die Funktionenfolge $f_n : (-1, 1) \rightarrow \mathbb{R}$, $n \in \mathbb{N}_+$,

$$f_n(x) := \begin{cases} -1 & \text{für } x < -\frac{1}{2n}, \\ 2nx & \text{für } -\frac{1}{2n} \leq x \leq \frac{1}{2n}, \\ 1 & \text{für } x > \frac{1}{2n}. \end{cases}$$

Offensichtlich gilt einerseits

$$\lim_{n \rightarrow \infty} f_n(x) = f(x) := \begin{cases} -1 & x < 0, \\ 0 & x = 0, \\ 1 & x > 0, \end{cases}$$

und andererseits $|f_n(\frac{1}{4n}) - f(x)| = \frac{1}{2}$ für alle $n \in \mathbb{N}_+$. Damit ist die Folge gegen f punktweise, aber nicht gleichmäßig konvergent. Weiter beobachten wir, dass $(f_n)_{n \in \mathbb{N}}$ eine Folge stetiger Funktionen ist, aber der Grenzwert f nicht stetig ist.

Wir sehen, dass bei punktweiser Konvergenz Stetigkeit und Beschränktheit von Funktionen nicht erhalten bleiben. Die gleichmäßige Konvergenz hat bessere Eigenschaften.

Lemma 1.93 (Beschränktheit und gleichmäßige Konvergenz). *Sei D eine nichtleere Menge und $(f_n)_{n \in \mathbb{N}}$ eine Folge in $(B(D), \|\cdot\|_\infty)$. Sei f_n gleichmäßig konvergent gegen $f : D \rightarrow \mathbb{R}$. Dann gilt $f \in B(D)$.*

Beweis. Aufgrund der gleichmäßigen Konvergenz existiert $n \in \mathbb{N}$ mit

$$|f_n(x) - f(x)| \leq 1 \quad \text{für alle } x \in D.$$

Aus

$$|f(x)| \leq |f(x) - f_n(x)| + |f_n(x)| \leq 1 + \|f_n\|_\infty,$$

folgt nun durch Übergang zum Supremum $\|f\|_\infty \leq 1 + \|f_n\|_\infty$. □

Lemma 1.94 (Stetigkeit und gleichmäßige Konvergenz). *Sei $D \subseteq \mathbb{R}^d$ nichtleer und $(f_n)_{n \in \mathbb{N}} \subseteq C(D)$ gleichmäßig gegen ein $f \in \text{Abb}(D; \mathbb{R})$ konvergent. Dann gilt $f \in C(D)$.*

Beweis. Wir weisen das $\varepsilon - \delta$ -Kriterium für f nach. Sei $x_0 \in D$ und $\varepsilon > 0$. Dann existiert $n \in \mathbb{N}$ mit

$$|f_n(x) - f(x)| < \frac{\varepsilon}{3} \quad \text{für alle } x \in D.$$

Da f_n stetig in x_0 ist, existiert ein $\delta > 0$, so dass

$$|f_n(x) - f_n(x_0)| < \frac{\varepsilon}{3} \quad \text{für alle } x \in B(x_0; \delta).$$

Für alle $x \in B(x_0; \delta)$ folgt nun mit der Dreiecksungleichung

$$|f(x) - f(x_0)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)| \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

□

3.5.2020

#16

3.5.2020

Für Folgen in $B(D)$ sind Konvergenz bzgl. der Supremumsnorm und gleichmäßige Konvergenz äquivalent.

Satz 1.95 (Supremumsnorm und gleichmäßige Konvergenz). *Sei D eine nichtleere Menge, $f_n, f \in B(D)$ für $n \in \mathbb{N}$. Dann sind äquivalent*

(a) $\|f_n - f\|_\infty \rightarrow 0$.

(b) $f_n \rightarrow f$ gleichmäßig.

Der Beweis ist offensichtlich.

Bemerkung 1.96.

(a) Es ist sehr hilfreich zu wissen, dass die gleichmäßige Konvergenz in $B(D)$ durch eine Norm charakterisiert werden kann. Zum Beispiel wissen wir nun, dass eine gleichmäßig konvergente Folge in $B(D)$ eine Cauchy-Folge ist, oder dass der Grenzwert einer gleichmäßig konvergenten Folge in $S := \{f \in B(D) : \|f\|_\infty = 1\}$ (der Einheitssphäre) auch in S liegen muss.

(b) Die punktweise Konvergenz in $B(D)$ kann im Allgemeinen nicht durch eine Norm charakterisiert werden. Für allgemeine D existiert nicht einmal eine Metrik, mit der die punktweise Konvergenz charakterisiert werden könnte. Falls D jedoch eine endliche Menge ist, dann sind punktweise und gleichmäßige Konvergenz (und damit auch Konvergenz bzgl. der Supremumsnorm) äquivalent.

Wir haben bereits in der Übung bewiesen, dass $B(D)$ (der Raum der reellen, beschränkten Funktionen auf einer Menge D) zusammen mit der Supremumsnorm $\|\cdot\|_\infty$ ein Banachraum ist.

Satz 1.97 ($(C(K), \|\cdot\|_\infty)$ ist ein Banachraum). Sei $K \subseteq \mathbb{K}^d$ kompakt. Dann ist $C(K)$ ein abgeschlossener Unterraum von $(B(K); \|\cdot\|_\infty)$ und damit insbesondere ein Banachraum.

Beweis. 1. Argument für Unterraum: Aus K kompakt und $f \in C(K)$ stetig folgt mit Satz 1.57 (Satz vom Minimum und Maximum): $\|f\|_\infty < \infty$ und damit $f \in B(K)$. Damit ist der Vektorraum $C(K)$ ein Unterraum von $B(K)$.

2. Argument für abgeschlossen: Sei nun $(f_n) \subseteq C(K)$ eine konvergente Folge mit Grenzwert $f \in B(K)$. Nach Satz 1.95 konvergiert dann f_n gleichmäßig gegen f . Aus Lemma 1.94 folgt nun $f \in C(K)$. □

Bemerkung 1.98. Der Satz gilt auch, wenn wir K durch eine kompakte Teilmenge eines metrischen Raumes ersetzen. Auf die Kompaktheit kann jedoch nicht verzichtet werden:

Für $D = (0, 1)$ ist $C(D)$ kein Unterraum von $B(D)$, da beispielsweise $f(x) := \frac{1}{x}$ in $C(D)$ aber nicht in $B(D)$ enthalten ist. Für allgemeine Definitionsbereiche $D \subset \mathbb{K}^d$ kann man stattdessen den Raum der beschränkten, stetigen Funktionen $C_b(D) := C(D) \cap B(D)$, der mit der Supremumsnorm ein abgeschlossener Unterraum von $B(D)$ ist, definieren.

Mit $(B(D), \|\cdot\|_\infty)$ und $(C(D), \|\cdot\|_\infty)$ haben wir zwei unendlichdimensionale Banachräume kennengelernt. Zum Abschluss dieses Kapitels geben wir ein Beispiel für einen unvollständigen normierten Raum.

Beispiel 1.99. Wir betrachten den Vektorraum $C^1(I)$ mit $I := [-1, 1]$. Als Unterraum von $(C(I), \|\cdot\|_\infty)$ ist $(C^1(I), \|\cdot\|_\infty)$ ein normierter Raum. Wir behaupten, dass $C^1(I)$ nicht abgeschlossen ist (und damit kein Banachraum sein kann). Um das zu sehen konstruieren wir eine Folge $(u_n) \subseteq C^1(I)$, die gleichmäßig konvergiert, deren Grenzwert aber nicht in $C^1(I)$ liegt. Setze $u_0(x) := |x|$. Es gilt

$u_0 \in C(I) \setminus C^1(I)$, da u_0 nicht differenzierbar in 0 ist. Wir “glätten die Kante bei 0” und betrachten hierzu $u : \mathbb{R} \rightarrow \mathbb{R}$,

$$u(x) := \begin{cases} \frac{1}{2}x^2 + \frac{1}{2} & \text{falls } |x| \leq 1 \\ u_0(x) & \text{sonst.} \end{cases}$$

Es gilt $u \in C^1(\mathbb{R})$. Für $n \in \mathbb{N}_+$ betrachten wir die skalierte Funktion $u_n : I \rightarrow \mathbb{R}$, $u_n(x) := \frac{1}{n}u(nx)$. Offensichtlich gilt $(u_n) \subseteq C^1(I)$, sowie

$$|u_n(x) - u_0(x)| \leq \frac{2}{n} \quad \text{für alle } x \in I.$$

Also gilt $\|u_n - u_0\|_\infty \rightarrow 0$.

Die Supremumsnorm ist “zu schwach” um aus $C^1(I)$ einen Banachraum zu machen. Man kann zeigen, dass $C^1(I)$ mit der stärkeren Norm

$$\|f\|_{C^1(I)} := \|f\|_\infty + \|f'\|_\infty$$

ein Banachraum ist.

Bemerkung 1.100 (Verallgemeinerungen). Die Aussagen in diesem Kapitel gelten auch in allgemeineren Situation und die Beweise übertragen sich zumeist direkt. Insbesondere kann

- in Definition 1.89, Lemma 1.93, Lemma 1.94, Satz 1.95 und Satz 1.97 kann der Bildraum $\mathbb{K}^d$ (bzw. $\mathbb{R}$) durch einen beliebigen normierten Vektorraum ersetzt werden.
- in Lemma 1.94 kann darüber hinaus das Definitionsgebiet D durch einen beliebigen metrischen Raum ersetzt werden.

Kontrollfrage 5.1. Sei D eine nichtleere Menge, $f_n, f \in \text{Abb}(D; \mathbb{K})$ für alle $n \in \mathbb{N}$. Wir betrachten die Aussagen:

- (a) f_n ist punktweise konvergent gegen f .
- (b) f_n ist gleichmäßig konvergent gegen f .
- (c) $\forall \varepsilon > 0 \forall x \in D \exists n_0 \in \mathbb{N} \forall n \geq n_0 : |f_n(x) - f(x)| < \varepsilon$.
- (d) $\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall x \in D \forall n \geq n_0 : |f_n(x) - f(x)| < \varepsilon$.

Welche der folgenden Aussagen sind wahr?

(A) (a) $\Rightarrow$ (b)

$\times$: Siehe Bsp. 1.91

(B) (b) $\Rightarrow$ (a)

$\checkmark$: Siehe Bem. 1.90

(C) (a) $\Leftrightarrow$ (c)

$\checkmark$: Andere Schreibweise für Def. 1.89 a)

(D) $(a) \Leftrightarrow (d)$

$\times$: Würde sonst (A) und (F) widersprechen.

(E) $(b) \Leftrightarrow (c)$

$\times$: Würde sonst (A) und (C) widersprechen.

(F) $(b) \Leftrightarrow (d)$

$\checkmark$: Andere Schreibweise für Def. 1.89 b)

Kontrollfrage 5.2. Sei $K \subset \mathbb{R}^d$ kompakt und nichtleer. Wir betrachten $C(K)$ mit der Supremumsnorm, sowie die Menge

$$A := \{f \in C(K) : 0 \leq f(x) \leq 1 \text{ für alle } x \in K\}$$

Welche der folgenden Aussagen sind wahr?

(A) A ist beschränkt in $C(K)$.

$\checkmark$: $\forall f \in A : \|f\|_\infty \leq 1$.

(B) A ist offen in $C(K)$.

$\times$: Um die konstante Nullfunktion kann man keine Kugel, die auch in A ist, legen.

(C) A ist abgeschlossen in $C(K)$.

$\checkmark$: Die Norm $\|\cdot\|$ ist stetig auf $(C(K), \|\cdot\|_\infty)$ und A ist das Urbild unter $\|\cdot\|$ von der abgeschlossenen Menge $[0, 1]$.

Kontrollfrage 5.3. Wir betrachten die Funktionenfolge $f_n : \mathbb{R} \rightarrow \mathbb{R}$, $n \in \mathbb{N}$,

$$f_n(x) := \begin{cases} 1 & \text{falls } x < \frac{1}{n+1}, \\ 0 & \text{sonst.} \end{cases}$$

Welche der folgenden Aussagen sind wahr?

(A) (f_n) ist gleichmäßig konvergent.

$\times$: Der gleichmäßige Grenzwert, falls existent, müsste mit f aus (B) übereinstimmen, aber: Für beliebig große $n \in \mathbb{N}$ gibt es immernoch $x > 0$ mit $f_n(x) = 1$.

(B) (f_n) ist punktweise konvergent gegen

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

$\checkmark$: Für jedes $x > 0$ gilt $x > \frac{1}{n+1}$, wenn n nur groß genug wird. Für $x \leq 0$ ist hingegen $f_n(x) = 1$ für alle $n \in \mathbb{N}$.

(C) (f_n) ist punktweise konvergent gegen

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

$\times$: Siehe (B).

Kontrollfrage 5.4. Die für $n \in \mathbb{N}_+$ definierten Funktionen $f_n : \mathbb{R} \rightarrow \mathbb{R}$, $f_n(x) := x/n$ konvergieren für $n \rightarrow \infty$

(A) auf $\mathbb{R}$ gleichmäßig,

$\times$: Der gleichmäßige Grenzwert müsste mit dem punktweisen aus (B) übereinstimmen. Aber:
 $\forall n \in \mathbb{N} : f_n(n) = 1$.

(B) auf $\mathbb{R}$ punktweise,

$\checkmark$: Der Grenzwert ist die Nullfunktion.

(C) auf $[0, 1)$ gleichmäßig,

$\checkmark$: $\|f_n\|_\infty = 1/n \rightarrow 0$, weil nur $[0, 1)$ betrachtet wird.

(D) auf $[0, 1)$ punktweise.

$\checkmark$: Folgt aus (C).

Kontrollfrage 5.5. Sei $f \in \text{Abb}(\mathbb{R}; \mathbb{R})$. Was gilt für beliebige Folgen $(f_n) \subseteq \text{Abb}(\mathbb{R}; \mathbb{R})$ mit $\|f_n - f\|_\infty \rightarrow 0$?

(A) $(f_n) \subseteq C(\mathbb{R}) \Rightarrow f \in C(\mathbb{R})$

$\checkmark$: Lemma 1.94.

(B) $(f_n) \subseteq B(\mathbb{R}) \Rightarrow f \in C(\mathbb{R})$

$\times$: Gegenbeispiel: Eine konstante Folge, deren Folgenglieder allesamt eine einzige beschränkte, aber nicht stetige Funktion sind.

(C) $(f_n) \subseteq C(\mathbb{R}) \Rightarrow f \in B(\mathbb{R})$

$\times$: Gegenbeispiel: Eine konstante Folge, deren Folgenglieder allesamt eine einzige stetige, aber nicht beschränkte Funktion sind.

(D) $(f_n) \subseteq B(\mathbb{R}) \Rightarrow f \in B(\mathbb{R})$

$\checkmark$: Lemma 1.93.

Kontrollfrage 5.6. Was stimmt?

(A) Es gibt eine Folge $(f_n) \subseteq \text{Abb}(\mathbb{R}; \mathbb{R})$, die gleichmäßig und punktweise konvergiert, wobei der gleichmäßige Grenzwert sich aber vom punktweisen unterscheidet.

$\times$: Siehe Bem. 1.90.

(B) In der Definition der gleichmäßigen Konvergenz in $\text{Abb}(\mathbb{R}; \mathbb{R}^d)$ ist es egal, welche Norm man für $\mathbb{R}^d$ verwendet.

$\checkmark$: Siehe Bem. 1.90.

(C) Es gibt Funktionenfolgen, die gleichmäßig konvergieren, aber nicht punktweise.

$\times$: Siehe Bem. 1.90.

(D) Jede Folge $(f_n) \subseteq C([0, 1])$, welche eine Cauchyfolge bezüglich der $\|\cdot\|_\infty$ -Norm ist, konvergiert gleichmäßig gegen ein $f \in \text{Abb}([0, 1])$.

$\checkmark$: Siehe Satz 1.97.

(E) Jede Folge $(f_n) \subseteq C([0, 1])$, welche eine Cauchyfolge bezüglich der $\|\cdot\|_\infty$ -Norm ist, konvergiert gleichmäßig gegen ein $f \in C([0, 1])$.

$\checkmark$: Siehe Satz 1.97.

3.5.2020

#17

3.5.2020

2 Differentialrechnung im Mehrdimensionalen

In diesem Kapitel beschäftigen wir uns mit der Differenzierbarkeit von Abbildungen der Form

$$f : D \subset \mathbb{R}^d \rightarrow \mathbb{R}^n.$$

Wir sprechen von

- Kurve, falls $d = 1$ und $n \geq 1$,
- Skalarer Funktion / Skalarfeld, falls $d \geq 1$ und $n = 1$,
- Vektorwertiger Abbildung, falls $d \geq 1$ und $n \geq 1$,
- Vektorfeld, falls $d = n \geq 1$.

2.1 Kurven im $\mathbb{R}^d$

Ziele:

- Begriffe: Differenzierbare Kurve, Tangentialvektor.
- Approximation differenzierbarer Kurven durch Streckenzüge.
- Länge einer rektifizierbaren Kurve.

Definition 2.1 (Kurve). Sei $I \subseteq \mathbb{R}$ ein Intervall, (M, d) ein metrischer Raum.

- Eine stetige Funktion $f : I \rightarrow M$ heißt Kurve (in M).
- Die Bildmenge $\text{Bild}(f)$ heißt Spur der Kurve f .
- Die Elemente $f(t)$ der Spur von f heißen Kurvenpunkte von f .
- Ein Kurvenpunkt x heißt Doppelpunkt von f , falls es $s, t \in I$ gibt mit $t \neq s$ und $x = f(t) = f(s)$.

Besonders wichtig ist für uns der Fall $M = \mathbb{R}^d$. In diesem Fall ist $f : I \rightarrow \mathbb{R}^d$ gerade dann eine Kurve, wenn $f \in C(I; \mathbb{R}^d)$.

Beispiel aus der Physik:

- Kurven im $\mathbb{R}^3$ zur Beschreibung der zeitlichen Bewegung eines Körpers im Raum.
- Kurven im $\mathbb{R}^{6n}$ zur Beschreibung der Dynamik eines n -Teilchen-Systemes im Phasenraum.

Beispiele 2.2. (a) **Gerade:** Für $a, v \in \mathbb{R}^d$, $v \neq 0$, wird durch

$$f : \mathbb{R} \rightarrow \mathbb{R}^d, \quad f(t) := a + tv$$

eine Kurve definiert. Ihre Spur ist die Gerade durch a in Richtung v . Interpretiert man t als Zeit und $f(t) \in \mathbb{R}^d$, $d = 3$, als die Position eines Körpers im Raum, dann beschreibt f eine gleichförmige, geradlinige Bewegung. Sei $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ eine stetige Bijektion (z.B. $\varphi(t) = t^3$). Dann ist $g := f \circ \varphi$ eine weitere Kurve. f und g haben die gleich Spur, aber g beschreibt im Allgemeinen keine gleichförmige Bewegung mehr.

(b) **Kreis(bogen):** Für $r > 0$ und $L > 0$ wird durch

$$f : [0, L] \rightarrow \mathbb{R}^2, \quad f(t) := r(\cos t, \sin t)$$

eine Kurve definiert, deren Spur auf dem Kreis um 0 mit Radius $r > 0$ liegt. Für $L \geq 2\pi$ ist die Spur der gesamte Kreis und der Punkt $(r, 0)$ ist ein Doppelpunkt der Kurve.

(c) **Schraublinie:** Für $r > 0$, $c > 0$, $I \subseteq \mathbb{R}$ Intervall wird durch

$$f : I \rightarrow \mathbb{R}^3, \quad f(t) := (r \cos t, r \sin t, ct)$$

eine Kurve definiert, deren Spur eine Schraublinie beschreibt. f ist injektiv und besitzt keine Doppelpunkte.

(d) **Verbindungsstrecke.** Sei $(X, \|\cdot\|)$ ein normierter Raum. Seien $t_0 < t_1$ und $x_0, x_1 \in X$ gegeben. Dann wird durch

$$f : [a, b] \rightarrow X, \quad f(t) := x_0 + \frac{t-a}{b-a}(x_1 - x_0)$$

eine Kurve in X definiert. Sie beschreibt die Verbindungsstrecke von x_0 nach x_1 .

(e) **Streckenzug.** Sei $(X, \|\cdot\|)$ ein normierter Raum. Seien Parameterpunkte $t_0 < t_1 < \dots < t_n$ sowie Streckenpunkte $x_0, \dots, x_n \in X$ gegeben. Dann wird durch

$$f : [t_0, t_n] \rightarrow X, \quad f(t) := x_{i-1} + \frac{t-t_{i-1}}{t_i-t_{i-1}}(x_i - x_{i-1}) \quad \text{falls } t \in [t_{i-1}, t_i]$$

eine Kurve in X definiert. Sie beschreibt den Streckenzug durch die Punkte $x_0, \dots, x_n$.

(f) Wir betrachten den metrischen Raum

$$M := O(2) \subseteq \mathbb{R}^{2 \times 2}, \quad d(Q, R) := \|Q - R\|_\infty.$$

Dann wird durch

$$f : \mathbb{R} \rightarrow O(2), \quad f(t) := \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}$$

eine Kurve definiert.

Definition 2.3 (Wegzusammenhängend).

Ein metrischer Raum (M, d) heißt wegzusammenhängend, falls für alle $x, y \in M$ eine Kurve $f : [0, 1] \rightarrow M$ existiert mit $f(0) = x$ und $f(1) = y$.

Beispiele 2.4.

(a) Sei $I \subseteq \mathbb{R}$. Dann gilt: I Intervall $\Leftrightarrow I$ wegzusammenhängend.

(b) Sei X ein normierter Raum und $K \subseteq X$ konvex. Dann ist K wegzusammenhängend.

(c) $O(d) \subseteq \mathbb{R}^{d \times d}$ ist nicht wegzusammenhängend. Argument: Ist g eine beliebige Kurve in $O(d)$. Dann ist $u(t) := \det(g(t))$ stetig und es gilt $u(t) \in \{-1, 1\}$ für alle t . Damit muss u konstant sein entlang der Kurve. Es kann also keine Kurve existieren, die eine Matrix in $SO(d)$ mit einer in $O(d) \setminus SO(d)$ verbindet.

Kontrollfrage 5.7. Die Kurve

$$f : [0, 3\pi] \rightarrow \mathbb{R}^{2 \times 2}, \quad f(t) := \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}$$

hat

(A) keine Doppelpunkte.

×: Siehe (C).

(B) genau einen Doppelpunkt.

×: Siehe (C).

(C) unendlich viele Doppelpunkte.

✓: $f(t) = f(t + 2\pi)$ für alle $t \in [0, \pi]$.

Kontrollfrage 5.8. Für $r > 0$ betrachte man den metrischen Raum

$$M := \{x \in \mathbb{R}^2 : \min\{\|x - e_1\|_2, \|x + e_1\|_2\} \leq r\}$$

Welche der folgenden Aussagen sind richtig?

(A) $r < \frac{1}{2} \Rightarrow M$ ist wegzusammenhängend.

×: Die Menge wird durch die y -Achse in zwei Kreisscheiben getrennt. Ähnlich zu Beispiel 2.4 c) kann man einen Widerspruch konstruieren.

(B) $r \geq 1 \Rightarrow M$ ist nicht wegzusammenhängend.

$\times$: Die Menge besteht aus zwei sich überlappenden Kreisen. Und beide sind konvex und damit wegzusammenhängend.

3.5.2020

#18

3.5.2020

Wir möchten nun die Differenzierbarkeit für Kurven in $\mathbb{R}^d$ definieren.

Wir erinnern zunächst an den Differenzierbarkeitsbegriff aus der Analysis 1: Sei $I \subseteq \mathbb{R}$ ein Intervall (offen, halboffen oder abgeschlossen). Eine Funktion $u : I \rightarrow \mathbb{R}$ heißt differenzierbar in $t \in I$, falls der Grenzwert

$$u'(t) := \lim_{\substack{h \rightarrow 0 \\ h \neq 0 \wedge t+h \in I}} \frac{u(t+h) - u(t)}{h}$$

existiert. In diesem Fall nennen wir $u'(t)$ die Ableitung von u .⁹

Aus Bequemlichkeit schreiben wir statt $\lim_{\substack{h \rightarrow 0 \\ h \neq 0 \wedge t+h \in I}}$ von nun an einfach $\lim_{h \rightarrow 0}$.

Die Definition für Kurven ist analog:

Definition 2.5. Sei $I \subseteq \mathbb{R}$ ein Intervall, $f : I \rightarrow \mathbb{R}^d$ eine Kurve.

(a) f heißt differenzierbar in $t \in I$, falls der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(t+h) - f(t)}{h} \quad \text{in } \mathbb{R}^d \text{ existiert.}$$

In diesem Fall setzen wir $f'(t) := \lim_{h \rightarrow 0} \frac{f(t+h) - f(t)}{h}$. Wir nennen $f'(t)$ den Tangentialvektor an f in t bzw. Ableitung von f in t .

(b) Ein Punkt $t \in I$ heißt regulärer Punkt von f , falls f in t differenzierbar ist mit $f'(t) \neq 0$. In diesem Fall heißt

$$\frac{f'(t)}{\|f'(t)\|_2}$$

Tangenteneinheitsvektor an f in t . In diesem Fall heißt die Gerade

$$g : \mathbb{R} \rightarrow \mathbb{R}^d, \quad g(h) := f(t) + hf'(t)$$

Tangente an f im Punkt t .

Ein Punkt $t \in I$ heißt singulärer Punkt von f , falls f in t differenzierbar ist mit $f'(t) = 0$.

(c) f heißt differenzierbar, wenn f differenzierbar für alle $t \in I$ ist.

(d) f heißt stetig differenzierbar, wenn f differenzierbar für alle $t \in I$ ist und die Abbildung

$$f' : I \rightarrow \mathbb{R}^d, \quad t \mapsto f'(t)$$

stetig ist.

⁹Beachten Sie, dass wir hier — im Gegensatz zur Analysis 1 — auch die Ableitung für Randpunkte definieren.

(e) f heißt regulär, wenn f stetig differenzierbar ist und $f'(t) \neq 0$ für alle $t \in I$.

Bemerkung 2.6. Im Fall $d = 1$ ist die Definition mit jener aus Analysis 1 identisch.

Ist $f : I \rightarrow \mathbb{R}^d$ eine Kurve und $v \in \mathbb{R}^d$, dann wird durch $f_v(t) := f(t) \cdot v$ eine Funktion $f_v : I \rightarrow \mathbb{R}$ definiert. Insbesondere bezeichnen wir mit $f_i := f_{e_i}$ für $i = 1, \dots, d$ die Komponenten von f . Für diese $\mathbb{R}$ -wertige Funktionen haben wir den Begriff der Differenzierbarkeit bereits in der Analysis 1 eingeführt.

Lemma 2.7. Für eine Kurve $f : I \rightarrow \mathbb{R}^d$ sind äquivalent:

- (a) f ist (stetig) differenzierbar.
- (b) Für alle $i = 1, \dots, d$ ist f_i (stetig) differenzierbar.
- (c) Für alle $v \in \mathbb{R}^d$ ist $f_v : I \rightarrow \mathbb{R}$ (stetig) differenzierbar.

Falls f in $t \in I$ differenzierbar ist, gilt

$$f'(t) = (f'_1(t), \dots, f'_d(t)), \quad f'_v(t) = f'(t) \cdot v = \sum_{i=1}^d f'_i(t) v_i.$$

Beweis. Das folgt aus Lemma 1.22. □

Beispiele 2.8. (a) **Gerade:** Für $a \in \mathbb{R}^d$ und $v \in \mathbb{R}^d \setminus \{0\}$ ist

$$f : \mathbb{R} \rightarrow \mathbb{R}^d, \quad f(t) := a + tv$$

stetig differenzierbar. Es gilt:

$$f'(t) = v \quad \text{Tangentenvektor}, \quad \frac{v}{\|v\|_2} \quad \text{Tangenteneinheitsvektor}.$$

(b) **Kreis:** Für $r > 0$ und $L > 0$ ist

$$f : [0, L] \rightarrow \mathbb{R}^2, \quad f(t) := r(\cos t, \sin t)$$

stetig differenzierbar. Es gilt

$$f'(t) = r(-\sin t, \cos t), \quad \|f'(t)\|_2 = r, \quad \frac{f'(t)}{\|f'(t)\|_2} = (-\sin t, \cos t)$$

(c) **Streckenzug.** Der Streckenzug

$$f : [t_0, t_n] \rightarrow \mathbb{R}^d, \quad f(t) := x_{i-1} + \frac{t - t_{i-1}}{t_i - t_{i-1}}(x_i - x_{i-1}) \quad \text{falls } t \in [t_{i-1}, t_i]$$

ist für alle $t \in [t_0, t_n] \setminus \{t_1, \dots, t_{n-1}\}$ differenzierbar mit Tangentialvektor $\frac{x_i - x_{i-1}}{t_i - t_{i-1}}$ für alle $t \in (t_{i-1}, t_i)$.

Lemma 2.9 (Parametertransformation). *Seien $I, J \subseteq \mathbb{R}$ Intervalle, $f : I \rightarrow \mathbb{R}^d$ eine stetig differenzierbare Kurve und $\varphi : J \rightarrow I$ eine stetig differenzierbare Funktion. Dann ist $g := f \circ \varphi : J \rightarrow \mathbb{R}^d$ eine stetig differenzierbare Kurve und es gilt*

$$g'(t) = \varphi'(t)f'(\varphi(t))$$

Beweis. Das folgt durch komponentenweise Betrachtung aus der Kettenregel für skalare Funktionen. $\square$

Lemma 2.10. *Sei $f : I \rightarrow \mathbb{R}^d$ eine Kurve, $t \in I$ und $v \in \mathbb{R}^d$. Es sind äquivalent:*

(a) *f ist differenzierbar in t mit $f'(t) = v$.*

(b) *Die Gerade*

$$g : \mathbb{R} \rightarrow \mathbb{R}^d, \quad g(h) := f(t) + hv$$

erfüllt die Restgliedabschätzung

$$(2.1) \quad r(h) := f(t+h) - g(h), \quad \lim_{h \rightarrow 0} \frac{\|r(h)\|}{h} = 0.$$

Beweis. Das folgt durch komponentenweise Betrachtung aus dem Weierstrass'schem Zerlegungssatz (Analysis 1, Satz 7.4). $\square$

Definition 2.11 (Schnittpunkte und Schnittwinkel). *Seien $I, J \subseteq \mathbb{R}$ Intervalle und $f : I \rightarrow \mathbb{R}^d$, $g : J \rightarrow \mathbb{R}^d$ reguläre Kurven. Ein Punkt $x \in \text{Bild}(f) \cap \text{Bild}(g)$ heißt Schnittpunkt von f und g . Für einen Schnittpunkt $x = f(t) = g(s)$ mit $t \in I$ und $s \in J$ bezeichnet*

$$\angle(f'(t), g'(s))$$

den Schnittwinkel.

Bemerkung 2.12. *Achtung: Kurven müssen nicht surjektiv sein und so kann ein Schnittpunkt x zweier Kurven zu verschiedenen Kurvenabschnitten derselben Kurve gehören. Je nach betrachtetem Kurvenabschnitt können sich verschiedene Schnittwinkel ergeben. Wir können auch Schnittpunkte von Kurven mit sich selbst untersuchen.*

Wir untersuchen nun die Interpolation von Kurven durch Streckenzüge.

Lemma 2.13. *Sei $f : [a, b] \rightarrow \mathbb{R}^d$ eine Kurve. Dann sind äquivalent*

(a) *f ist eine Verbindungsstrecke der Form*

$$f(t) = f(a) + \frac{t-a}{b-a}(f(b) - f(a)) \quad \text{für alle } t \in [a, b].$$

(b) *f ist stetig differenzierbar in $[a, b]$ und f' ist konstant.*

Beweis. 1. Die Richtung (a) $\Rightarrow$ (b) ist offensichtlich.

2. Wir zeigen, dass aus der Negation von (a) die Negation von (b) folgt. Betrachte

$$u(t) := f(t) - \left(f(a) + \frac{t-a}{b-a}(f(b) - f(a)) \right).$$

Aus der Negation von (a) folgt die Existenz von $t \in (a, b)$ und $i \in \{1, \dots, d\}$ mit $u_i(t) \neq 0$. Aus dem Mittelwertsatz folgt nun die Existenz von $s_1 \in (0, t)$ mit

$$u'_i(s_1) = \frac{u_i(t) - u_i(a)}{t - a} \neq 0.$$

Andererseits gilt $u_i(b) = u_i(a)$ und damit existiert $s_2 \in (a, b)$ mit $u'_i(s_2) = 0$. Also ist u'_i nicht konstant. Es folgt, dass f' nicht konstant ist. □

Bemerkung 2.14. Der Mittelwertsatz gilt nicht für Kurven in $\mathbb{R}^d$ mit $d > 1$, d.h. es gibt stetig differenzierbare Kurven $f : [a, b] \rightarrow \mathbb{R}^d$ mit

$$\forall t \in (a, b) : \frac{f(b) - f(a)}{b - a} \neq f'(t).$$

Beispielsweise gilt für die Schraublinie

$$f : \mathbb{R} \rightarrow \mathbb{R}^3, \quad f(t) = (\cos(t), \sin(t), t),$$

$$\frac{f(2\pi) - f(0)}{2\pi} = (0, 0, 1) \neq f'(t) \text{ für alle } t \in \mathbb{R}.$$

Wie gut kann man eine Kurve durch eine Verbindungsstrecke approximieren?

Lemma 2.15. Sei $f : [a, b] \rightarrow \mathbb{R}^d$ eine stetig differenzierbare Kurve und $g : [a, b] \rightarrow \mathbb{R}^d$ die interpolierende Verbindungsstrecke

$$g(t) := f(a) + \frac{t-a}{b-a}(f(b) - f(a)).$$

Dann gilt

$$\sup_{t \in [a, b]} \|f(t) - g(t)\|_\infty \leq 2(b-a) \sup_{t \in [a, b]} \|f'(t)\|_\infty.$$

Beweis. Es genügt für jede Komponenten $i = 1, \dots, d$ zu zeigen, dass

$$\sup_{t \in [a, b]} |f_i(t) - g_i(t)| \leq 2(b-a) \sup_{t \in [a, b]} |f'_i(t)|.$$

Setze $u(t) := f_i(t) - g_i(t)$.

1. Behauptung:

$$\sup_{t \in [a, b]} |u'(t)| \leq 2 \sup_{t \in [a, b]} |f'_i(t)|.$$

Aus dem Zwischenwertsatz folgt

$$\exists s \in (a, b) : \quad g'_i(t) = \frac{f_i(b) - f_i(a)}{b - a} = f'_i(s) \quad \text{für alle } t \in [a, b].$$

Hieraus folgt $\sup_{t \in [a, b]} |g'_i(t)| \leq \sup_{t \in [a, b]} |f'_i(t)|$ und wegen $u' = f'_i - g'_i$ die Behauptung.

2. Da $t \mapsto |u(t)|$ stetig auf dem kompaktem Intervall $[a, b]$ ist, existiert $s \in [a, b]$ mit

$$\sup_{t \in [a, b]} |u(t)| = |u(s)|.$$

Wegen $u(a) = u(b) = 0$ können wir annehmen, dass $s \in (a, b)$. Aus dem Mittelwertsatz folgt

$$\exists s_2 \in (a, s) : u'(s_2) = \frac{u(s) - u(a)}{s - a} \stackrel{u(a)=0}{=} \frac{u(s)}{s - a},$$

also

$$|u(s)| \leq |b - a| \sup_{t \in [a, b]} |u'(t)|.$$

Zusammen mit Schritt 1 folgt

$$\sup_{t \in [a, b]} |u(t)| \leq 2(b - a) \sup_{t \in [a, b]} |f'_i(t)|.$$

□

Wie gut kann man eine Kurve durch einen Streckenzug interpolieren?

Lemma 2.16. Sei $f : [a, b] \rightarrow \mathbb{R}^d$ eine stetig differenzierbare Kurve. Für $n \in \mathbb{N}_+$ bezeichne $g_n : [a, b] \rightarrow \mathbb{R}^d$ den Streckenzug mit

$$t_k = a + \frac{k}{n}(b - a), \quad x_k = f(t_k), \quad k = 0, \dots, n.$$

Dann gilt

$$\sup_{t \in [a, b]} \|g_n(t) - f(t)\|_\infty \leq 2 \frac{b - a}{n} \sup_{t \in [a, b]} \|f'(t)\|_\infty.$$

Beweis. Siehe Übung. □

Bemerkung 2.17. Auf dem Raum der stetigen Kurven $C([a, b]; \mathbb{R}^d)$ wird durch

$$(2.2) \quad \|u\|_\infty := \sup_{t \in [a, b]} \|u(t)\|_\infty$$

eine Norm definiert. Mit dieser ist $C([a, b]; \mathbb{R}^d)$ ein Banachraum. Die Aussage von Lemma 2.16 lautet nun

$$\|f - g_n\|_\infty \leq 2 \frac{b - a}{n} \|f'\|_\infty.$$

Die Norm auf $\mathbb{R}^d$ auf der rechten Seite von (2.2) kann durch jede andere Norm $\|\cdot\|_*$ auf $\mathbb{R}^d$ ersetzt werden. Aus der Äquivalenz der Normen auf $\mathbb{R}^d$ folgt in der Situation von Lemma 2.16 nun

$$\|f - g_n\| \leq C \frac{b - a}{n} \|f'\|.$$

wobei

$$\|u\| := \sup_{t \in [a, b]} \|u(t)\|_*$$

und C eine Konstante bezeichnet, die nur von der Norm $\|\cdot\|_*$ abhängig ist.

Bemerkung 2.18 (Uneinheitliche Bezeichnungen).

Der Begriff “Kurve” wird nicht einheitlich in der Mathematik verwendet. Oft ist in der Literatur auch mit (geometrischer) Kurve die Spur (d.h. das Bild) einer stetigen Abbildung $f : I \rightarrow \mathbb{R}^d$ gemeint und die Abbildung f (die wir in dieser Vorlesung Kurve nennen) als Weg oder als (parametrisierte) Kurve bezeichnet. Eine geometrische Kurve (also Spur) kann durch verschiedene Wege (parametrisierte Kurven) beschrieben werden.

Wie kann man die euklidische Länge $L(f)$ einer (parametrisierten) Kurve $f : [a, b] \rightarrow \mathbb{R}^d$ definieren? Bevor wir die Länge $L(f)$ formal einführen überlegen wir uns, welche Eigenschaften die Abbildung $f \mapsto L(f)$ haben soll:

- Für eine Verbindungsstrecke zweier Punkte $x_0, x_1 \in \mathbb{R}^d$,

$$f : [a, b] \rightarrow \mathbb{R}^d, \quad f(t) := x_0 + \frac{t-a}{b-a}(x_1 - x_0)$$

sollte die Länge gerade der Abstand der Endpunkte sein,

$$L(f) \stackrel{!}{=} \|x_1 - x_0\|_2.$$

Wir haben hier den euklidischen Abstand gewählt, da dieser invariant unter euklidischen Transformationen ist. Die Schreibweise $\stackrel{!}{=}$ soll verdeutlichen, dass diese Identität später—nachdem wir $L(f)$ definiert haben—noch überprüft werden muss.

- Die Länge sollte additiv bezüglich der Zusammensetzung von Kurven sein: Ist $f : [a, b] \rightarrow \mathbb{R}^d$ eine Kurve und $s \in (a, b)$, so sollte

$$(2.3) \quad L(f) = L(f|_{[a, s]}) + L(f|_{[s, b]})$$

gelten.¹⁰

- Ein Streckenzug $f : [a, b] \rightarrow \mathbb{R}^d$ mit Parameterpunkten $t_0 = a < \dots < t_n = b$ und Streckenpunkten $x_0, \dots, x_n \in \mathbb{R}^d$ ist die Zusammensetzung von Verbindungsstrecken. Es soll daher gelten:

$$(2.4) \quad L(f) \stackrel{!}{=} \sum_{k=1}^n L(f|_{[t_{k-1}, t_k]}) = \sum_{k=1}^n \|x_k - x_{k-1}\|_2.$$

- Ausserdem möchten wir, dass die kürzeste Verbindung zweier Punkte $x_1, x_2 \in \mathbb{R}^d$ gerade die Verbindungsstrecke ist. D.h. für alle Kurven $f : [a, b] \rightarrow \mathbb{R}^d$ mit $f(a) = x_1$ und $f(b) = x_2$ soll gelten

$$(2.5) \quad \|x_2 - x_1\|_2 \leq L(f).$$

¹⁰Für eine Funktion $f : [a, b] \rightarrow \mathbb{R}^d$ und $I \subseteq [a, b]$ Intervall, bezeichnet $f|_I$ die Einschränkung von f auf I , d.h. die Funktion $I \rightarrow \mathbb{R}^d, t \mapsto f(t)$.

Für Streckenzüge (und Verbindungsstrecken) definieren wir die Länge durch (2.4).

Beispiele 2.19. Sei $n \in \mathbb{N}_+$. Es bezeichne

$$R := \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix}, \quad \alpha := \frac{2\pi}{n},$$

die Rotation in $\mathbb{R}^2$ mit Drehwinkel $360^\circ/n$.

Sei $f_n : [0, 2\pi] \rightarrow \mathbb{R}^2$ der Streckenzug mit

$$t_k = \frac{k}{n} 2\pi, \quad x_k := R^k e_1, \quad k = 0, \dots, n.$$

Die Punkte x_k liegen auf dem Einheitskreis und bilden ein regelmäßiges n -Eck.

Es gilt

$$\begin{aligned} \|x_k - x_{k-1}\|_2 &= \|R^k e_1 - R^{k-1} e_1\|_2 = \|R^{k-1}(R e_1 - e_1)\|_2 \\ &= \|R e_1 - e_1\|_2 \quad (\text{Lemma 1.67}) \\ &= \|(\cos(\alpha) - 1, -\sin(\alpha))\|_2 = \sqrt{2(1 - \cos \alpha)} \\ &= 2 \sin(\alpha/2) \quad (\text{Definition von } \sin(\alpha/2) \text{ am Einheitskreis}) \end{aligned}$$

Also gilt für die Länge L_n des Streckenzuges f_n :

$$L_n = \sum_{k=1}^n \|f_n(t_k) - f_n(t_{k-1})\|_2 = n \|R e_1 - e_1\|_2 = 2n \sin\left(\frac{\pi}{n}\right).$$

Mit Hilfe der Regel von L'Hospital folgt

$$\lim_{n \rightarrow \infty} L_n = 2\pi. \quad \text{Das ist der Umfang des Einheitskreises!}$$

Bemerkung 2.20 (Kurvulängen $\neq$ Spurlänge). Seien $x_*, x^* \in \mathbb{R}^d$ zwei verschiedene Punkte. Wir betrachten den Streckenzug $f : [0, 1] \rightarrow \mathbb{R}^d$ mit Parameterpunkten $t_k = \frac{k}{2n}$ ($k = 0, 1, \dots, 2n$) und

$$f(t_k) = \begin{cases} x_* & k \text{ ist gerade} \\ x^* & \text{sonst.} \end{cases}$$

Die Länge von f ist $2n\|x_* - x^*\|_2$ und damit ein Vielfaches der Länge der Verbindungsstrecke der beiden Punkte!

Die Spur von f ist mit der Spur der Verbindungsstrecke g von x_* nach x^* identisch. Die (parametrisierten) Kurven f und g sind aber nicht gleich. Die oben definierte Länge ist also (zunächst) eine Eigenschaft der parametrisierten Kurve und nicht der Spur der Kurve!

Wir setzen unsere informelle Diskussion zur Definition der Länge fort. Um die Länge einer beliebigen Kurve $f : [a, b] \rightarrow \mathbb{R}^d$ zu definieren, interpolieren wir f durch einen Streckenzug. Sei also $f_n : [a, b] \rightarrow \mathbb{R}^d$ ein Streckenzug mit Parameterpunkten $a = t_0 < t_1 < \dots < t_n = b$ und Streckenpunkten $x_k = f(t_k)$. Beide Kurven verbinden die Kurvenpunkte $f(t_0), \dots, f(t_n)$. Der Streckzug f_n verbindet diese Punkte jedoch durch "gerade" Verbindungsstrecken. Daher sollte die Länge von f größer (gleich) der Länge von f_n sein. Wir werden daher die Länge von f als das Supremum über die Längen aller Streckenzüge definieren, die f interpolieren.

Definition 2.21 (Zerlegung). Sei $[a, b] \subseteq \mathbb{R}$ ein beschränktes, nichtleeres Intervall.

(a) Eine endliche Menge der Form

$$Z = \{t_0, \dots, t_n : a = t_0 < t_1 < \dots < t_n = b\} \quad (n \in \mathbb{N}_+)$$

heißt Zerlegung von $[a, b]$.

Wir definieren die Feinheit von Z durch¹¹

$$|Z| := \max\{t_k - t_{k-1} : 1 \leq k \leq n\}.$$

Konvention: Wir benutzen für eine Zerlegung Z die Schreibweise $Z = \{t_0 < \dots < t_n\}$ um auszudrücken, dass Z aus den Punkten $t_0, \dots, t_n$ besteht und diese geordnet sind, d.h. $t_0 < t_1 < \dots < t_n$.

(b) Wir bezeichnen mit $\mathcal{Z}[a, b]$ die Menge aller Zerlegungen von $[a, b]$.

(c) Für eine Kurve $f : [a, b] \rightarrow \mathbb{R}^d$ und eine Zerlegung $Z = \{t_0 < \dots < t_n\} \in \mathcal{Z}[a, b]$ definieren wir

$$L(f; Z) := \sum_{k=1}^n \|f(t_k) - f(t_{k-1})\|_2.$$

$L(f; Z)$ ist gerade die in (2.4) informell definierte Länge des Streckenzuges mit Parameterpunkten $t_0, \dots, t_n$ und Streckenpunkten $f(t_0), \dots, f(t_n)$.

Wir untersuchen nun einige Eigenschaften von $L(f; Z)$.

Bemerkung 2.22. • Auf $\mathcal{Z}[a, b]$ ist durch die Relation $\subseteq$ eine partielle Ordnung gegeben. Gilt $Z_1 \subseteq Z_2$ so sagen wir auch, dass Z_2 feiner als Z_1 ist.

- Die Vereinigung $Z_1 \cup Z_2$ zweier Zerlegungen ist eine Zerlegung von $[a, b]$. Diese ist feiner als Z_1 und als Z_2 .
- Spezielle Zerlegungen sind äquidistante Zerlegungen der Form

$$t_k = a + \frac{k}{n}(b - a), \quad k = 0, \dots, n.$$

Das Funktional $L(f; Z)$ ist additiv in folgendem Sinne:

Lemma 2.23. Sei $f : [a, b] \rightarrow \mathbb{R}^d$ eine Kurve, $Z \in \mathcal{Z}[a, b]$ und $s \in Z \setminus \{a, b\}$. Setze $I := [a, s]$ und $J := [s, b]$. Dann sind $Z \cap I$ und $Z \cap J$ Zerlegungen von $[a, s]$ und $[s, b]$. Weiter gilt

$$L(f; Z) = L(f|_I; Z \cap I) + L(f|_J; Z \cap J).$$

Aus Bequemlichkeit schreiben wir von nun an einfach $L(f; Z \cap I)$ statt $L(f|_I; Z \cap I)$.

¹¹Korrektur 17.5.2020

Beweis. Sei $Z = \{t_0 < \dots < t_n\}$ und $s = t_\ell$ mit $\ell \in \{1, \dots, n-1\}$. Dann gilt

$$\begin{aligned} L(f; Z) &= \sum_{k=1}^n \|f(t_k) - f(t_{k-1})\|_2 = \sum_{k=1}^{\ell} \|f(t_k) - f(t_{k-1})\|_2 + \sum_{k=\ell+1}^n \|f(t_k) - f(t_{k-1})\|_2 \\ &= L(f; Z \cap I) + L(f; Z \cap J). \end{aligned}$$

□

Lemma 2.24 (Dreiecksungleichung). *Sei $f : [a, b] \rightarrow \mathbb{R}^d$ eine Kurve und $s \in (a, b)$. Dann gilt*

$$L(f; \{a, b\}) \leq L(f; \{a, s, b\}).$$

Beweis. Das folgt aus der Dreiecksungleichung für $\|\cdot\|_2$:

$$L(f; \{a, b\}) = \|f(b) - f(a)\|_2 \leq \|f(s) - f(a)\|_2 + \|f(b) - f(s)\|_2 = L(f; \{a, s, b\}).$$

□

Lemma 2.25 (Monotonie). *Sei $f : [a, b] \rightarrow \mathbb{R}^d$ eine Kurve und $Z_1, Z_2 \in \mathcal{Z}[a, b]$. Dann gilt:*

$$Z_1 \subseteq Z_2 \quad \Rightarrow \quad L(f; Z_1) \leq L(f; Z_2).$$

Beweis. Durch Iteration lässt sich die Behauptung auf den Fall $Z_2 \setminus Z_1 = \{s\}$ zurückführen. Sei also $Z_1 = \{t_0 < t_1 < \dots < t_n\}$ und $Z_2 = Z_1 \cup \{s\}$ für ein $s \in [a, b] \setminus Z_1$. Wähle ℓ mit $s \in Z_2 \cap [t_{\ell-1}, t_\ell]$. Dann gilt mit Lemma 2.24 (Dreiecksungleichung)

$$L(f; Z_1 \cap [t_{\ell-1}, t_\ell]) = L(f; \{t_{\ell-1}, t_\ell\}) \leq L(f; \{t_{\ell-1}, s, t_\ell\}) = L(f; Z_2 \cap [t_{\ell-1}, t_\ell]).$$

Mit Lemma 2.23 (Additivität) gilt nun

$$\begin{aligned} L(f; Z_1) &= L(f; Z_1 \cap [t_0, t_{\ell-1}]) + L(f; Z_1 \cap [t_{\ell-1}, t_\ell]) + L(f; Z_1 \cap [t_\ell, t_n]) \\ &\leq L(f; Z_2 \cap [t_0, t_{\ell-1}]) + L(f; Z_2 \cap [t_{\ell-1}, t_\ell]) + L(f; Z_2 \cap [t_\ell, t_n]) \\ &= L(f; Z_2). \end{aligned}$$

□

Ein Streckenzug besitzt eine natürlichen Zerlegung – die Menge der Parameterpunkte. Wir können jedoch auch feinere Zerlegungen betrachten. Hierbei ändert sich aber die Länge nicht:

Lemma 2.26. *Sei $Z_1 = \{t_0 < \dots < t_n\}$ eine Zerlegung von $[a, b]$. Sei $f : [a, b] \rightarrow \mathbb{R}^d$ ein Streckenzug mit Parameterpunkten $t_0, \dots, t_n \in Z_1$ und Streckenpunkten $x_0, \dots, x_n$. Sei Z_2 eine weitere Zerlegung mit $Z_1 \subseteq Z_2$. Dann gilt*

$$L(f; Z_2) = L(f; Z_1) = \sum_{k=1}^n \|x_k - x_{k-1}\|_2.$$

Beweis. Aufgrund der Additivität (Lemma 2.23) genügt es für alle $k = 1, \dots, n$ zu zeigen, dass

$$L(f; Z_2 \cap [t_{k-1}, t_k]) = \|f(t_{k-1}) - f(t_k)\|_2.$$

Wir definieren Parameterpunkte $s_0, \dots, s_m$ durch $Z_2 \cap [t_{k-1}, t_k] = \{s_0 := t_{k-1} < s_1 < \dots < s_m := t_k\}$. Dann gilt

$$L(f; Z_2 \cap [t_{k-1}, t_k]) = \sum_{i=1}^m \|f(s_i) - f(s_{i-1})\|_2.$$

Auf $[s_0, s_m] = [t_{k-1}, t_k]$ ist f eine Verbindungsstrecke,

$$f(s) = f(s_0) + \frac{s - s_0}{s_m - s_0} (f(s_m) - f(s_0)), \quad s \in [s_0, s_m].$$

Also gilt

$$\begin{aligned} \sum_{i=1}^m \|f(s_i) - f(s_{i-1})\|_2 &= \sum_{i=1}^m \left\| \frac{s_i - s_{i-1}}{s_m - s_0} (f(s_m) - f(s_0)) \right\|_2 \\ &= \|f(s_m) - f(s_0)\|_2 \underbrace{\sum_{i=1}^m \left| \frac{s_i - s_{i-1}}{s_m - s_0} \right|}_{=1}. \end{aligned}$$

□

Beispiel 2.27 (Dyadische Zerlegung). Sei $f : [a, b] \rightarrow \mathbb{R}^d$ eine Kurve. Für $n \in \mathbb{N}$ bezeichne Z_n die äquidistante Zerlegung mit Parameterpunkten

$$t_k = a + \frac{k}{2^n} (b - a), \quad k = 0, \dots, 2^n.$$

Dann gilt

$$Z_0 \subseteq Z_1 \subseteq Z_2 \subseteq Z_3 \subseteq \dots$$

Aus Lemma 2.25 folgt, dass die durch

$$\ell_n := L(f; Z_n)$$

definierte Folge in $\mathbb{R}$ monoton wachsend ist. Also ist (ℓ_n) entweder konvergent oder uneigentlich gegen $+\infty$ konvergent. Beispiel: Im Fall $f : [0, 2\pi] \rightarrow \mathbb{R}^2$, $f(t) := (\cos(t), -\sin(t))$ sind wir in der Situation von Beispiel 2.19 und es gilt $\ell_n = 2 \cdot 2^n \sin(\frac{\pi}{2^n}) \rightarrow 2\pi$.

Wir definieren nun die Länge einer Kurve:

Definition 2.28 (Länge und Rektifizierbarkeit). Für eine Kurve $f : [a, b] \rightarrow \mathbb{R}^d$ setzen wir

$$L(f) := \sup_{Z \in \mathcal{Z}[a, b]} L(f; Z).$$

f heißt rektifizierbar, falls $L(f) < \infty$. Wir nennen $L(f)$ die (euklidische) Länge von f .

Wir überprüfen, dass (2.3) erfüllt ist:

Lemma 2.29 (Additivität bzgl. Zusammensetzung). *Sei $f : [a, b] \rightarrow \mathbb{R}^d$ eine Kurve und $s \in (a, b)$. Dann gilt*

$$L(f) = L(f|[a, s]) + L(f|[s, b]).$$

Insbesondere ist f rektifizierbar, genau dann wenn sowohl $f|[a, s]$ als auch $f|[s, b]$ rektifizierbar sind.

Beweis. Die Aussage über die Rektifizierbarkeit folgt direkt aus der behaupteten Identität. Wir beweisen die Identität.

1. Es gilt

$$Z_1 \in \mathcal{Z}[a, s] \wedge Z_2 \in \mathcal{Z}[s, b] \quad \Rightarrow \quad Z_1 \cup Z_2 \in \mathcal{Z}[a, b].$$

Mit Lemma 2.23 folgt

$$L(f; Z_1) + L(f; Z_2) = L(f; Z_1 \cup Z_2) \leq L(f).$$

Mit Supremumsbildung über $Z_1 \in \mathcal{Z}[a, s]$ und $Z_2 \in \mathcal{Z}[s, b]$ folgt

$$L(f|[a, s]) + L(f|[s, b]) \leq L(f).$$

2. Es gilt

$$Z \in \mathcal{Z}[a, b] \quad \Rightarrow \quad \begin{cases} Z_1 := \{s\} \cup (Z \cap [a, s]) \in \mathcal{Z}[a, s] \\ Z_2 := \{s\} \cup (Z \cap [s, b]) \in \mathcal{Z}[s, b] \end{cases}$$

Mit Lemma 2.23 folgt

$$L(f; Z) \leq L(f; Z \cup \{s\}) = L(f; Z_1) + L(f; Z_2) \leq L(f|[a, s]) + L(f|[s, b]).$$

Supremumsbildung in Z führt auf

$$L(f) \leq L(f|[a, s]) + L(f|[s, b]).$$

□

Wir überprüfen, dass (2.4) erfüllt ist:

Lemma 2.30 (Länge eines Streckenzuges). *Sei $f : [a, b] \rightarrow \mathbb{R}^d$ ein Streckenzug mit Parameterpunkten $t_0, \dots, t_n \in Z_1$ und Streckenpunkten $x_0, \dots, x_n$. Dann gilt*

$$L(f) = \sum_{k=1}^n \|x_k - x_{k-1}\|_2.$$

Beweis. Für die Zerlegung $Z_1 := \{t_0, \dots, t_n\}$ gilt

$$L(f; Z_1) = \sum_{k=1}^n \|f(t_k) - f(t_{k-1})\|_2 = \sum_{k=1}^n \|x_k - x_{k-1}\|_2,$$

und damit

$$L(f) \geq L(f; Z_1).$$

Wir beweisen nun $L(f) \leq L(f; Z_1)$ durch ein Widerspruchsargument: Annahme $L(f) > L(f; Z_1)$. Dann existiert $Z_2 \in \mathcal{Z}[a, b]$ mit

$$L(f; Z_2) > L(f; Z_1).$$

Setze $Z_3 := Z_2 \cup Z_1$. Aus der Monotonie folgt

$$L(f; Z_3) \geq L(f; Z_2) > L(f; Z_1).$$

Andererseits, folgt aus Lemma 2.26 wegen $Z_1 \subseteq Z_3$

$$L(f; Z_3) = L(f; Z_1),$$

und damit ein Widerspruch. □

Wir überprüfen, dass (2.5) erfüllt ist:

Lemma 2.31. *Sei $f : [a, b] \rightarrow \mathbb{R}^d$ eine Kurve. Es bezeichne g die Verbindungsstrecke von $f(a)$ nach $f(b)$. Dann gilt*

$$L(f) = \|f(b) - f(a)\|_2 \quad \Rightarrow \quad \text{Bild}(f) = \text{Bild}(g).$$

Beweis. Siehe Übung. □

10.5.2020

#20

10.5.2020

Satz 2.32 (Rektifizierbarkeit von C^1 -Kurven). *Sei $f : [a, b] \rightarrow \mathbb{R}^d$ eine stetig differenzierbare Kurve. Dann ist f rektifizierbar und es gilt¹²*

$$L(f) = \lim_{n \rightarrow \infty} \sum_{k=1}^n \|f(t_k) - f(t_{k-1})\|_2, \quad t_k := a + \frac{k}{n}(b-a).$$

Für den Beweis benötigen wir nachfolgendes technisches Lemma:

Lemma 2.33. *Sei $f : [a, b] \rightarrow \mathbb{R}^d$ eine stetig differenzierbare Kurve. Sei $\varepsilon > 0$. Dann existiert $\delta > 0$, so dass für Zerlegungen $Z_1, Z_2 \in \mathcal{Z}[a, b]$ mit $|Z_1|, |Z_2| < \delta$ die Abschätzung*

$$|L(f; Z_1) - L(f; Z_2)| < \varepsilon$$

gilt.

Beweis. Sei $\varepsilon > 0$. Die Funktion $t \mapsto \|f'(t)\|_\infty$ ist auf $[a, b]$ stetig und damit auch gleichmäßig stetig (da $[a, b]$ kompakt). Also existiert $\delta > 0$, so dass für alle $\xi, \eta \in [a, b]$,

$$(2.6) \quad |\xi - \eta| < \delta \Rightarrow \|f'(\xi) - f'(\eta)\|_\infty < \frac{\varepsilon}{2\sqrt{d}(b-a)}.$$

Es gelte $|Z_1|, |Z_2| < \delta$. Es genügt zu zeigen: Mit $Z = Z_1 \cup Z_2$ gilt

$$|L(f; Z_1) - L(f; Z)| < \varepsilon/2 \text{ und } |L(f; Z_2) - L(f; Z)| < \varepsilon/2,$$

Wir zeigen dies nur für Z_1 . Sei also $Z_1 = \{t_0 < \dots < t_n\}$ und setze $I_k := [t_{k-1}, t_k]$.

¹²Korrektur am 23.5.2020.

1. (Lokale Abschätzung). Behauptung: Es gilt

$$|L(f; Z_1 \cap I_k) - L(f; Z \cap I_k)| \leq (t_k - t_{k-1})\sqrt{d}\varepsilon.$$

Argument: Sei $Z \cap I_k = \{s_0 := t_{k-1} < s_1 < \dots < s_m := t_k\}$. Es gilt

$$\begin{aligned} L(f; Z_1 \cap I_k) &= \|f(s_m) - f(s_0)\|_2 \\ &= \sum_{\ell=1}^m (s_\ell - s_{\ell-1}) \left\| \frac{f(s_m) - f(s_0)}{s_m - s_0} \right\|_2, \end{aligned}$$

sowie

$$\begin{aligned} L(f; Z \cap I_k) &= \sum_{\ell=1}^m \|f(s_\ell) - f(s_{\ell-1})\|_2 \\ &= \sum_{\ell=1}^m (s_\ell - s_{\ell-1}) \left\| \frac{f(s_\ell) - f(s_{\ell-1})}{s_\ell - s_{\ell-1}} \right\|_2, \end{aligned}$$

Also

$$(2.7) \quad L(f; Z_1 \cap I_k) - L(f; Z \cap I_k) = \sum_{\ell=1}^m (s_\ell - s_{\ell-1}) \underbrace{\left(\left\| \frac{f(s_m) - f(s_0)}{s_m - s_0} \right\|_2 - \left\| \frac{f(s_\ell) - f(s_{\ell-1})}{s_\ell - s_{\ell-1}} \right\|_2 \right)}_{=: T_\ell}.$$

Mit der inversen Dreiecksungleichung und der Abschätzung $\|\cdot\|_2 \leq \sqrt{d}\|\cdot\|_\infty$ folgt

$$|T_\ell| \leq \sqrt{d} \left\| \left(\frac{f(s_m) - f(s_0)}{s_m - s_0} \right) - \left(\frac{f(s_\ell) - f(s_{\ell-1})}{s_\ell - s_{\ell-1}} \right) \right\|_\infty.$$

Fixiere eine Koordinate $i = 1, \dots, d$. Wegen des Mittelwertsatzes existieren $\xi, \eta \in (s_0, s_m)$ mit

$$e_i \cdot \left(\frac{f(s_m) - f(s_0)}{s_m - s_0} - \frac{f(s_\ell) - f(s_{\ell-1})}{s_\ell - s_{\ell-1}} \right) = f'_i(\xi) - f'_i(\eta).$$

Wegen $|\xi - \eta| \leq t_k - t_{k-1} \leq |Z| < \delta$ folgt nun wegen (2.6) $|f'_i(\xi) - f'_i(\eta)| < \frac{\varepsilon}{2\sqrt{d}(b-a)}$. Also insgesamt

$$|T_\ell| \leq \frac{\varepsilon}{2(b-a)}.$$

Aus (2.7) folgt nun

$$|L(f; Z_1 \cap I_k) - L(f; Z \cap I_k)| \leq \sum_{\ell=1}^m (s_\ell - s_{\ell-1}) |T_\ell| \leq \frac{t_k - t_{k-1}}{b-a} \frac{\varepsilon}{2}.$$

2. (Globale Abschätzung). Aus der Additivität folgt

$$\begin{aligned} |L(f; Z_1) - L(f; Z)| &= \left| \sum_{k=1}^n (L(f; Z_1 \cap I_k) - L(f; Z \cap I_k)) \right| \\ &\leq \sum_{k=1}^n |L(f; Z_1 \cap I_k) - L(f; Z \cap I_k)| \\ &\stackrel{1.}{\leq} \frac{\varepsilon}{2(b-a)} \sum_{k=1}^n (t_k - t_{k-1}) = \varepsilon/2. \end{aligned}$$

□

Beweis von Satz 2.32. Wir bezeichnen mit

$$Z_n := \{t_k = a + \frac{k}{n}(b-a) : k = 0, \dots, n\}.$$

die äquidistante Zerlegung von $[a, b]$ mit Feinheit $|Z_n| = \frac{b-a}{n}$.

1. Behauptung: Die Folge $L_n := L(f; Z_n)$ ist eine Cauchyfolge. Argument: Sei $\varepsilon > 0$ und wähle $\delta > 0$ gemäß Lemma 2.33. Wähle $N \in \mathbb{N}$ mit $\frac{b-a}{N} < \delta$. Dann gilt

$$|Z_n|, |Z_m| < \delta \quad \text{für alle } n, m \in \mathbb{N} \text{ mit } n, m \geq N,$$

und aus Lemma 2.33 folgt

$$|L_n - L_m| < \varepsilon \quad \text{für alle } n, m \geq N.$$

2. Behauptung:

$$L(f) = \lim_{n \rightarrow \infty} L_n =: L.$$

Der Grenzwert existiert wegen Schritt 1. Die Abschätzung “ $\geq$ ” ist trivial, da $L(f) \geq L(f; Z_n)$ für alle $n \in \mathbb{N}$. Annahme $L(f) > L$. Dann existiert $Z \in \mathcal{Z}[a, b]$ mit $L(f; Z) > L$. Sei nun $\varepsilon := \frac{L(f; Z) - L}{2}$ und wähle $\delta > 0$ gemäß Lemma 2.33 und $n \in \mathbb{N}$ mit $\frac{b-a}{n} < \delta$. Dann gilt mit $\hat{Z} := Z \cup Z_n$ zum einen

$$(2.8) \quad L(f; \hat{Z}) - L \geq L(f; Z) - L = 2\varepsilon;$$

andererseits gilt wegen $|\hat{Z}| \leq |Z_n| < \delta$ mit Lemma 2.33,

$$|L(f; \hat{Z}) - L(f; Z_m)| < \varepsilon \quad \text{für alle } m \geq n.$$

Ein Grenzübergang $m \rightarrow \infty$ liefert

$$|L(f; \hat{Z}) - L| < \varepsilon,$$

was im Widerspruch zu (2.8) steht.

□

Bemerkung 2.34. Wir werden später zeigen, dass für eine stetig differenzierbare Kurve $f : [a, b] \rightarrow \mathbb{R}^d$ die Länge mit

$$L(f) = \int_a^b \|f'(t)\|_2 dt$$

berechnet werden kann.

Korollar 2.35. Sei $f : [a, b] \rightarrow \mathbb{R}^d$ eine stückweise stetig differenzierbare Kurve, d.h. f ist eine Kurve und es existiert $\{t_0 < \dots < t_n\} \in \mathcal{Z}[a, b]$, so dass $f|_{[t_{k-1}, t_k]}$ stetig differenzierbar ist. Dann ist f rektifizierbar.

Beweis. Beweis in der Übung. □

Beispiel 2.36 (Nicht rektifizierbare Kurve). *Betrachte*

$$f : [0, 1] \rightarrow \mathbb{R}^d, \quad f(t) := \begin{cases} (t, t \cos(\frac{\pi}{t})) & \text{für } t > 0, \\ (0, 0) & \text{für } t = 0. \end{cases}$$

In der Übung zeigen wir, dass f eine Kurve ist, die nicht rektifizierbar ist.

Satz 2.37 (Parametertransformation). *Sei $f : [a, b] \rightarrow \mathbb{R}^d$ eine Kurve. Sei $\varphi : [a', b'] \rightarrow [a, b]$ ein Homöomorphismus. Dann ist $f \circ \varphi : [a', b'] \rightarrow \mathbb{R}^d$ eine Kurve und es gilt:*

$$L(f) = L(f \circ \varphi).$$

Insbesondere ist f genau dann rektifizierbar, wenn g rektifizierbar ist.

Beweis. Es genügt die Identität zu zeigen. Zunächst gilt: φ stetige Bijektion $\Rightarrow \varphi$ streng monoton. Wir behandeln den streng monoton wachsenden Fall. Der fallende Fall wird in der Übung besprochen.

- Es gilt

$$\begin{aligned} Z \in \mathcal{Z}[a, b] & \Leftrightarrow \varphi^{-1}(Z) \in \mathcal{Z}[a', b'] \\ \varphi(Z') \in \mathcal{Z}[a, b] & \Leftrightarrow Z' \in \mathcal{Z}[a', b'], \end{aligned}$$

da φ, φ^{-1} streng monoton und bijektiv sind.

- Sei $Z' = \{t_0 < \dots < t_n\} \in \mathcal{Z}[a', b']$ und die Zerlegung von $[a, b]$

$$Z := \varphi(Z') = \{s_k := \varphi(t_k) : k = 0, \dots, n\}.$$

Dann gilt

$$\begin{aligned} L(f; Z) &= \sum_{k=1}^{\infty} \|f(s_k) - f(s_{k-1})\|_2 = \sum_{k=1}^{\infty} \|(f \circ \varphi)(t_k) - (f \circ \varphi)(t_{k-1})\|_2 \\ &= L(f \circ \varphi; Z') \end{aligned}$$

Supremumsbildung führt auf

$$L(f) \leq L(f \circ \varphi).$$

Das selbe Argument mit vertauschten Rollen von f und $f \circ \varphi$ führt auf

$$L(f \circ \varphi) \leq L(f).$$

□

Bemerkung 2.38. *Bei der Definition der Länge $L(f)$ und approximativen Länge $L(f; Z)$ kann die euklidische Norm $\|\cdot\|_2$ durch eine beliebige Norm $\|\cdot\|_*$ auf $\mathbb{R}^d$ ersetzt werden. Alle Aussagen gelten auch in diesem Fall - nur die Invarianz der Länge unter euklidischen Transformationen kann hierbei verloren gehen. Der Begriff der Differenzierbarkeit und der Länge von Kurven kann auch (ganz analog) für Kurven in Banachräumen eingeführt werden.*

Kontrollfrage 6.1. Betrachten Sie die Kurven

$$f_1 : [0, 2\pi] \rightarrow \mathbb{R}^2,$$

$$f_1(t) = (\sin(t), \cos(t)),$$

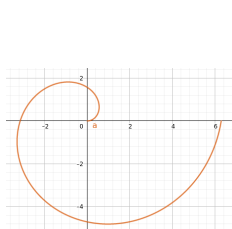
$$f_2 : [0, 2\pi] \rightarrow \mathbb{R}^2,$$

$$f_2(t) = (t \cos(t), t \sin(t)),$$

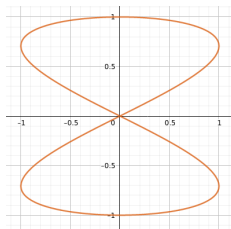
$$f_3 : [0, 2\pi] \rightarrow \mathbb{R}^2,$$

$$f_3(t) = (\sin(2t), \cos(t)),$$

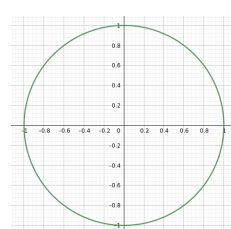
sowie die nachfolgenden Abbildungen von Spuren



(a)



(b)



(c)

Welche Zuordnung ist richtig?

(A) $f_1 \leftrightarrow (a), \quad f_2 \leftrightarrow (b), \quad f_3 \leftrightarrow (c)$

×

(B) $f_1 \leftrightarrow (a), \quad f_2 \leftrightarrow (c), \quad f_3 \leftrightarrow (b)$

×

(C) $f_1 \leftrightarrow (b), \quad f_2 \leftrightarrow (a), \quad f_3 \leftrightarrow (c)$

×

(D) $f_1 \leftrightarrow (b), \quad f_2 \leftrightarrow (c), \quad f_3 \leftrightarrow (a)$

×

(E) $f_1 \leftrightarrow (c), \quad f_2 \leftrightarrow (a), \quad f_3 \leftrightarrow (b)$

✓: Kann auch mit Geogebra überprüft werden.

(F) $f_1 \leftrightarrow (c), \quad f_2 \leftrightarrow (b), \quad f_3 \leftrightarrow (a)$

×

Kontrollfrage 6.2. Sei $f : [0, 1] \rightarrow \mathbb{R}^d$ ein Streckenzug. Welche der nachfolgenden Aussagen sind dann im Allgemeinen richtig?

(A) f ist rektifizierbar.

✓: Siehe Lemma 2.30

(B) $L(f) \leq \|f(1) - f(0)\|_2.$

×: „ $\geq$ “ wegen Definition als Supremum wäre richtig.

(C) $L(f) = \|f(1) - f(0)\|_2$, falls $\text{Bild}(f) = [[f(0), f(1)]]$.

×: Siehe Bem. 2.20.

Kontrollfrage 6.3. Sei $f : [-1, 1] \rightarrow \mathbb{R}^2$, $f(t) = (t, |t|)$. Welche der nachfolgenden Aussagen sind richtig?

(A) f ist stetig differenzierbar,

×: Siehe Lemma 2.7.

(B) f ist rektifizierbar.

✓: Siehe Korollar 2.35

(C) $L(f) = 2$.

×: $2\sqrt{2}$ ist die Länge.

Kontrollfrage 6.4. Sei $f : [-1, 2] \rightarrow \mathbb{R}^d$ eine rektifizierbare Kurve und

$$g : [-1, 2] \rightarrow \mathbb{R}^d, \quad g(t) := \begin{cases} f(0) + t(f(1) - f(0)) & \text{falls } t \in [0, 1] \\ f(t) & \text{sonst.} \end{cases}$$

Welche der nachfolgenden Aussagen sind richtig?

(A) g ist eine Kurve.

✓: Stetigkeit

(B) g ist rektifizierbar.

✓: Siehe Lemma 2.29.

(C) $L(f) \leq L(g)$.

×: Im Allgemeinen falsch. (Siehe auch Lemma 2.29 und Lemma 2.31.)

(D) $L(g) \leq L(f)$.

✓: Lemma 2.29 und Definition als Supremeum.

Kontrollfrage 6.5. Seien $f, g : [a, b] \rightarrow \mathbb{R}^d$ stetig differenzierbare Kurven mit $f(a) = g(a)$, $f(b) = g(b)$. Welche der nachfolgenden Aussagen sind richtig?

(A) Es gilt: $\text{Bild}(f) = \text{Bild}(g) \Rightarrow L(f) = L(g)$.

×: Siehe Bem. 2.20.

(B) Es gilt $L(f) = L(g)$, falls sowohl f' als auch g' konstant sind.

✓: Siehe Lemma 2.13.

(C) Ist f eine Verbindungsstrecke, dann gilt $L(f) \leq L(g)$.

✓: Definition von L als Supremum.

Kontrollfrage 6.6. Sei $f : [a, b] \rightarrow \mathbb{R}^d$ eine rektifizierbare Kurve, $s \in (a, b)$. Welche der nachfolgenden Aussagen sind dann im Allgemeinen richtig?

(A) $L(f|[a, s]) + L(f|[s, b]) = L(f)$.

✓: Siehe Lemma 2.29.

(B) $L(f|[a, s]) < L(f)$.

✗: Gegenbeispiel: konstantes f .

Kontrollfrage 6.7. Sei $f : [a, b] \rightarrow \mathbb{R}^d$ eine Kurve, $Q \in O(d)$ und $c \in \mathbb{R}^d$. Betrachte

$$g : [a, b] \rightarrow \mathbb{R}^d, \quad g(t) := Qf(t) + c.$$

Welche der nachfolgenden Aussagen sind richtig?

(A) g ist eine Kurve.

✓: Stetigkeit überträgt sich.

(B) f ist rektifizierbar $\Leftrightarrow g$ ist rektifizierbar.

✓: Siehe Lemma 1.67.

(C) Sind f und g rektifizierbar, dann gilt $L(g) = L(f)$.

✓: Siehe Lemma 1.67.

Kontrollfrage 6.8. Sei $f : [0, 1] \rightarrow \mathbb{R}^d$ eine rektifizierbare Kurve. Für $n \in \mathbb{N}_+$ bezeichne

$$Z_n := \{t_k := \frac{k}{n} : k = 0, \dots, n\}$$

eine äquidistante Zerlegung, sowie $f_n : [0, 1] \rightarrow \mathbb{R}^d$ den Streckenzug mit Parameterpunkten $t_k \in Z_n$ und Streckenpunkten $f(t_k)$. Welche der folgenden Aussagen sind dann richtig?

(A) $L(f_n) \geq \|f(1) - f(0)\|_2$.

✓: Siehe Lemma 2.25.

(B) $L(f_n) \geq L(f)$.

✗: I.A. wäre „ $\leq$ “ wegen der Definition von L als Supremum richtig.

(C) $L(f_n) \leq L(f)$.

✓: Wegen Definition von L als Supremum.

(D) Die Folge $\left(L(f_n)\right)_{n \in \mathbb{N}}$ ist Cauchyfolge, falls f stetig differenzierbar ist¹³.

¹³siehe Korrektur der Angabe via Rundmail

✓: Lemma 2.32 anwendbar.

(E) $L(f_n) \rightarrow L(f)$ für $n \rightarrow \infty$, falls f stetig differenzierbar ist¹⁴.

✓: ist Lemma 2.32 anwendbar.

10.5.2020

#21

17.5.2020

2.2 Totale Ableitung

Ziel:

- Totale Differenzierbarkeit für Abbildungen zwischen normierten Räumen
- Rechenregeln
- Schrankensatz

In diesem Kapitel bezeichnen $(X, \|\cdot\|_X)$ und $(Y, \|\cdot\|_Y)$ normierte Räume (über $\mathbb{R}$ oder $\mathbb{C}$) und $\Omega \subseteq X$ eine offene Menge. Wir führen den Begriff der Differenzierbarkeit für Funktionen $f : \Omega \subseteq X \rightarrow Y$ ein.

Wir haben in Analysis 1 die Differenzierbarkeit reeller Funktionen $f : \Omega \subseteq \mathbb{R} \rightarrow \mathbb{R}$ einer Veränderlichen diskutiert (das entspricht dem Fall $X = Y = \mathbb{R}$): f ist im Punkt $x \in \Omega$ differenzierbar, falls der Grenzwert des Differenzenquotienten

$$(2.9) \quad \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

existiert. Der Differenzenquotient und die Grenzwertbildung sind auch dann wohlgeformt, wenn wir den Bildraum durch einen allgemeinen normierten Vektorraum $(Y, \|\cdot\|_Y)$ ersetzen. Daher kann diese Definition ganz analog auch auf den Fall einer Funktion einer Veränderlichen mit Werten in einem mehrdimensionalen Bildraum ausgedehnt werden. Bei der Diskussion differenzierbarer Kurven sind wir so vorgegangen.

Im Fall einer Funktion mehrerer Veränderlichen ($\dim(X) > 1$) ist der Differenzenquotient nicht wohlgeformt, da man nicht durch einen Vektor teilen kann! Die Verallgemeinerung des Ableitungsbegriffes für Abbildungen auf normierten Räumen basiert auf der Weierstraßschen Zerlegungsformel, die wir für Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$ in AN10, Satz 7.4 eingeführt haben. Sie besagt, dass die Existenz des Grenzwertes in (2.9) und die Konvergenz des Differenzenquotienten gegen ein $a \in \mathbb{R}$ äquivalent dazu sind, dass der Rest $r(h)$ in der Zerlegungsformel

$$f(x+h) - (f(x) + ah) = r(h)$$

gegen Null für $h \rightarrow 0$ konvergiert. In diesem Fall haben wir $a \in \mathbb{R}$ die Ableitung von f im Punkt x genannt. Für die nachfolgende Diskussion ist es wichtig, die Ableitung nicht als Zahl $a \in \mathbb{R}$, sondern als lineare Abbildung

$$T : \mathbb{R} \rightarrow \mathbb{R}, \quad h \mapsto ah$$

¹⁴sieh Korrektur der Angabe via Rundmail

zu betrachten. Für $a \neq 0$ kann der Funktionsgraph von $y \mapsto f(x) + a(y - x)$ als Tangente an den Funktionsgraphen von f im Punkt x verstanden werden.

Wir werden Ableitungen von Abbildungen zwischen zwei normierten Räumen X, Y als stetige, lineare Abbildungen von X nach Y definieren. Wir bezeichnen mit

$$\mathcal{L}(X; Y)$$

den Vektorraum aller stetigen, linearen Abbildungen $T : X \rightarrow Y$. Durch

$$\|T\| := \sup_{\|x\|_X \leq 1} \|T(x)\|_Y, \quad T \in \mathcal{L}(X; Y)$$

wird auf $\mathcal{L}(X; Y)$ ein Norm (die Operatornorm) definiert (siehe Hausaufgabe 4.2).

Beim ersten Lesen der nachfolgenden Definitionen und Aussagen stellt man sich am besten vor, dass $X = \mathbb{R}^n$ und $Y = \mathbb{R}^m$.

Definition 2.39 (Totale Ableitung). *Sei $\Omega \subseteq X$ eine offene Menge. Sei $f : \Omega \rightarrow Y$ eine Abbildung.*

(a) *f heißt differenzierbar in $x \in \Omega$, falls es eine stetige, lineare Abbildung $T : X \rightarrow Y$ und einen Rest $r : \{h \in X : x + h \in \Omega\} \rightarrow Y$ gibt, so dass*

$$(2.10) \quad f(x + h) - (f(x) + T(h)) = r(h)$$

für alle h so, dass $x + h \in \Omega$ und

$$(2.11) \quad \lim_{\substack{h \rightarrow 0 \\ h \neq 0, h+x \in \Omega}} \frac{\|r(h)\|_Y}{\|h\|_X} = 0.$$

In diesem Fall schreiben wir $Df(x) := T$ und nennen $Df(x)$ die Ableitung von f in x .

(b) *f heißt differenzierbar, falls f in jedem Punkt $x \in \Omega$ differenzierbar ist. In diesem Fall heißt*

$$Df : \Omega \subseteq X \rightarrow \mathcal{L}(X; Y), \quad x \mapsto Df(x)$$

die Ableitung von f .

(c) *f heißt stetig differenzierbar, falls f differenzierbar ist und die Abbildung $Df : \Omega \subseteq X \rightarrow \mathcal{L}(X; Y)$ stetig ist. Wir bezeichnen mit $C^1(\Omega; Y)$ die Menge aller stetig differenzierbaren Abbildungen von $\Omega \subseteq X$ nach Y .*

Bemerkung 2.40. • *Wir nennen (2.11) die Restgliedabschätzung der Weierstraßschen Zerlegungsformel. Sie ist äquivalent zu:*

$$\lim_{k \rightarrow \infty} \frac{\|r(h_k)\|_Y}{\|h_k\|_X} = 0$$

für alle Folgen $(h_k)_{k \in \mathbb{N}} \subseteq \{h \in X \setminus \{0\} : h + x \in \Omega\}$ mit $h_k \rightarrow 0$ (d.h. $\|h_k\|_X \rightarrow 0$). Statt (2.11) schreiben wir von nun an einfach $\lim_{h \rightarrow 0} \frac{\|r(h)\|_Y}{\|h\|_X} = 0$. Die Aussage (2.11) bedeutet, dass der Rest r "schneller als linear verschwindet für $h \rightarrow 0$ ".

- Die Schreibweise (2.11) ist äquivalent zu:

$$r(h) = \|h\|_X \rho(h), \quad \lim_{h \rightarrow 0} \rho(h) = 0.$$

- Formel (2.10) heißt Weierstraßsche Zerlegungsformel. Im Fall $X = Y = \mathbb{R}$ ist die Definition konsistent mit Satz 7.4 aus AN10. Im Fall $X = \mathbb{R}$ und $Y = \mathbb{R}^m$ erhalten wir Aussage 2.1 zu differenzierbaren Kurven.
- Die Ableitung $Df(x)$ ist (falls existent) eindeutig bestimmt, was direkt aus der Forderung folgt, dass der Rest “schneller als linear verschwindet”.
- Die Abbildung

$$y \mapsto f(x) + Df(x)(y - x)$$

ist die bestmögliche Approximation von f bei x durch eine affine Abbildung.

- Statt differenzierbar sagt man auch Fréchet-differenzierbar oder total differenzierbar. $Df(x)$ nennt man auch Fréchetableitung oder totale Ableitung.

Proposition 2.41 (Notwendige Bedingung). Ist $f : \Omega \subseteq X \rightarrow Y$ differenzierbar in $x \in \Omega$, dann ist f stetig in x .

Beweis. Sei $(x_k) \subseteq \Omega$ eine Folge mit Grenzwert $x \in \Omega$. Dann gilt

$$f(x_k) = f(x) + Df(x)(x_k - x) + r(x_k - x) \rightarrow f(x).$$

□

Beispiele 2.42. (a) **Konstante Funktion.** Seien $c \in Y$ und

$$f : X \rightarrow Y, x \mapsto c.$$

Es gilt

$$r(h) = f(x+h) - f(x) = 0 \quad \text{für alle } x, h \in X,$$

also $Df(x) = 0$ für alle $x \in X$; f ist stetig differenzierbar.

(b) **Affine Funktion.** Seien $c \in Y$, $L : X \rightarrow Y$ linear,

$$f : X \rightarrow Y, \quad f(x) := c + L(x).$$

Es gilt

$$r(h) = f(x+h) - (f(x) + L(h)) = 0 \quad \text{für alle } x, h \in X,$$

also ist f differenzierbar mit $Df(x) = L$, falls L stetig ist. In diesem Fall gilt auch: Da die Abbildung

$$X \rightarrow \mathcal{L}(X; Y), \quad x \mapsto L$$

konstant ist, ist f stetig differenzierbar.

Beispiel: Seien $A \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^m$

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad f(x) := c + Ax \quad \Rightarrow \quad Df(x)(h) = Ah.$$

(c) **Quadratische Funktion.** Sei X ein Hilbertraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$, $Y := \mathbb{R}$.

$$f : X \rightarrow \mathbb{R}, \quad f(x) := \langle x, x \rangle.$$

Es gilt

$$r(h) = f(x+h) - f(x) = \langle x+h, x+h \rangle - \langle x, x \rangle = \langle x, h \rangle + \langle x, h \rangle + \langle h, h \rangle \quad \text{für alle } x, h \in X.$$

Es gilt

$$|\langle h, h \rangle| \leq \|h\|_X^2.$$

Also ist f differenzierbar mit

$$Df(x) : Y \rightarrow \mathbb{R}, \quad Df(x)(y) := 2\langle x, y \rangle.$$

Die Abbildung $x \mapsto Df(x)$ erfüllt

$$\begin{aligned} \|Df(x) - Df(x')\| &= \sup_{y \in X, \|y\|_X \leq 1} |Df(x)(y) - Df(x')(y)| = 2 \sup_{y \in X, \|y\|_X \leq 1} |\langle x - x', y \rangle| \\ &\leq 2\|x - x'\|_X. \end{aligned}$$

also ist $Df : X \rightarrow \mathcal{L}(X; \mathbb{R})$ stetig und f stetig differenzierbar.

Beispiel: Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch positiv definit.

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad f(x) := x \cdot Ax \quad \Rightarrow \quad Df(x)(h) = 2x \cdot Ah.$$

Proposition 2.43 (Summen und Produktregel). Seien $f, g : \Omega \subseteq X \rightarrow Y$ und $\varphi : \Omega \subseteq X \rightarrow \mathbb{K}$ differenzierbar im Punkt $a \in \Omega$. Dann gelten:

(a) (Summenregel). $f + g : \Omega \subseteq X \rightarrow Y$, $x \mapsto f(x) + g(x)$ ist differenzierbar in a und es gilt

$$D(f + g)(a) = Df(a) + Dg(a).$$

(b) (Produktregel). $u : \Omega \subseteq X \rightarrow Y$, $u(x) := \varphi(x)f(x)$ ist differenzierbar in a und es gilt

$$Du(a)(h) = D\varphi(a)(h)f(a) + \varphi(a)Df(a)(h) \quad \text{für alle } h \in X.$$

Beweis. Siehe Übung. □

	17.5.2020
#22	17.5.2020

Proposition 2.44 (Kettenregel). Seien X, Y, Z normierte Räume. Seien $U \subseteq X$ und $V \subseteq Y$ offene Mengen,

$$f : U \subseteq X \rightarrow Y, \quad g : V \subseteq Y \rightarrow Z, \quad \text{Bild}(f) \subseteq V.$$

Sei f im Punkt $x \in U$ und g im Punkt $y := f(x)$ differenzierbar. Dann ist $g \circ f : U \rightarrow Z$ im Punkt x differenzierbar und es gilt

$$D(g \circ f)(x) = Dg(f(x)) \circ Df(x).$$

Beweis. Nach Voraussetzung gilt

$$\begin{aligned} f(x+h) &= f(x) + Df(x)(h) + \|h\|_X \rho_1(h) & \text{mit } \lim_{h \rightarrow 0} \rho_1(h) = 0, \\ g(y+\tilde{h}) &= g(y) + Dg(y)(\tilde{h}) + \|\tilde{h}\|_Y \rho_2(\tilde{h}) & \text{mit } \lim_{\tilde{h} \rightarrow 0} \rho_2(\tilde{h}) = 0. \end{aligned}$$

Mit $\tilde{h} := Df(x)(h) + \|h\|_X \rho_1(h)$ folgt

$$\begin{aligned} (g \circ f)(x+h) &= (g \circ f)(x) + (Dg(y) \circ Df(x))(h) \\ &\quad + \underbrace{\|h\|_X Dg(y)(\rho_1(h)) + \|\tilde{h}\|_Y \rho_2(\tilde{h})}_{=: R(h)}. \end{aligned}$$

Da $Df(x) : X \rightarrow Y$ stetig ist, gilt

$$\begin{aligned} \|\tilde{h}\|_Y &\leq \|Df(x)(h)\|_Y + \|h\|_X \|\rho_1(h)\|_Y \\ &\leq \|h\|_X \left(\|Df(x)\| + \|\rho_1(h)\|_Y \right), \end{aligned}$$

und wir folgern

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

und damit die Differenzierbarkeit von $g \circ f$ in x . □

Beispiel 2.45. Betrachte $u : \mathbb{R}^d \rightarrow \mathbb{R}$, $x \mapsto \|x\|_2 = (g \circ f)(x)$ mit $f(x) = x \cdot x$ und $g(y) = \sqrt{y}$. Für alle $\|x\|_2 > 0$ gilt dann

$$\begin{aligned} Du(x)(h) &= (Dg(f(x)) \circ Df(x))(h) = \underbrace{Dg(f(x))}_{= \frac{1}{2}(f(x))^{-\frac{1}{2}}} \underbrace{(Df(x)(h))}_{= 2x \cdot h} \\ &= \frac{x \cdot h}{\|x\|_2}. \end{aligned}$$

Für eine stetig differenzierbare Funktion $f : [a, b] \rightarrow \mathbb{R}$ haben wir die Abschätzung $\|f(t) - f(s)\| \leq \max_{\xi \in [a, b]} |f'(\xi)| |t - s|$ bewiesen. Dies gilt auch im Mehrdimensionalen:

Satz 2.46 (Schränkensatz). Sei $f : \Omega \subseteq X \rightarrow Y$ stetig differenzierbar, $K \subseteq \Omega$ eine konvexe, kompakte Menge. Dann gilt für alle $x, y \in K$,

$$\|f(x) - f(y)\|_Y \leq C \|x - y\|_X$$

mit

$$C := \sup_{\xi \in K} \|Df(\xi)\| < \infty.$$

Hierbei bezeichnet $\|Df(\xi)\|$ die Operatornorm von $Df(\xi) \in \mathcal{L}(X; Y)$

Beweis.

1. Zunächst gilt $C < \infty$, denn: $K \rightarrow \mathbb{R}$, $\xi \mapsto \|Df(\xi)\|$ ist stetig und K kompakt. Sei $u : [0, 1] \rightarrow X$, $u(t) := x + t(y - x)$. Es gilt $\text{Bild}(u) \subseteq K$, da K konvex. Für $\varepsilon > 0$ betrachte

$$F_\varepsilon : [0, 1] \rightarrow \mathbb{R}, \quad F_\varepsilon(t) := \|f(u(t)) - f(u(0))\|_Y - t(C + \varepsilon)\|x - y\|_X.$$

Es gilt: F_ε ist stetig, $F_\varepsilon(0) = 0$, und

$$F_\varepsilon(1) \leq 0 \Leftrightarrow \|f(x) - f(y)\|_Y \leq C\|x - y\|_X + \varepsilon\|x - y\|_X.$$

Es genügt also zu zeigen:

$$(2.12) \quad F_\varepsilon(1) \leq 0 \quad \text{für alle } \varepsilon > 0,$$

Dann folgt die Behauptung durch Grenzübergang $\varepsilon \rightarrow 0$.

2. Wir beweisen (2.12). Annahme: Die Aussage gilt nicht. Dann existiert $\varepsilon > 0$ mit $F_\varepsilon(1) > 0$. Wähle c mit $F_\varepsilon(0) = 0 < c < F_\varepsilon(1)$ und setze

$$t_0 := \max\{t \in [0, 1] : F_\varepsilon(t) \leq c\}.$$

Nach Konstruktion gilt $0 < t_0 < 1$ sowie

$$(\star) \quad \varphi(t) := \frac{F_\varepsilon(t) - F_\varepsilon(t_0)}{t - t_0} > 0 \quad \text{für alle } t > t_0.$$

Mit der inversen Dreiecksungleichung folgt

$$\varphi(t) \leq \left\| \frac{f(u(t)) - f(u(t_0))}{t - t_0} \right\|_Y - (C + \varepsilon)\|x - y\|_X.$$

Mit der Weierstraßschen Zerlegungsformel gilt für $t > t_0$:

$$\begin{aligned} \frac{f(u(t)) - f(u(t_0))}{t - t_0} &= \frac{f(u(t_0) + (t - t_0)(y - x)) - f(u(t_0))}{t - t_0} \\ &= Df(u(t_0))(y - x) + \|x - y\|_X \rho\left((t - t_0)(y - x)\right) \end{aligned}$$

mit $\rho(h) \rightarrow 0$ für $h \rightarrow 0$. Also folgt

$$\begin{aligned} \lim_{\substack{t \rightarrow t_0 \\ t > t_0}} \left\| \frac{f(u(t)) - f(u(t_0))}{t - t_0} \right\|_Y &= \|Df(u(t_0))(y - x)\|_Y \\ &\leq \|Df(u(t_0))\| \|y - x\|_X \\ &\leq C\|y - x\|_X, \end{aligned}$$

und damit $\limsup_{t \rightarrow t_0, t > t_0} \varphi(t) \leq -\varepsilon\|y - x\|_X$. Also existiert $t_1 \in (t_0, 1]$ mit $\varphi(t_1) \leq 0$ im Widerspruch zu $(\star)$. □

Kontrollfrage 7.1. Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine differenzierbare Funktion mit $f'(x) = 2x$ für alle $x \in \mathbb{R}$. Welche der nachfolgenden Aussagen ist richtig?

(A) f ist total differenzierbar.

✓: Siehe Bemerkung 2.40, Punkt 3.

(B) Die totale Ableitung von f im Punkt $x = 1$ ist $Df(x) = 2$

×: Die totale Ableitung ist eine lineare Abbildung, keine reelle Zahl.

(C) Die totale Ableitung von f im Punkt $x = 1$ ist $Df(x) : \mathbb{R} \rightarrow \mathbb{R}, y \mapsto 2y$.

✓: Siehe auch Bemerkung 2.40, Punkt 3.

Kontrollfrage 7.2. Sei $(X, \|\cdot\|)$ ein normierter Raum und $L : X \rightarrow \mathbb{R}$ eine lineare Abbildung. Welche der nachfolgenden Aussagen ist richtig?

1. L ist differenzierbar.

×: Differenzierbare Abbildungen müssen stetig sein. (Siehe Proposition 2.41.) Aber im Unendlichdimensionalen gibt es unstetige lineare Abbildungen, siehe Bemerkung 1.26.

2. L ist differenzierbar, falls X endlichdimensional ist.

✓: Siehe Beispiel 2.42 b) und Bemerkung 1.26.

Kontrollfrage 7.3. Sei $F \in \mathbb{R}^{n \times n}$ und betrachte

$$f : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}, \quad f(A) := AAF.$$

Welche der nachfolgenden Aussagen ist richtig?

1. f ist differenzierbar und es gilt

$$Df(A)(H) = 2AFH.$$

×

2. f ist differenzierbar und es gilt

$$Df(A)(H) = HAF + AHF.$$

✓: Betrachte $f(A + H) - f(A)$ und identifiziere den in H linearen Teil und den Restterm!

3. f ist differenzierbar und es gilt

$$Df(A)(H) = 2HFA.$$

×

2.3 Partielle Ableitung

Ziele:

- Partielle Ableitung für Funktionen $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$.
- Kriterium für die (totale) Differenzierbarkeit
- Jacobi-Matrix
- Lemma von Schwarz

Wir betrachten nun den Spezialfall $X = \mathbb{R}^n$ und $Y = \mathbb{R}^m$. Im ganzen Kapitel bezeichnet $\Omega \subseteq \mathbb{R}^n$ eine offene Menge. Sei $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$. Dann ist $Df(x)$ eine lineare Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$ und kann durch eine $m \times n$ -Matrix repräsentiert werden (siehe Bemerkung 1.2). Wir werden diese Matrix als die Jacobi-Matrix von f an der Stelle x bezeichnen und werden sehen, dass ihre Einträge durch Ableiten reeller Funktionen einer Veränderlichen berechnet werden können. Dies führt auf den Begriff der partiellen Ableitung und einem praktischen Kriterium um die Differenzierbarkeit von f zu überprüfen.

Um Abbildungen der Form

$$f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$$

zu verstehen, möchten wir f auf skalare Funktionen einer reellen Variablen zurückführen. Hierbei müssen wir sowohl den Bildraum $\mathbb{R}^m$ als auch den Raum des Definitionsbereiches $\mathbb{R}^n$ betrachten.

Wir kümmern uns zunächst um den Bildraum. Wie bei Kurven (Vgl. Lemma 2.7) definieren wir für $v \in \mathbb{R}^m$ die skalare Funktion

$$f_v : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}, \quad f_v(x) := v \cdot f(x).$$

Die skalaren Funktionen

$$f_i : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}, \quad f_i(x) := e_i \cdot f(x)$$

mit der Standardbasis $e_1, \dots, e_m$ von $\mathbb{R}^m$ heißen Komponentenfunktionen von f und es gilt

$$f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix} \quad \text{für alle } x \in \Omega.$$

15

Lemma 2.47. Sei $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$, $v \in \mathbb{R}^m$. Dann sind äquivalent:

- (a) f ist differenzierbar in $x \in \Omega$ (bzw. stetig differenzierbar).
- (b) Für alle $i = 1, \dots, m$ ist f_i differenzierbar in $x \in \Omega$ (bzw. stetig differenzierbar).

¹⁵Da wir in diesem Kapitel Matrix-Vektor-Produkte vielfach verwendet werden, empfiehlt es sich die Elemente von $a \in \mathbb{R}^m$ als Spaltenvektoren $\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}$ vorzustellen. Manchmal benutzen wir um Platz zu sparen die Schreibweise $(a_1 \ \cdots \ a_m)^T$.

(c) Für alle $v \in \mathbb{R}^m$ ist $f_v : \Omega \rightarrow \mathbb{R}$ differenzierbar in $x \in \Omega$ (bzw. stetig differenzierbar).

Falls f in $x \in \Omega$ differenzierbar ist, dann gilt

$$Df(x)(h) = \begin{pmatrix} Df_1(x)(h) \\ \vdots \\ Df_m(x)(h) \end{pmatrix} \quad \text{für alle } h \in \mathbb{R}^n,$$

sowie für $v \in \mathbb{R}^m$,

$$Df_v(x)(h) = v \cdot Df(x)(h) \quad \text{für alle } h \in \mathbb{R}^n,$$

Beweis. Das folgt aus der Charakterisierung der Konvergenz in $\mathbb{R}^m$ mittels Konvergenz der Komponenten, siehe Lemma 1.22. $\square$

Als nächstes beschäftigen wir uns mit dem Definitionsbereich. Seien $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine Abbildung, $x \in \Omega$ und $e \in \mathbb{R}^n$ eine Richtung. Für $r > 0$ klein genug¹⁶ wird durch

$$(-r, r) \rightarrow \mathbb{R}^m, \quad t \mapsto f(x + te)$$

eine Kurve im $\mathbb{R}^m$ definiert. Diese beschreibt Veränderungen von f “längs der Geraden” $t \mapsto x + te$ im $\mathbb{R}^n$. Im Spezialfall $e = e_i$ sind diese Kurven parallel zu den Koordinatenachsen. Man kann nun die Differenzierbarkeit der Kurven $t \mapsto f(x + te)$ untersuchen. Dies führt auf die Begriffe der “partiellen Ableitung” und der “Richtungsableitung”:

Definition 2.48 (Richtungsableitung, partielle Ableitung).

Sei $\Omega \subseteq \mathbb{R}^n$ offen, $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine Funktion, $x \in \Omega$.

(a) f heißt im Punkt x in Richtung $e \in \mathbb{R}^n$ differenzierbar, falls die Funktion

$$(-r, r) \rightarrow \mathbb{R}^m, \quad t \mapsto f(x + te)$$

in $t = 0$ differenzierbar ist. (Hierbei ist $r > 0$ so klein gewählt, dass $B(x; r\|e\|) \subset \Omega$). In diesem Fall wird der Tangentialvektor bei $t = 0$, d.h.

$$\frac{\partial f}{\partial e}(x) := \lim_{t \rightarrow 0} \frac{f(x + te) - f(x)}{t}$$

Richtungsableitung von f im Punkte x in Richtung e genannt.

(b) f heißt im Punkt $x \in \Omega$ nach x_j (mit $j \in \{1, \dots, n\}$) partiell differenzierbar, falls f in x in Richtung $e_j \in \mathbb{R}^n$ differenzierbar ist. In diesem Fall nennt man

$$\frac{\partial}{\partial x_j} f(x) := \frac{\partial f}{\partial e_j}(x) = \lim_{t \rightarrow 0} \frac{f(x_1, \dots, x_j + t, \dots, x_n) - f(x)}{t}$$

partielle Ableitung von f nach x_j im Punkt x .

(c) f heißt im Punkt x partiell differenzierbar, falls f in x nach $x_1, \dots, x_n$ partiell differenzierbar ist.

¹⁶Da Ω offen ist, existiert $r > 0$, sodass $x + te \in \Omega$ für alle $t \in [-r, r]$

(d) f heißt partiell nach x_j differenzierbar, wenn sie in jedem Punkt partiell nach x_j differenzierbar ist. In diesem Fall heißt die Abbildung

$$\frac{\partial}{\partial x_j} f : \Omega \rightarrow \mathbb{R}^m, \quad x \mapsto \frac{\partial}{\partial x_j} f(x)$$

partielle Ableitung von f nach x_j .

(e) f heißt partiell differenzierbar, falls f partiell nach x_j für alle $j = 1, \dots, n$ differenzierbar ist.

(f) f heißt stetig partiell differenzierbar, falls f partiell differenzierbar ist und alle partiellen Ableitungen $\frac{\partial}{\partial x_j} f : \Omega \rightarrow \mathbb{R}^m$ stetig sind.

Beispiele 2.49. (a) $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $f(x) := Ax$ mit $A \in \mathbb{R}^{m \times n}$, $e \in \mathbb{R}^n$,

$$\frac{\partial f}{\partial e}(x) = \lim_{t \rightarrow 0} \frac{f(x + te) - f(x)}{t} = Ae, \quad \frac{\partial}{\partial x_j} f(x) = Ae_j = \begin{pmatrix} A_{1j} \\ \vdots \\ A_{mj} \end{pmatrix} \in \mathbb{R}^m$$

(b) $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, $f(x) = \begin{pmatrix} x_1^2 x_2 x_3 + \sin(x_1) \\ x_2^3 \cos(x_3) \end{pmatrix}$.

$$\begin{aligned} \frac{\partial}{\partial x_1} f_1(x) &= 2x_1 x_2 x_3 + \cos(x_1) \\ \frac{\partial}{\partial x_1} f_2(x) &= 0 \\ \frac{\partial}{\partial x_1} f(x) &= \begin{pmatrix} 2x_1 x_2 x_3 + \cos(x_1) \\ 0 \end{pmatrix} \\ \frac{\partial}{\partial x_2} f(x) &= \begin{pmatrix} x_1^2 x_3 \\ 3x_2^2 \cos(x_3) \end{pmatrix} \\ \frac{\partial}{\partial x_3} f(x) &= \begin{pmatrix} x_1^2 x_2 \\ -x_2^3 \sin(x_3) \end{pmatrix} \end{aligned}$$

(c) $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f(x) := \|x\|_2$.

$$\frac{\partial}{\partial x_j} f(x) = \frac{1}{2}(x \cdot x)^{-\frac{1}{2}} 2e_j \cdot x = \frac{x_j}{\|x\|_2}.$$

(d) Die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x) = |x_1| x_2$ ist nach x_2 partiell differenzierbar mit

$$\frac{\partial}{\partial x_2} f(x) = |x_1|.$$

f ist in $x = (0, x_2)$ mit $x_2 \neq 0$ nicht partiell nach x_1 differenzierbar, denn

$$\lim_{\substack{t \rightarrow 0 \\ t > 0}} \frac{|t| x_2}{t} = x_2 \neq \lim_{\substack{t \rightarrow 0 \\ t < 0}} \frac{|t| x_2}{t} = -x_2.$$

Lemma 2.50 (Summen und Produktregel für partielle Ableitungen). *Seien $f, g : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ und $\varphi : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ partiell nach x_j differenzierbar im Punkt $x \in \Omega$. Dann gelten:*

(a) $f + g$ ist partiell nach x_j differenzierbar in x und es gilt

$$\frac{\partial}{\partial x_j}(f + g)(x) = \frac{\partial}{\partial x_j}f(x) + \frac{\partial}{\partial x_j}g(x)$$

(b) φf ist partiell nach x_j differenzierbar in x und es gilt

$$\frac{\partial}{\partial x_j}(\varphi f)(x) = \frac{\partial}{\partial x_j}\varphi(x) f(x) + \varphi(x) \frac{\partial}{\partial x_j}f(x).$$

Beweis. Das folgt durch Anwendung der Rechenregeln für die Ableitung aus Analysis 1 auf die reellen Funktionen einer Veränderlichen

$$u(t) := (f_i + g_i)(x + te_j) \quad \text{und} \quad v(t) := \varphi(x + te_j)f_i(x + te_j), \quad (i = 1, \dots, m).$$

□

Definition 2.51 (Jacobi-Matrix). *Sei $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ in $x \in \Omega$ partiell differenzierbar. Die Matrix*

$$J_f(x) := \begin{pmatrix} \frac{\partial}{\partial x_1}f_1(x) & \cdots & \frac{\partial}{\partial x_n}f_1(x) \\ \vdots & & \vdots \\ \frac{\partial}{\partial x_1}f_m(x) & \cdots & \frac{\partial}{\partial x_n}f_m(x) \end{pmatrix} \in \mathbb{R}^{m \times n}$$

heißt Jacobi-Matrix (oder auch Funktionalmatrix) von f in x .

Satz 2.52. *Sei $\Omega \subset \mathbb{R}^n$ offen, $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine Abbildung.*

(a) *Ist f im Punkt $x \in \Omega$ differenzierbar, dann ist f im Punkt x partiell differenzierbar und es gilt*

$$Df(x)(h) = J_f(x)h \quad \text{für alle } h \in \mathbb{R}^n.$$

(b) *Ist f in einer Umgebung von x partiell differenzierbar und sind alle partiellen Ableitungen stetig in x , dann ist f differenzierbar in x .*

Beweis.

(a) Betrachte für $j = 1, \dots, n$, $u(t) := f(x + te_j)$. Nach Definition ist f genau dann partiell nach x_j differenzierbar in x , wenn u differenzierbar bei $t = 0$ ist. Nach der Kettenregel ist u differenzierbar im Punkt $t = 0$ und

$$\frac{\partial}{\partial x_j}f(x) = u'(0) = Df(x)(e_j).$$

Da dies für alle $j = 1, \dots, n$ gilt, folgt $J_f(x)h = Df(x)(h)$.

- (b) Da f genau dann differenzierbar im Punkt x ist, wenn alle Komponenten f_i differenzierbar in x sind (siehe Lemma 2.47), genügt es den Fall $m = 1$ zu betrachten. Setze $J_f(x) = \left(\frac{\partial}{\partial x_1} f(x), \dots, \frac{\partial}{\partial x_n} f(x) \right) \in \mathbb{R}^{1 \times n}$. Wir müssen zeigen, dass die Restgliedabschätzung der Weierstraßschen Zerlegungsformel mit $T(h) := J_f(x)h$ erfüllt ist. Wähle $\delta > 0$, so dass f in $\{x + h : \|h\|_\infty < \delta\}$ partiell differenzierbar ist. Sei nun $h \in \mathbb{R}^n \setminus \{0\}$ mit $\|h\|_\infty < \delta$. Wir definiere iterativ Vektoren $h^{(0)}, \dots, h^{(n)} \in \mathbb{R}^n$ durch

$$h^{(0)} := 0, \quad h^{(k)} := h^{(k-1)} + h_k e_k.$$

Es gilt $h = h^{(n)}$ und

$$f(x + h) - f(x) = \sum_{k=1}^n f(x + h^{(k)}) - f(x + h^{(k-1)}) \quad (\text{Teleskopsumme}).$$

Mit dem (eindimensionalen) Mittelwertsatz angewendet auf $\frac{u(h_k) - u(0)}{h_k}$ mit $u(t) := f(x + h^{(k-1)} + te_k)$ folgt

$$f(x + h^{(k)}) - f(x + h^{(k-1)}) = h_k \frac{u(h_k) - u(0)}{h_k} = h_k u'_k(\tau) \quad \text{mit } \tau \in (0, h_k).$$

Mit $\xi_{k,h} := x + h^{(k-1)} + \tau e_k$ gilt $u'_k(\tau) = \frac{\partial}{\partial x_k} f(\xi_{k,h})$ (nach Definition der partiellen Ableitung), also

$$f(x + h^{(k)}) - f(x + h^{(k-1)}) = h_k \frac{\partial}{\partial x_k} f(\xi_{k,h}).$$

Es folgt

$$\begin{aligned} f(x + h) - f(x) &= \sum_{k=1}^n h_k \frac{\partial}{\partial x_k} f(\xi_{k,h}) \\ &= J_f(x)h + \sum_{k=1}^n h_k \left(\frac{\partial}{\partial x_k} f(\xi_{k,h}) - \frac{\partial}{\partial x_k} f(x) \right). \end{aligned}$$

Also

$$|f(x + h) - f(x) - J_f(x)h| \leq \|h\|_\infty \sum_{k=1}^n \left| \frac{\partial}{\partial x_k} f(\xi_{k,h}) - \frac{\partial}{\partial x_k} f(x) \right|.$$

Nun gilt $\xi_{k,h} \rightarrow x$ für $h \rightarrow 0$ und da die partiellen Ableitungen von f bei x stetig sind, folgt

$$\lim_{h \rightarrow 0} \sum_{k=1}^n \left| \frac{\partial}{\partial x_k} f(\xi_{k,h}) - \frac{\partial}{\partial x_k} f(x) \right| = 0.$$

Also ist f differenzierbar in x .

□

Korollar 2.53. $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ ist genau dann stetig differenzierbar, wenn f stetig partiell differenzierbar ist.

Bemerkung 2.54. Mit Satz 2.52 haben wir ein Kriterium um die Differenzierbarkeit einer Funktion f in einem Punkt x zu überprüfen: Zunächst stellen wir die Jacobi-Matrix $J_f(x)$ auf (falls dies nicht möglich ist, kann f nicht in x differenzierbar sein). Dann überprüfen wir, ob die lineare Abbildung $T(h) := J_f(x)h$ die Bedingung aus Definition 2.39 (a) erfüllt. Eine hinreichende (aber nicht notwendige) Bedingung hierfür ist die Stetigkeit der partiellen Ableitungen.

Beispiel 2.55 (Polarkoordinaten in der Ebene). Die Abbildung

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \quad f(r, \alpha) := (r \cos \alpha, r \sin \alpha)$$

repräsentiert die Transformation von Polarkoordinaten auf kartesische Koordinaten. Es gilt

$$J_f(r, \alpha) = \begin{pmatrix} \cos \alpha & -r \sin \alpha \\ \sin \alpha & r \cos \alpha \end{pmatrix}$$

Damit sind alle partiellen Ableitungen von f stetig. Aus Satz 2.52 folgt, dass f stetig differenzierbar ist.

	17.5.2020
#24	17.5.2020

Lemma 2.56 (Kettenregel für partielle Ableitungen). Seien $U \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^m$ offene Mengen,

$$f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad g : V \subseteq \mathbb{R}^m \rightarrow \mathbb{R}^d, \quad \text{Bild}(f) \subseteq V.$$

(a) Sei f im Punkt $x \in U$ partiell nach x_j differenzierbar, und g in $y := f(x)$ differenzierbar. Dann ist $g \circ f : U \rightarrow \mathbb{R}^d$ im Punkt x partiell nach x_j differenzierbar und es gilt

$$\frac{\partial}{\partial x_j}(g \circ f)(x) = J_g(f(x)) \frac{\partial}{\partial x_j} f(x).$$

(b) Sei f im Punkte $x \in U$ partiell differenzierbar, und g in $y := f(x)$ differenzierbar. Dann ist $g \circ f : U \rightarrow \mathbb{R}^d$ im Punkt x partiell differenzierbar und es gilt

$$J_{g \circ f}(x) = J_g(f(x)) J_f(x).$$

Beweis. (a) Setze

$$u(t) := g(f(x + te_j)).$$

Aufgrund der partiellen Differenzierbarkeit von f (in x) und der Differenzierbarkeit (in $f(x)$) von g gilt

$$f(x + te_j) = f(x) + t \underbrace{\frac{\partial}{\partial x_j} f(x)}_{=: h(t)} + |t| \rho(t) \quad \text{mit } \rho(t) \rightarrow 0 \text{ für } t \rightarrow 0.$$

$$u(t) - u(0) = g(f(x) + h(t)) - g(f(x)) = Dg(f(x))(h(t)) + \|\rho(t)\| \rho_g(h(t))$$

mit $\rho_g(h) \rightarrow 0$ für $h \rightarrow 0$. Nun gilt

$$\begin{aligned} & \frac{Dg(f(x))(h(t))}{t} \\ = & Dg(f(x))\left(\frac{\partial}{\partial x_j}f(x)\right) + Dg(f(x))(\rho(t))\frac{|t|}{t} \quad (\text{wegen Linearität von } Dg(f(x))) \\ \rightarrow & J_g(f(x))\frac{\partial}{\partial x_j}f(x) + 0 \quad \text{für } t \rightarrow 0 \quad (\text{wegen } \lim_{t \rightarrow 0} \rho(t) = 0). \end{aligned}$$

Also ist u differenzierbar in 0 und $\frac{\partial}{\partial x_j}(g \circ f)(x) = u'(0) = J_g(f(x))\frac{\partial}{\partial x_j}f(x)$.

(b) Das ist eine Übung. □

Bemerkung 2.57. Die Kettenregel kann man sich einfach merken, indem man den Spezialfall

$$f(x) := Ax, \quad g(y) := By, \quad (g \circ f)(x) = BAx$$

mit $A \in \mathbb{R}^{m \times n}$ und $B \in \mathbb{R}^{d \times m}$ betrachtet. Dann gilt offensichtlich

$$J_{g \circ f}(x) = BA, \quad J_g(y) = B, \quad J_f(x) = A.$$

Korollar 2.58. Sei $g : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ in $x \in \Omega$ partiell differenzierbar, $e \in \mathbb{R}^n$. Dann gilt

$$\frac{\partial}{\partial e}g(x) = J_g(x)e.$$

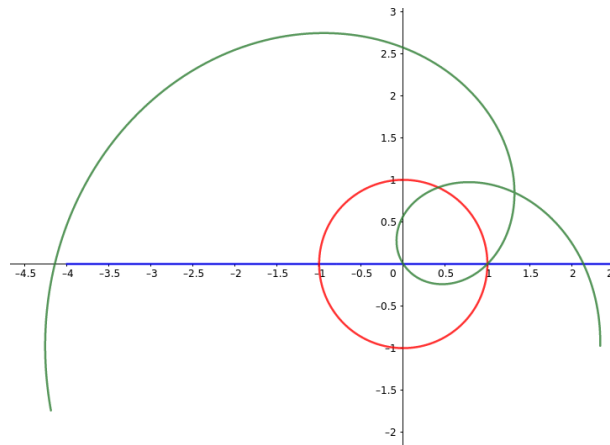
Beweis. Folgt aus Lemma 2.56 mit $f(t) := x + te$. □

Beispiel 2.59. Sei $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$ stetig differenzierbar und $u : [a, b] \rightarrow \mathbb{R}^d$ eine stetig differenzierbare Kurve. Dann ist $v := f \circ u$ eine stetig differenzierbare Kurve und es gilt

$$v'(t) = J_f(u(t))u'(t).$$

Beispiel: Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ die Transformation aus Beispiel 2.55, und seien $e \in \mathbb{R}^2$,

$$u : \mathbb{R} \rightarrow \mathbb{R}^2, \quad u(t) := (1, 0) + te, \quad v := f \circ u.$$



Die Abbildung zeigt die Spur von v mit $e = e_1$ (blau), $e = e_2$ (rot), $e = \frac{1}{\sqrt{2}}(1, 1)$ (grün).

Bemerkung 2.60. Aus Satz 2.52 und Proposition 2.41 folgt

$$\begin{aligned} f \text{ stetig partiell differenzierbar} &\Leftrightarrow f \text{ stetig differenzierbar,} \\ f \text{ differenzierbar} &\Rightarrow f \text{ partiell differenzierbar,} \\ f \text{ differenzierbar} &\Rightarrow f \text{ stetig.} \end{aligned}$$

Alle anderen möglichen Implikationen (außer jenen die sich durch Transitivität ergeben) zwischen diesen 4 Begriffen gelten nicht!

Beispiel 2.61. Die Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = \begin{cases} (0, 0) & x = y = 0 \\ \frac{x^2 y}{x^2 + y^2} & \text{sonst} \end{cases}$$

ist nicht differenzierbar in $(0, 0)$, aber partiell differenzierbar. Siehe Übung.

Partielles Ableiten kann iteriert werden.

Beispiele:

$$\begin{aligned} f : \mathbb{R}^d \rightarrow \mathbb{R}^2, \quad f(x) &:= \begin{pmatrix} x_1^2 x_2 \\ x_1 x_2 \end{pmatrix}, \\ u : \mathbb{R}^d \rightarrow \mathbb{R}^2, \quad u(x) &:= \frac{\partial}{\partial x_1} f(x) = \begin{pmatrix} 2x_1 x_2 \\ x_2 \end{pmatrix}, \\ v : \mathbb{R}^d \rightarrow \mathbb{R}^2, \quad v(x) &:= \frac{\partial}{\partial x_2} \frac{\partial}{\partial x_1} f(x) := \frac{\partial}{\partial x_2} u(x) = \begin{pmatrix} 2x_1 \\ 1 \end{pmatrix} \end{aligned}$$

Definition 2.62 (Höhere partielle Ableitung). Sei $\Omega \subseteq \mathbb{R}^n$ offen. $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ heißt $(k+1)$ -mal partiell differenzierbar in Ω , falls f k -mal partiell differenzierbar in Ω ist und alle k -ten partiellen Ableitungen der Form

$$\frac{\partial}{\partial x_{j_k}} \cdots \frac{\partial}{\partial x_{j_1}} f, \quad j_\ell \in \{1, \dots, n\}, \quad \ell \in \{1, \dots, k\},$$

partiell differenzierbar sind.

f heißt k -mal stetig differenzierbar, falls alle partiellen Ableitung der Ordnung $\leq k$ existieren und stetig sind.

Wir definieren

$$\begin{aligned} C^0(\Omega; \mathbb{R}^m) &:= C(\Omega; \mathbb{R}^m) \\ C^k(\Omega; \mathbb{R}^m) &:= \{f : \text{Die Funktion } f : \Omega \rightarrow \mathbb{R}^m \text{ ist } k\text{-mal stetig partiell differenzierbar.}\} \\ C^\infty(\Omega; \mathbb{R}^m) &:= \bigcap_{k=0}^{\infty} C^k(\Omega; \mathbb{R}^m) \end{aligned}$$

Im Fall $m = 1$ schreibt man einfach $C^k(\Omega)$ statt $C^k(\Omega; \mathbb{R}^m)$ für $k = 0, \dots, \infty$.

Satz 2.63 (Lemma von Schwarz). Sei $\Omega \subseteq \mathbb{R}^n$, $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine Funktion, $k, j \in \{1, \dots, n\}$. Falls die partielle Ableitungen

$$\frac{\partial}{\partial x_k} \frac{\partial}{\partial x_j} f, \quad \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_k} f$$

existieren und im Punkt $x \in \Omega$ stetig sind, dann gilt

$$\frac{\partial}{\partial x_k} \frac{\partial}{\partial x_j} f(x) = \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_k} f(x).$$

Beweis.

Es genügt denn Fall $n = 2$ und $x = 0$ zu betrachten.¹⁷ Wähle $\delta > 0$ mit $B(0; \delta) \subseteq \Omega$ (Kugel bzgl. Maximumsnorm). Sei $x = (x_1, x_2) \in B(0; \delta)$ mit $x_1 \neq 0$ und $x_2 \neq 0$.

1. Behauptung: Es existieren $\xi \in (0, x_1)$ und $\eta \in (0, x_2)$ mit

$$(2.13) \quad f(x_1, x_2) - f(x_1, 0) = f(0, x_2) - f(0, 0) + x_1 x_2 \frac{\partial}{\partial x_2} \frac{\partial}{\partial x_1} f(\xi, \eta).$$

Argument: Mit dem Mittelwertsatz angewendet auf

$$u(t) := f(t, x_2) - f(t, 0),$$

folgt

$$u(x_1) = u(0) + x_1 u'(\xi) \quad \text{für ein } \xi \in (0, x_1),$$

also

$$(2.14) \quad f(x_1, x_2) - f(x_1, 0) = f(0, x_2) - f(0, 0) + x_1 \left(\frac{\partial}{\partial x_1} f(\xi, x_2) - \frac{\partial}{\partial x_1} f(\xi, 0) \right).$$

Mit dem Mittelwertsatz angewendet auf $v(s) := \frac{\partial}{\partial x_1} f(\xi, s)$ folgt

$$v(x_2) = v(0) + x_2 v'(\eta) \quad \text{für ein } \eta \in (0, x_2),$$

und in Kombination mit (2.14) die Behauptung.

2. Wir wiederholen Schritt 1 mit vertauschten Rollen von x_1 und x_2 und erhalten

$$f(x_1, x_2) - f(0, x_2) = f(x_1, 0) - f(0, 0) + x_1 x_2 \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} f(\tilde{\xi}, \tilde{\eta}),$$

für ein $\tilde{\xi} \in (0, x_1)$ und ein $\tilde{\eta} \in (0, x_2)$. Zusammen mit (2.13) folgt

$$\frac{\partial}{\partial x_2} \frac{\partial}{\partial x_1} f(\xi, \eta) = \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} f(\tilde{\xi}, \tilde{\eta}),$$

Für $\delta \rightarrow 0$ gilt $\|x\|_\infty \rightarrow 0$ und damit $(\xi, \eta) \rightarrow 0$ sowie $(\tilde{\xi}, \tilde{\eta}) \rightarrow 0$. Aus der Stetigkeit der zweiten partiellen Ableitungen folgt nun

$$\frac{\partial}{\partial x_2} \frac{\partial}{\partial x_1} f(0, 0) = \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} f(0, 0).$$

¹⁷andernfalls betrachte $u(s, t) := f(x + se_k + te_j)$.

□

Durch Iteration erhalten wir die Aussagen, dass die Reihenfolge beim k -maligen partiellen Ableiten einer C^k -Funktion keine Rolle spielt:

Korollar 2.64. Sei $f \in C^k(\Omega; \mathbb{R}^m)$. Seien $j_1, \dots, j_k \in \{1, \dots, n\}$ und $\pi : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ bijektiv. Dann gilt

$$\frac{\partial}{\partial x_{j_1}} \cdots \frac{\partial}{\partial x_{j_k}} f = \frac{\partial}{\partial x_{j_{\pi(1)}}} \cdots \frac{\partial}{\partial x_{j_{\pi(k)}}} f$$

Kontrollfrage 7.4. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ im Punkt $x \in \mathbb{R}^2$ differenzierbar mit

$$Df(x)(h) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} h, \quad \text{für alle } h \in \mathbb{R}^2,$$

dann ist f im Punkt x partiell differenzierbar und es gilt

(A) $\frac{\partial}{\partial x_1} f_2(x) = b.$

$\times: \frac{\partial}{\partial x_1} f_2(x) = c$

(B) $\frac{\partial}{\partial x_1} f_1(x) = a.$

✓: Siehe Satz 2.52.

(C) $\frac{\partial}{\partial x_2} f(x) = (b, d)^\top.$

✓ Siehe Satz 2.52.

(D) $\frac{\partial}{\partial x_1} f(x) = Df(x)(e_1).$

✓ Siehe Satz 2.52.

Kontrollfrage 7.5. Welche der nachfolgenden Aussagen sind im Allgemeinen für eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ richtig?

(A) f ist nicht stetig $\Rightarrow f$ ist nicht differenzierbar.

✓: Siehe Proposition 2.41.

(B) f ist partiell differenzierbar $\Rightarrow f$ ist differenzierbar.

$\times$: Siehe Bemerkung 2.60.

(C) f ist differenzierbar $\Rightarrow f$ ist partiell differenzierbar.

✓: Siehe Bemerkung 2.60.

(D) f ist partiell differenzierbar $\Rightarrow f$ ist stetig.

$\times$: Siehe Beispiel 2.61.

(E) f ist stetig partiell differenzierbar $\Rightarrow f$ ist stetig.

✓: Siehe Bemerkung 2.60.

Kontrollfrage 7.6. Für alle differenzierbaren Funktionen $f, g \in C^1(\mathbb{R}^2)$ ist die Funktion $\mathbb{R}^2 \rightarrow \mathbb{R}^2$, $x \mapsto (f(x), g(x))$

(A) differenzierbar

✓: Siehe Lemma 2.47.

(B) stetig differenzierbar

✓: Siehe Lemma 2.47.

(C) nicht unbedingt differenzierbar

×: Siehe (A).

Kontrollfrage 7.7. Wenn für eine Funktion $f \in \text{Abb}(\mathbb{R}^2; \mathbb{R}^2)$ im Punkt $x \in \mathbb{R}^2$ die Aussage¹⁸

$$\forall y \in \mathbb{R}^2 : f(y) = f(x) + \begin{pmatrix} 3 & 4 \\ 2 & 0 \end{pmatrix} (y - x) + \begin{pmatrix} \|y - x\|_2^3 \\ 0 \end{pmatrix}$$

gilt, dann ist im Allgemeinen

(A) f linear,

×: Kubischer Term!

(B) f partiell im Punkt x differenzierbar,

✓: Folgt aus (C) mit Bemerkung 2.60.

(C) f differenzierbar im Punkt x ,

✓: Begründbar (D) oder direkt mit der Weierstraßschen Zerlegungsformel, wobei der kubische Restterm mit „ $\|y - x\|_2^3$ “ die Bedingung (2.11) erfüllt.

(D) f differenzierbar.

✓: Verknüpfung differenzierbarer Funktionen (Siehe Lemma 2.50.)

Kontrollfrage 7.8. Welche der nachfolgenden Aussagen sind richtig?

(A) Die partielle Ableitung nach x_2 von $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x) = |x_1|x_2 + x_2^2$, ist nicht differenzierbar.

✓: $\frac{\partial}{\partial x_2} f(x) = |x_1| + 2x_2$ ist nicht einmal partiell nach x_1 differenzierbar.

(B) Für $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x) = x_1(1 + x_2^2)^\pi$ gilt $\frac{\partial}{\partial x_1} \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} f(x) = 0$

✓: Am schnellsten mit Satz 2.63 sichtbar.

(C) Für $f : \mathbb{R}^d \rightarrow \mathbb{R}$, $f(x) = \|x\|_2^2$ und $v \in \mathbb{R}^d$ gilt $\frac{\partial}{\partial v} f(x) = 2v \cdot x$.

✓: Beispiel 2.42 c) + Korollar 2.58.

2.4 Skalarfeld, Niveaumenge, Gradient

Definition 2.65 (Skalarfeld und Gradient). *Sei $\Omega \subseteq \mathbb{R}^n$ offen.*

(a) *Ein $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ heißt skalarfeld, oder Skalarfeld, auf Ω .*

(b) *Für $c \in \mathbb{R}$ wird*

$$N_c(f) := \{x \in \Omega : f(x) = c\}$$

Niveaumenge von f (zum Niveau c) genannt.

(c) *Sei $\Omega \subseteq \mathbb{R}^n$ offen, $f : \Omega \rightarrow \mathbb{R}$ partiell differenzierbar. Dann heißt*

$$\nabla f(x) := \left(\frac{\partial}{\partial x_1} f(x), \dots, \frac{\partial}{\partial x_n} f(x) \right) \in \mathbb{R}^n$$

der Gradient von f im Punkt x .

Für ein Skalarfeld f und $x \in \Omega$ ist $Df(x)$ eine lineare Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}$. Für alle

$$h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \in \mathbb{R}^n$$

gilt (siehe Satz 2.52)

$$\begin{aligned} Df(x)(h) &= \sum_{j=1}^n h_j Df(x)(e_j) = J_f(x)h = \sum_{j=1}^n h_j \frac{\partial}{\partial x_j} f(x) \\ &= \nabla f(x) \cdot h \end{aligned}$$

Beispiel 2.66. *Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch. Für $f(x) := x \cdot Ax$ gilt $Df(x)(h) = 2Ax \cdot h$, also $\nabla f(x) = 2Ax$.*

Korollar 2.67 (Richtungsableitung und Gradient). *Sei $\Omega \subseteq \mathbb{R}^n$ offen, $f : \Omega \rightarrow \mathbb{R}$ und $e \in \mathbb{R}^n$. Sei f in $x \in \Omega$ differenzierbar, dann gilt*

$$\frac{\partial}{\partial e} f(x) = \nabla f(x) \cdot e.$$

Bemerkung 2.68. *Die Aussage gilt nicht, wenn f in x nur partiell differenzierbar ist (siehe Beispiel 2.61).*

Für $f, g : \Omega \rightarrow \mathbb{R}$ und $\lambda \in \mathbb{R}$ gelten die Rechenregeln

$$\nabla(f + \lambda g) = \nabla f + \lambda \nabla g, \quad \nabla(f \cdot g) = g \cdot \nabla f + f \cdot \nabla g.$$

(Das folgt direkt aus den Rechenregeln für die totale Ableitung).

Sei nun $g : \mathbb{R}^m \rightarrow \mathbb{R}$ differenzierbar und $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $F(x) = Ax$. Setze

$$u := g \circ F.$$

Nach der Kettenregel ist $u : \mathbb{R}^n \rightarrow \mathbb{R}$ differenzierbar und es gilt

$$u(x) = (g \circ F)(x) = f(Ax), \quad Du(x)(h) = D(g \circ F)(x)(h) = Dg(Ax)(F(h)) = J_g(Ax)Ah$$

Aus

$$\nabla u(x) \cdot h = Du(x)(h) = J_g(Ax)Ah = \nabla g(Ax) \cdot Ah = A^\top \nabla g(Ax) \cdot h,$$

folgt $\nabla(g \circ F)(x) = J_F^\top(x) \nabla g(F(x))$.

Lemma 2.69. Seien $g : \mathbb{R}^m \rightarrow \mathbb{R}$ und $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ differenzierbar. Dann ist $g \circ F$ differenzierbar und es gilt

$$\nabla(g \circ F)(x) = J_F^\top(x) \nabla g(F(x)).$$

Beweis. Übung. □

Beispiel 2.70 (Topografische Karte). Wir beschreiben eine Gebirgslandschaft in einem Wandera-real. Wir repräsentieren das Areal durch ein Gebiet $\Omega \subseteq \mathbb{R}^2$. Wir beschreiben die Höhe eines Ortspunktes über Nullniveau mit einem Skalarfeld $u : \Omega \rightarrow \mathbb{R}$.

In topografischen Karten sind die Niveaulinien $N_c(u)$ für verschiedene Niveaus als “Höhenlinien” eingezeichnet.

Wir nehmen an dass u differenzierbar ist. Dann gilt für alle Richtungen $e \in S^1 := \{y \in \mathbb{R}^2, \|y\|_2 = 1\}$,

$$\frac{\partial}{\partial e} f(x) = \lim_{h \rightarrow 0} \frac{u(x + he) - u(x)}{h} = \nabla u(x) \cdot e.$$

Die rechte Seite gibt an, wie sich die Höhe ändert, wenn wir ein (infinitesimal) kleines Stück von x in Richtung e wandern.

Wenn $\nabla u(x) = 0$, dann ändert sich die Höhe nicht.

Falls $\nabla u(x) \neq 0$, dann wird der Ausdruck

$$S^1 \rightarrow \mathbb{R}, \quad e \mapsto \nabla u(x) \cdot e$$

durch $e := -\frac{\nabla u(x)}{\|\nabla u(x)\|_2}$ minimiert. Der normierte Vektor

$$-\frac{\nabla u(x)}{\|\nabla u(x)\|_2}$$

zeigt in die Richtung des steilsten Abstiegs und

$$\frac{\nabla u(x)}{\|\nabla u(x)\|_2}$$

in die Richtung des steilsten Anstiegs.

Sei nun $f : [a, b] \rightarrow \Omega$ eine differenzierbare Kurve, die z.B. unsere Wanderoute beschreibt. Die Höhenänderung entlang des Weges ist durch

$$(u \circ f)'(t) = \nabla u(f(t)) \cdot f'(t)$$

beschrieben. Es kann nun gelten, dass

$$(2.15) \quad (u \circ f)'(t) = 0 \quad \text{für alle } t \in (a, b),$$

d.h. in jedem Kurvenpunkt sind der Tangentialvektor $f'(t)$ und der Gradient $\nabla u(f(t))$ orthogonal. Aus (2.15) folgt, dass $u \circ f$ konstant ist, d.h. die Höhe ändert sich nicht entlang des Weges. Die Spur von f liegt in diesem Fall in einer Niveaumenge von u .

Definition 2.71 (Tangentialvektor an eine Niveaumenge). Sei $\Omega \subseteq \mathbb{R}^n$ offen, $f : \Omega \rightarrow \mathbb{R}$ differenzierbar, $x \in \Omega$ mit $\nabla f(x) \neq 0$. Setze $c := f(x)$. Ein Vektor $v \in \mathbb{R}^n$ heißt Tangentialvektor an $N_c(f)$ im Punkt x , falls $\nabla f(x) \cdot v = 0$.

Lemma 2.72. In der Situation von Definition 2.71 gilt

$$T(x) := \{v \in \mathbb{R}^n : v \text{ ist Tangentialvektor an } N_c(f) \text{ im Punkt } x\}$$

ist ein $n - 1$ -dimensionaler Unterraum des $\mathbb{R}^n$. $T(x)$ heißt Tangentialraum und $x + T(x)$ Tangentialebene an $N_c(f)$.

Beweis. Siehe lineare Algebra. □

Lemma 2.73. In der Situation von Definition 2.71 gilt: Ist $g : [a, b] \rightarrow \mathbb{R}^n$ eine differenzierbare Kurve mit $\text{Bild}(g) \subseteq N_c(f)$, dann gilt $g'(t) \in T(g(t))$ für alle $t \in [a, b]$.

Beweis. Es gilt $u(t) := f(g(t)) = c$ für alle $t \in [a, b]$, also $u'(t) = \nabla f(g(t)) \cdot g'(t) = 0$. □

Satz 2.74 (Mittelwertsatz für Skalarfelder). Sei $\Omega \subset \mathbb{R}^n$ offen, $f : \Omega \rightarrow \mathbb{R}$ differenzierbar. Seien $x, y \in \Omega$ mit $[[x, y]] \subseteq \Omega$. Dann existiert $\xi \in [[x, y]]$ mit

$$f(y) - f(x) = \nabla f(\xi) \cdot (y - x).$$

Beweis. Betrachte $u : [0, 1] \rightarrow \mathbb{R}$, $u(t) := f(x + t(y - x))$. Dann ist u differenzierbar (nach der Kettenregel). Es gilt $u(1) = f(y)$ und $u(0) = f(x)$. Die Behauptung folgt nun aus dem Mittelwertsatz und der Identität

$$u'(t) = Df(x + t(y - x))(y - x) = \nabla f(x + t(y - x)) \cdot (y - x).$$

□

Korollar 2.75. Sei $\Omega \subseteq \mathbb{R}^n$ offen und konvex, $f : \Omega \rightarrow \mathbb{R}$ differenzierbar mit $\nabla f(x) = 0$ für alle $x \in \Omega$. Dann ist f konstant.

Beweis. Sei f nicht konstant. Dann existieren $x, y \in \Omega$ mit $f(x) \neq f(y)$. Wegen der Konvexität gilt $[[x, y]] \subseteq \Omega$. Aus dem Mittelwertsatz für Skalarfelder folgt

$$\nabla f(\xi) \cdot (y - x) = f(y) - f(x) \neq 0$$

für ein $\xi \in \Omega$. Ein Widerspruch zu $\nabla f(z) = 0$ für alle $z \in \Omega$. □

Bemerkung 2.76. Das Konzept des Gradienten kann auch auf differenzierbare Funktionen $f : H \rightarrow \mathbb{R}$ mit einem Hilbertraum H ausgedehnt werden: Für $x \in H$ ist $Df(x) \in \mathcal{L}(H; \mathbb{R})$. Die Elemente in $\mathcal{L}(H; \mathbb{R})$ nennt man Funktionale. In der Funktionalanalysis zeigt man, dass jedes Funktional $T \in \mathcal{L}(H; \mathbb{R})$ in der Form $T(h) = \langle a, h \rangle$ für ein $a \in H$ und alle $h \in H$ dargestellt werden kann. Der Vektor $a \in H$ ist hierbei eindeutig bestimmt. Insbesondere, existiert ein Vektor $\nabla f(x) \in H$ — der Gradient von f im Punkt x — mit

$$Df(x)(h) = \langle \nabla f(x), h \rangle \quad \text{für alle } h \in H.$$

Der Gradient ist vom Skalarprodukt abhängig: Ist $\langle \cdot, \cdot \rangle^*$ ein äquivalentes Skalarprodukt auf H und $a^* \in H$ so gewählt, dass $\langle a^*, h \rangle^* = Df(x)(h)$ für alle $h \in H$, dann gilt im Allgemeinen nicht $a^* = \nabla f(x)$!

24.5.2020

#26

24.5.2020

2.5 Taylorentwicklung

Wir betrachten zunächst eine Funktion $f \in C^\infty(\mathbb{R})$. Wir haben gesehen, dass f in einem Punkt $a \in \mathbb{R}$ durch die affine Funktion $y \mapsto f(a) + f'(a)(y - a)$ bestmöglich approximiert wird. Affine Funktionen sind Polynome vom Grad 1. Wir möchten nun f bei x durch Polynome vom Grad k approximieren. Da wir mehr Freiheitsgrade erlauben, erwarten wir, dass f besser genähert werden kann.

Definition 2.77 (Taylorpolynom). Sei $I \subseteq \mathbb{R}$ offenes Intervall, $f \in C^k(I)$ und $a \in I$. Dann heißt das durch

$$T_k(x; a) := \sum_{\ell=0}^k \frac{f^{(\ell)}(a)}{\ell!} (x - a)^\ell, \quad x \in \mathbb{R},$$

definierte Polynom das k -te Taylorpolynom (auch: Taylorpolynom der Ordnung k) von f im Punkt a .

Beispiel 2.78.

(a) 3-te Taylorpolynom im Punkt a :

$$T_3(x; a) = f(a) + f'(a)(x - a) + \frac{1}{2}f''(a)(x - a)^2 + \frac{1}{6}f'''(a)(x - a)^3.$$

(b) $f(x) = x^2$,

$$f^{(0)}(a) = f(a), \quad f^{(1)}(a) = f'(a), \quad f^{(2)}(a), \quad f^{(\ell)}(a) = 0 \text{ für } \ell > 2.$$

Taylorpolynom im Punkt a :

$$T_1(x; a) = a^2 + 2a(x - a), \quad T_k(x; a) = a^2 + 2a(x - a) + \frac{1}{2}2(x - a)^2 \quad (k \geq 2)$$

Beobachtung: $T_k(x; a) = x^2$ für $k \geq 2$.

(c) $f(x) = \exp(x)$, Taylorpolynom bei $a = 0$. Wegen $\exp'(x) = \exp(x)$ und $\exp(0) = 1$ gilt¹⁹

$$T_k(x; 0) = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \cdots + \frac{1}{k!}x^k = \sum_{\ell=0}^k \frac{1}{\ell!}x^\ell.$$

Dies ist die abgebrochene Reihenentwicklung der Exponentialfunktion! $\rightarrow$ Die Reihenentwicklung der Exponentialfunktion kann als Eselsbrücke für die Definition des Taylorpolynoms dienen.

(d) $f(x) = \sin(x)$, $a = 0$.

$$\begin{aligned} T_0(x; 0) &= \sin(0) = 0 \\ T_1(x; 0) &= 0 + \cos(0)x = x \\ T_2(x; 0) &= 0 + \cos(0)x - \frac{1}{2}\sin(0)x^2 = x \\ T_3(x; 0) &= 0 + \cos(0)x - \frac{1}{2}\sin(0)x^2 - \frac{1}{6}\cos(0)x^3 = x - \frac{1}{6}x^3. \end{aligned}$$

¹⁹Korrektur: 29.05.2020

Bemerkung 2.79 (Geometrische Bedeutung). • Der Graph des Taylorpolynoms der Ordnung 1 an f im Punkt a ist die Tangente an den Graphen von f im Punkt $(a, f(a))$.

- Der Graph des Taylorpolynoms der Ordnung 2 an f im Punkt a ist eine Parabel. Für $\frac{1}{2}f''(a) > 0$ ist die Parabel nach oben geöffnet, für $\frac{1}{2}f''(a) < 0$ ist die Parabel nach unten geöffnet. Im Fall $\frac{1}{2}f''(a) = 0$ ist die Parabel mit der Tangente an den Graphen von f im Punkt $(a, f(a))$ identisch.
- Der Graph des Taylorpolynoms der Ordnung 2 wird auch Schmiegeparabel an den Graphen von f im Punkt $(a, f(a))$ genannt.

Satz 2.80 (Eindimensionale Taylorentwicklung). Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $f \in C^{k+1}(I)$, $a, x \in I$ mit $x \neq a$. Dann existiert $\tau \in (a, x) \cup (x, a)$ mit

$$f(x) = \underbrace{\sum_{\ell=0}^k \frac{f^{(\ell)}(a)}{\ell!} (x-a)^\ell}_{\text{Taylorpolynom}} + \underbrace{\frac{1}{(k+1)!} f^{(k+1)}(\tau) (x-a)^{k+1}}_{=\text{Restglied}}$$

Bemerkung 2.81. Es gibt verschiedene Möglichkeit das Restglied aufzuschreiben. Das in Satz 2.80 verwendete Restglied heißt Lagrangesches Restglied.

Beweisidee.

Der Fall $k = 0$ entspricht dem Mittelwertsatz.

Der Fall $k = 1$ entspricht der Weierstraßschen Zerlegungsformel.

Der Fall $k > 1$ folgt durch Iteration aus dem Mittelwertsatz.

Beweis. O.B.d.A. sei $x > a$. Setze

$$\begin{aligned} R(x) &:= f(x) - \sum_{\ell=0}^k \frac{f^{(\ell)}(a)}{\ell!} (x-a)^\ell \\ Q(x) &:= \frac{(x-a)^{k+1}}{(k+1)!}. \end{aligned}$$

Dann gilt

$$\begin{aligned} R^{(\ell)}(a) &= 0 && \text{für alle } 0 \leq \ell \leq k, \\ Q_n^{(\ell)}(x) &= 0 \Leftrightarrow x = a && \text{für alle } 0 \leq \ell \leq k, \\ R^{(k+1)}(x) &= f^{(k+1)}(x) && \text{für alle } x \in I, \\ Q^{(k+1)}(x) &= 1 && \text{für alle } x \in I \end{aligned}$$

Es genügt zu zeigen dass

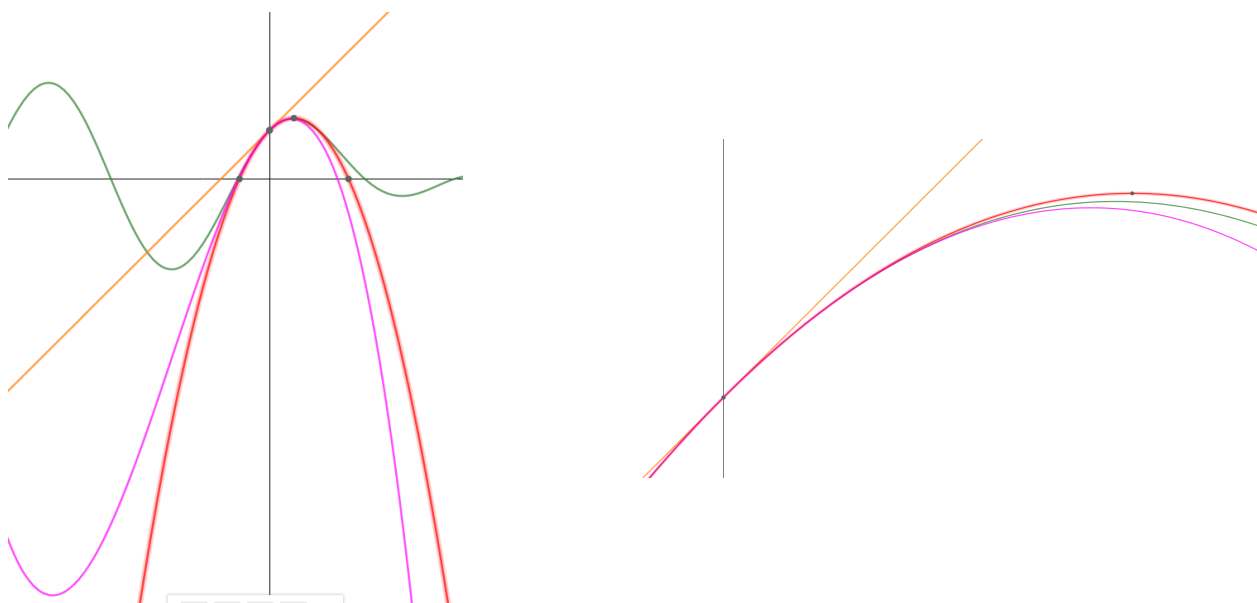
$$\frac{R(x)}{Q(x)} = f^{(k+1)}(\tau_{k+1})$$

für ein $\tau_{k+1} \in (a, x)$.

Aus einer iterierten Anwendung des Mittelwertsatzes (AN10, Satz 7.15) folgt die Existenz von Zwischenstellen $a < \tau_k < \dots < \tau_1 < x$ mit

$$\begin{aligned} \frac{R(x)}{Q(x)} &= \frac{R(x) - R(a)}{Q(x) - Q(a)} \\ &= \frac{R'(\tau_1)}{Q'(\tau_1)} = \frac{R'(\tau_1) - R'(a)}{Q'(\tau_1) - Q'(a)} \\ &= \frac{R''(\tau_2)}{Q''(\tau_2)} = \frac{R''(\tau_2) - R''(a)}{Q''(\tau_2) - Q''(a)} \\ &= \dots \\ &= \frac{R^{(k+1)}(\tau_{k+1})}{Q^{(k+1)}(\tau_{k+1})} = f^{(k+1)}(\tau_{k+1}). \end{aligned}$$

□



Graph(f) grün, Graph($T_1(\cdot; 0)$) orange, Graph($T_2(\cdot; 0)$) rot, Graph($T_3(\cdot; 0)$) magenta.

Wir verallgemeinern die Taylorentwicklung auf Skalarfelder $f : \mathbb{R}^n \rightarrow \mathbb{R}$.

Idee: Wende die eindimensionale Taylorentwicklung auf die Funktion

$$g : [0, 1] \rightarrow \mathbb{R}, \quad g(t) := f(a + th)$$

an.

Ableitungen von g lassen sich als Linearkombination von partiellen Ableitungen von f ausdrücken. Zur “Buchhaltung” der auftretenden höheren partiellen Ableitung führen wird die Multiindex-Notation ein:

Definition 2.82 (Multiindex).

- Ein $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ heißt *Multiindex*. Für einen Multiindex $\alpha = (\alpha_1, \dots, \alpha_n)$ definieren wir die *Ordnung* $|\alpha| = \sum_{j=1}^n \alpha_j$ und die *Fakultät*

$$\alpha! := \alpha_1! \cdots \alpha_n!.$$

- Für $f \in C^{|\alpha|}(\Omega)$, $\Omega \subseteq \mathbb{R}^n$ offen, setzen wir

$$D^\alpha f(x) := \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n} f(x),$$

mit

$$\left(\frac{\partial}{\partial x_j}\right)^0 = 1, \quad \left(\frac{\partial}{\partial x_j}\right)^\ell = \underbrace{\frac{\partial}{\partial x_j} \cdots \frac{\partial}{\partial x_j}}_{\ell\text{-mal}}$$

- Für $x \in \mathbb{R}^n$ definieren wir

$$x^\alpha := \prod_{j=1}^n x_j^{\alpha_j}.$$

Beispiel: $n = 3$, $\alpha = (2, 1, 0)$,

$$|\alpha| = 3, \quad \alpha! = 2, \quad x^\alpha = x_1^2 x_2 x_3^0 = x_1^2 x_2.$$

$$D^\alpha f(x) = \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} f(x)$$

Lemma 2.83. Sei $\Omega \subseteq \mathbb{R}^n$ offen, $f \in C^k(\Omega)$, $x \in \Omega$, $h \in \mathbb{R}^n$ mit $[[x, x+h]] \subseteq \Omega$. Dann existiert ein offenes Intervall I mit $[0, 1] \subseteq I$ und durch

$$g(t) := f(x + th)$$

wird eine Funktion $g \in C^k(I)$ definiert mit

$$g^{(k)}(t) = \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha|=k}} \frac{k!}{\alpha!} D^\alpha f(x + th) h^\alpha.$$

Beweis. Da Ω offen ist, existiert $\delta > 0$ mit $x + th \in \Omega$ für alle $t \in (-\delta, 1 + \delta) =: I$.

Aus der Kettenregel folgt, dass g differenzierbar ist und

$$\begin{aligned} g'(t) &= \nabla f(x + th) \cdot h = \sum_{j=1}^n \frac{\partial}{\partial x_j} f(x + th) h_j \\ &= \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha|=1}} D^\alpha f(x + th) h^\alpha. \end{aligned}$$

Eine weitere Anwendung der Kettenregel auf g' zeigt, dass g zweimal differenzierbar ist und

$$(2.16) \quad g''(t) = (\nabla f(x + th) \cdot h)' = \sum_{\ell=1}^n \sum_{j=1}^n \frac{\partial}{\partial x_\ell} \frac{\partial}{\partial x_j} f(x + th) h_j h_\ell.$$

Die Summanden sind symmetrisch in j und ℓ , d.h. Terme mit Indizes $j \neq \ell$ kommen in der Summe zweifach vor (das ist genau dann der Fall, wenn $\max\{\alpha_1, \dots, \alpha_n\} = 1$). Da

$$\frac{2!}{\alpha!} = \begin{cases} 1 & \text{für } \max_{1 \leq j \leq n} \{\alpha_j\} = 2, \\ 2 & \text{für } \max_{1 \leq j \leq n} \{\alpha_j\} = 1, \end{cases}$$

folgt

$$(2.17) \quad g''(t) = \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha|=2}} \frac{2!}{|\alpha|} D^\alpha f(x+th) h^\alpha.$$

Die Argumentation kann induktiv fortgesetzt werden: Mit vollständiger Induktion folgt aus der Kettenregel die Differenzierbarkeit von $g^{(k)}$ und die Formel

$$g^{(k)}(t) = \sum_{j_1=1}^n \cdots \sum_{j_k=1}^n \frac{\partial}{\partial x_{j_1}} \cdots \frac{\partial}{\partial x_{j_k}} f(x+th) h_{j_k} \cdots h_{j_1}.$$

Jede der k -fachen partiellen Ableitungen entspricht nach Umsortieren und Zusammenfassen einer Ableitung der Form $D^\alpha f(x+th)$ mit einem Multiindex α der Ordnung k . Jeder solche Multiindex kommt in der k -fachen Summe mit der Anzahl

$$\binom{k}{\alpha_1} \binom{k-\alpha_1}{\alpha_2} \cdots \binom{k-\alpha_1-\cdots-\alpha_{n-1}}{\alpha_n} = \frac{k!}{\alpha_1! \alpha_2! \cdots \alpha_n!} = \frac{k!}{\alpha!}$$

vor. □

Satz 2.84 (Taylorformel für Skalarfelder). *Sei $\Omega \subseteq \mathbb{R}^n$ offen, $f \in C^{k+1}(\Omega)$, $a, x \in \Omega$ mit $[[a, x]] \subseteq \Omega$. Dann gibt es ein $\xi \in [[a, x]]$ mit*

$$f(x) = \underbrace{\sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| \leq k}} \frac{1}{\alpha!} D^\alpha f(a) (x-a)^\alpha}_{\text{Taylorpolynom}} + \underbrace{\sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha|=k+1}} \frac{1}{\alpha!} D^\alpha f(\xi) (x-a)^\alpha}_{\text{Restglied}}.$$

Das durch

$$T_k(x; a) := \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| \leq k}} \frac{1}{\alpha!} D^\alpha f(a) (x-a)^\alpha$$

definierte Polynom $T_k(\cdot; a) : \mathbb{R}^n \rightarrow \mathbb{R}$ heißt k -tes Taylorpolynom von f im Punkt a .

Beweis. Für $g(t) := f(a + t(x-a))$ existiert nach Satz 2.80 ein $\tau \in (0, 1)$ mit

$$\begin{aligned} f(x) &= g(1) = \sum_{\ell=0}^k \frac{g^{(\ell)}(0)}{\ell!} + \frac{1}{(k+1)!} g^{(k+1)}(\tau) \\ &\stackrel{\text{Lemma 2.83}}{=} \sum_{\ell=0}^k \frac{1}{\ell!} \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha|=\ell}} \frac{\ell!}{\alpha!} D^\alpha f(a) (x-a)^\alpha \\ &\quad + \frac{1}{(k+1)!} \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha|=k+1}} \frac{(k+1)!}{\alpha!} D^\alpha \underbrace{(a + \tau(x-a))}_{=:\xi} (x-a)^\alpha. \end{aligned}$$

□

Definition 2.85 (Hesse-Matrix). Sei $\Omega \subseteq \mathbb{R}^n$ offen und $f : \Omega \rightarrow \mathbb{R}$ zweimal partiell differenzierbar in $x \in \Omega$. Die durch

$$H_f(x) := \begin{pmatrix} \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_1} f(x) & \cdots & \frac{\partial}{\partial x_n} \frac{\partial}{\partial x_1} f(x) \\ \vdots & & \vdots \\ \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_n} f(x) & \cdots & \frac{\partial}{\partial x_n} \frac{\partial}{\partial x_n} f(x) \end{pmatrix}$$

definierte Matrix in $\mathbb{R}^{n \times n}$ heißt Hesse-Matrix von f im Punkt x .

Bemerkung 2.86. Der erste und zweite Entwicklungsterm in der Taylorentwicklung lassen sich mit Hilfe des Gradienten und der Hesse-Matrix darstellen:

$$\sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha|=1}} \frac{1}{\alpha!} D^\alpha f(a)(x-a)^\alpha = \nabla f(a) \cdot (x-a) \quad (1. \text{ Entwicklungsterm}),$$

$$\sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha|=2}} \frac{1}{\alpha!} D^\alpha f(a)(x-a)^\alpha = \frac{1}{2} (x-a) \cdot H_f(a)(x-a) \quad (2. \text{ Entwicklungsterm}).$$

Man beachte:

$$\sum_{\ell=1}^n \sum_{j=1}^n \frac{\partial}{\partial x_\ell} \frac{\partial}{\partial x_j} f(x) h_j h_\ell = h \cdot H_f(x) h.$$

Bemerkung 2.87. Ist $f \in C^2(\Omega)$, dann ist $H_f(a)$ symmetrisch (Lemma von Schwarz!). Sei $a \in \Omega$ mit $B(a; r) \subseteq \Omega$. Dann existiert nach der Taylorformel für Skalarfelder für jedes $h \in B(0; r)$ eine Zwischenstelle $\xi(h) \in [[a, a+h]]$ mit

$$f(a+h) = f(a) + \nabla f(a) \cdot h + \frac{1}{2} h \cdot H_f(\xi(h)) h.$$

Definiere

$$\rho(h) := \frac{1}{2 \|h\|_2^2} \left(h \cdot (H_f(\xi(h)) - H_f(a)) h \right) \quad \text{für } h \neq 0$$

und $\rho(0) := 0$ sonst. Dann gilt

$$f(a+h) = f(a) + \nabla f(a) \cdot h + \frac{1}{2} h \cdot H_f(a) h + \rho(h) \|h\|_2^2,$$

und

$$\lim_{h \rightarrow 0} \rho(h) = 0.$$

Das Restglied $\rho(h) \|h\|_2^2$ heißt Peano-Restglied.

Letzteres sieht man wie folgt: Sei $(h_k) \subseteq B(0; r)$ mit $h_k \rightarrow 0$. Setze $\xi_k := \xi(h_k)$. Dann gilt wegen $\xi_k \in [[a, a+h_k]]$, dass $\xi_k \rightarrow 0$. Da die zweiten partiellen Ableitungen von f stetig in a sind, folgt $H_f(\xi_k) \rightarrow H_f(a)$. Also folgt (mit der Cauchy-Schwarz Ungleichung)

$$\begin{aligned} \left| \frac{h_k}{\|h_k\|_2} \cdot (H_f(\xi_k) - H_f(a)) \frac{h_k}{\|h_k\|_2} \right| &\leq \left\| \frac{h_k}{\|h_k\|_2} \right\|_2 \| (H_f(\xi_k) - H_f(a)) \frac{h_k}{\|h_k\|_2} \|_2 \\ &\leq \|H_f(\xi_k) - H_f(a)\| \rightarrow 0. \end{aligned}$$

Korollar 2.88 (Quadratische Taylorentwicklung von Skalarfeldern). Sei $\Omega \subseteq \mathbb{R}^n$ offen, $f \in C^2(\Omega)$ und $B(a; r) \subseteq \Omega$. Dann existiert $\rho : B(0; r) \rightarrow \mathbb{R}$ mit $\lim_{h \rightarrow 0} \rho(h) = 0$ und

$$f(a + h) = f(a) + \nabla f(a) \cdot h + \frac{1}{2} h \cdot H_f(a) h + \rho(h) \|h\|_2^2$$

für alle $h \in B(0; r)$.

Bemerkung 2.89 (Geometrische Bedeutung).

(a) Für $k = 1$ beschreibt der Graph des Taylorpolynoms der Ordnung 1,

$$\{(x, x_{n+1}) : x_{n+1} = f(a) + \nabla f(a) \cdot (x - a), x \in \mathbb{R}^n\} \subseteq \mathbb{R}^{n+1}$$

die Tangentialebene an den Graphen von f im Punkt $(a, f(a))$.

(b) Für $k = 2$ beschreibt der Graph des Taylorpolynoms der Ordnung 2,

$$\{(x, x_{n+1}) : x_{n+1} = f(a) + \nabla f(a) \cdot (x - a) + \frac{1}{2} (x - a) \cdot H_f(a) (x - a), x \in \mathbb{R}^n\} \subseteq \mathbb{R}^{n+1}$$

eine Quadrik, die sich an den Graphen von f im Punkt $(a, f(a))$ anschmiegt. Sie wird Schmiegequadrik genannt.

Bemerkung 2.90. Sei $n = 2$, $\Omega \subseteq \mathbb{R}^2$, $f \in C^2(\Omega)$, $a \in \Omega$. Die Hesse-Matrix $H_f(a)$ ist symmetrisch. Nach dem Spektralsatz für symmetrische, reelle Matrizen sind alle Eigenwerte κ_1, κ_2 reell und es existiert eine ONB aus Eigenvektoren.

- Falls $\kappa_1 = \kappa_2 = 0$, dann ist die Schmiegequadrik flach.
Beispiel: $f(x) = e \cdot x$ für ein $e \in \mathbb{R}^n \setminus \{0\}$.
- Falls genau ein Eigenwert 0 ist, dann ist die Schmiegequadrik ein parabolischer Zylinder.
Beispiel: $f(x) = \pm x_1^2$.
- Falls beide Eigenwerte strikt positiv oder strikt negativ sind, dann ist die Schmiegequadrik ein elliptisches Paraboloid.
Beispiel: $f(x) = \pm(x_1^2 + x_2^2)$.
- Falls ein Eigenwert strikt negativ und der andere strikt positiv ist, dann ist die Schmiegequadrik ein hyperbolischer Paraboloid.
Beispiel: $f(x) = x_1^2 - x_2^2$.

Kontrollfrage 8.1. Sei $f \in C^2(\mathbb{R})$ und

$$g : \mathbb{R} \rightarrow \mathbb{R}, \quad g(x) := f(0) + f'(0)x.$$

Welche der nachfolgenden Aussagen sind richtig?

(A) Es gilt $\lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{f(h) - g(h)}{h} = 0$.

✓ : Definition der Ableitung durch Differenzenquotienten plus Stetigkeit. (Man beachte, dass g ein Taylorpolynom ersten Grades von f ist.)

(B) Es gilt $f(h) = g(h) + f'(\tau)$ für ein τ mit $|\tau| < |h|$.

×: Gegenbeispiel: $f(x) = x$.

(C) Es gilt $\lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{f(h) - g(h)}{h^2} = 0$.

×: Gegenbeispiel: Quadratische Funktion. ($f(x) = x^2$)

(D) Es gilt $\lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{f(h) - g(h) - ch^2}{h^2} = 0$ für ein $c \in \mathbb{R}$.

✓: $h \mapsto g(h) + \frac{1}{2}f''(0)h^2$ ist das Taylorpolynom zweiten Grades von f um 0. Wähle also $\frac{1}{2}f''(0) = c$ und nutze Satz 2.80.

Kontrollfrage 8.2. Welche der nachfolgenden Aussagen sind richtig?

(A) Das k te Taylorpolynom von $x \mapsto \exp(x)$ an der Stelle $x = 0$ ist

$$T_k(x; 0) = \sum_{\ell=0}^k \frac{x^\ell}{\ell!}$$

✓: $\exp(0) = 1$.

(B) Das k te Taylorpolynom von $x \mapsto \exp(x)$ an der Stelle $x = 2$ ist

$$T_k(x; 2) = \exp(2) \sum_{\ell=0}^k \frac{(x-2)^\ell}{\ell!}$$

✓: Siehe Definition 2.77.

Kontrollfrage 8.3. Wir betrachten

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x) = x_1^3 + x_1^2 + 2x_1x_2 + x_1 + 2x_2 + 1.$$

Welche der nachfolgenden Aussagen sind wahr? Das zweite Taylorpolynom von f im Punkt $(0, 0)$ ist gleich

(A) 1.

×

(B) $x_1 + 2x_2 + 1$.

×

(C) $x_1^2 + 2x_1x_2 + x_1 + 2x_2 + 1$.

✓

(D) $x_1^3 + x_1^2 + 2x_1x_2 + x_1 + 2x_2 + 1$.

×

Kontrollfrage 8.4. Sei $f \in C^3(\mathbb{R})$. Es gelte

$$f'(0) = f''(0) = 0 \quad \text{und} \quad f'''(0) > 0.$$

Welche der nachfolgenden Aussagen sind wahr?

(A) Es gilt

$$\lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{f(h)}{h^3} = \frac{1}{6} f'''(0)$$

×: Wenn auch noch $f(0) = 0$ gegeben wäre, würde die Aussage aus Satz 2.80 folgen. So ist sie aber im Allgemeinen falsch. (Es könnte als Grenzwert auch ∞ herauskommen.)

(B) f hat in 0 ein lokales Maximum.

×: Gegenbeispiel: $x \mapsto x^3$.

(C) f hat in 0 ein lokales Minimum

×: Gegenbeispiel: $x \mapsto x^3$.

(D) f kann in 0 kein lokales Extremum besitzen.

✓: Aus Satz 2.80 folgt

$$\forall h \in \mathbb{R} \exists \xi \text{ mit } |\xi| < |h| : f(h) = f(0) + \frac{1}{6} h^3 f'''(\xi).$$

Da aber f''' stetig ist und $f(0) = 0$ ist, gilt

$$\exists \epsilon > 0 \forall h \in (-\epsilon, \epsilon) \exists c > 0 : f(h) = f(0) + \frac{1}{6} h^3 c.$$

Die Behauptung ergibt sich nun daraus, dass der Term „ h^3 “ in der letzten Aussage negative und positive Werte annimmt.

24.5.2020

#27

24.5.2020

2.6 Lokale Extrema von Skalarfeldern

Ziel:

- Kriterien für lokale Extrema für Skalarfelder mit Hilfe des Gradienten und der Hesse-Matrix

Definition 2.91 (lokale Extrema). Sei $\Omega \subset \mathbb{R}^n$ offen, $f : \Omega \rightarrow \mathbb{R}$, $a \in \Omega$. Falls ein $r > 0$ existiert mit

$$\left. \begin{array}{l} f(a) \geq f(x) \\ f(a) \leq f(x) \end{array} \right\} \text{ für alle } x \in B(a; r), \text{ so heißt } a \text{ lokales } \begin{cases} \text{Maximum,} \\ \text{Minimum.} \end{cases}$$

Falls ein $r > 0$ existiert mit

$$\left. \begin{array}{l} f(a) > f(x) \\ f(a) < f(x) \end{array} \right\} \text{ für alle } x \in B(a; r), \text{ so heißt } a \text{ isoliertes, lokales } \begin{cases} \text{Maximum,} \\ \text{Minimum.} \end{cases}$$

$a \in \Omega$ heißt (isoliertes) lokales Extremum, falls a ein (isoliertes) lokales Minimum/Maximum ist.

Beispiel 2.92. Betrachte $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x) = -\cos(x_1) + cx_2^2$ mit $c \in \mathbb{R}$.

Der Punkt $a = (0, 0)$ ist

$$\begin{cases} \text{lokales Minimum} & \text{falls } c \geq 0 \\ \text{lokales, isoliertes Minimum} & \text{falls } c > 0 \\ \text{kein lokales Minimum} & \text{falls } c < 0 \end{cases}$$

Satz 2.93 (Notwendiges Kriterium für lokale Extrema). Sei $\Omega \subseteq \mathbb{R}^n$ offen, $f : \Omega \rightarrow \mathbb{R}$ partiell differenzierbar. Ist $a \in \Omega$ ein lokales Extremum von f , dann gilt $\nabla f(a) = 0$.

Beweis. Ist a ein lokales Extremum von f , dann ist $t = 0$ ein Extremum von $g : (-r, r) \rightarrow \mathbb{R}$, $g(t) := f(a + te_j)$ für hinreichend kleines $r > 0$. Es folgt $0 = g'(0) = \frac{\partial}{\partial x_j} f(a)$. $\square$

Bemerkung 2.94. Ein Punkt $a \in \Omega$ mit $\nabla f(a) = 0$ heißt stationärer Punkt von f .

Bemerkung 2.95. Eine Matrix $A \in \mathbb{R}^{n \times n}$ heißt

$$\begin{array}{lll} \text{positiv definit} & \Leftrightarrow & h \cdot Ah > 0 \text{ für alle } h \in \mathbb{R}^n \setminus \{0\}, \\ \text{positiv semidefinit} & \Leftrightarrow & h \cdot Ah \geq 0 \text{ für alle } h \in \mathbb{R}^n, \\ \text{negativ definit} & \Leftrightarrow & h \cdot Ah < 0 \text{ für alle } h \in \mathbb{R}^n \setminus \{0\}, \\ \text{negativ semidefinit} & \Leftrightarrow & h \cdot Ah \leq 0 \text{ für alle } h \neq 0, \\ \text{indefinit} & \Leftrightarrow & A \text{ ist weder positiv semidefinit noch negativ semidefinit.} \end{array}$$

In der Linearen Algebra wird bewiesen:

$$\begin{array}{lll} A \text{ positiv definit} & \Leftrightarrow & \text{alle Eigenwerte von } A \text{ sind strikt positiv,} \\ A \text{ positiv semidefinit} & \Leftrightarrow & \text{alle Eigenwerte von } A \text{ sind nicht negativ} \\ A \text{ indefinit} & \Leftrightarrow & A \text{ besitzt mind. einen strikt positiven} \\ & & \text{und mind. einen strikt negativen Eigenwerte.} \end{array}$$

Satz 2.96 (Kriterien für lokale Extrema). Sei $\Omega \subseteq \mathbb{R}^n$ offen, $f \in C^2(\Omega)$, $a \in \Omega$.

(a) (Notwendige Bedingung). Es gilt:

$$\begin{array}{ll} a \text{ lokales Minimum} & \Rightarrow H_f(a) \text{ ist positiv semidefinit} \\ a \text{ lokales Maximum} & \Rightarrow H_f(a) \text{ ist negativ semidefinit} \end{array}$$

(b) (Hinreichende Bedingung). Ist außerdem $\nabla f(a) = 0$, dann gilt

$$\begin{array}{ll} a \text{ isoliertes lokales Minimum} & \Leftarrow H_f(a) \text{ ist positiv definit,} \\ a \text{ isoliertes lokales Maximum} & \Leftarrow H_f(a) \text{ ist negativ definit,} \\ a \text{ ist kein lokales Extremum} & \Leftarrow H_f(a) \text{ ist indefinit.} \end{array}$$

Beweis. Es genügt die Aussagen zum Minimum zu beweisen (andernfalls betrachte $-f$).

1. Sei a ein lokales Minimum. Wähle $r > 0$, so dass $B(a; r) \subseteq \Omega$ und

$$f(a) \leq f(a + h) \quad \text{für alle } h \in B(0; r).$$

Gemäß Taylorentwicklung gilt

$$f(a) \leq f(a + h) = f(a) + \nabla f(a) \cdot h + \frac{1}{2} h \cdot H_f(a) h + \rho(h) \|h\|_2^2,$$

mit $\rho(h) \rightarrow 0$ für $h \rightarrow 0$.

Da a lokales Minimum ist, gilt $\nabla f(a) = 0$. Mit einer Substitution $h = \delta e$ mit $e \in \mathbb{R}^n$ und $\delta > 0$ klein genug, folgt

$$-\rho(\delta a) \delta^2 \|e\|_2^2 \leq \frac{\delta^2}{2} e \cdot H_f(a) e.$$

Der Grenzübergang $\delta \rightarrow 0$ führt auf

$$0 \leq \frac{1}{2} h \cdot H_f(a) h$$

für alle $e \in \mathbb{R}^n$. Also ist die Hesse-Matrix positiv semidefinit.

2. Sei nun $H_f(a)$ positiv definit. Dann gilt

$$m := \min_{\substack{h \in \mathbb{R}^n \\ \|h\|_2=1}} h \cdot H_f(a) h > 0.$$

Gemäß Taylorentwicklung und der Voraussetzung $\nabla f(a) = 0$ gilt

$$f(a) \leq f(a + h) = f(a) + \frac{1}{2} h \cdot H_f(a) h + \rho(h) \|h\|_2^2,$$

mit $\rho(h) \rightarrow 0$ für $h \rightarrow 0$.

Wähle nun $\varepsilon > 0$ so klein, dass

$$\rho(h) \leq \frac{1}{4} m \quad \text{für alle } h \in B(0; \varepsilon).$$

Dann gilt für $h \in B(0; \varepsilon) \setminus \{0\}$:

$$f(a + h) \geq f(a) + \frac{1}{2} m \|h\|_2^2 - \frac{1}{4} m \|h\|_2^2 = f(a) + \frac{1}{4} m \|h\|_2^2 > f(a)$$

also ist a ein isoliertes lokales Minimum.

3. Seien v_1, v_2 normierte Eigenvektoren von $H_f(a)$ zu Eigenwerten $\kappa_1 < 0 < \kappa_2$. Betrachte

$$g_i(t) := f(a + t v_i).$$

Dann gilt

$$g'_i(0) = \nabla f(a) \cdot v_i = 0, \quad g''_i(0) = \frac{1}{2} v_i \cdot H_f(a) v_i = \frac{1}{2} \kappa_i.$$

Also hat g_1 bei $t = 0$ ein lokales isoliertes Maximum und g_2 bei $t = 0$ ein lokales isoliertes Minimum.

Damit kann a kein Extremum von f sein.

□

Argument für:

$$a \text{ lokales Minimum} \quad \Rightarrow \quad H_f(a) \text{ ist positiv semidefinit.}$$

Argument für:

$$a \text{ isoliertes lokales Minimum} \quad \Leftarrow \quad H_f(a) \text{ ist positiv definit.}$$

Argument für:

$$a \text{ ist kein lokales Extremum} \quad \Leftarrow \quad H_f(a) \text{ ist indefinit.}$$

Beispiele 2.97. ²⁰

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x) := x_2^2(x_1 - 1) + x_1^2(x_1 + 1).$$

$$\nabla f(x) = (x_2^2 + 3x_1^2 + 2x_1, 2(x_1 - 1)x_2), \quad H_f(x) = \begin{pmatrix} 6x_1 + 2 & 2x_2 \\ 2x_2 & 2(x_1 - 1) \end{pmatrix}$$

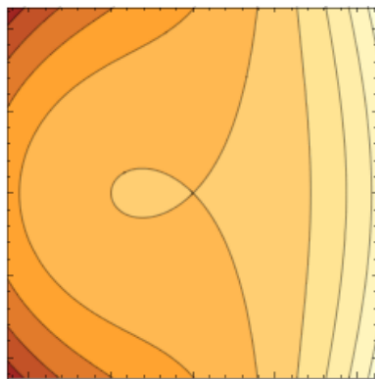
Wir möchten Extremstellen bestimmen.

- Stationäre Punkte bestimmen ergibt

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

- “Vorzeichen” von $H_f(P_i)$ überprüfen:

$$H_f(P_1) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}, \quad H_f(P_2) = \begin{pmatrix} -2 & 0 \\ 0 & -\frac{10}{3} \end{pmatrix}$$

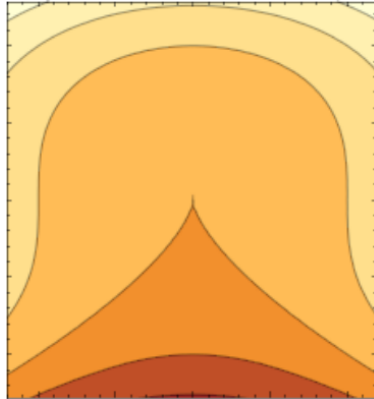


Konturplot von f .

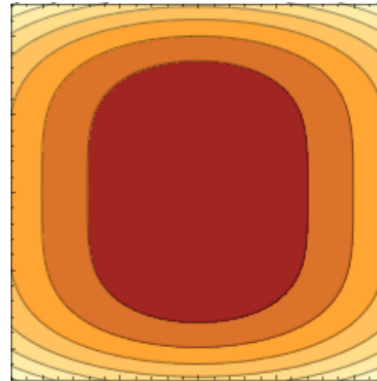
Bemerkung 2.98 (Semidefinite Hesse-Matrix). *Ist $H_f(a)$ semidefinit, so kann a ein lokales Extremum sein oder auch nicht.*

²⁰Korrektur: 29.05.2020

- $f(x) = x_1^2 + x_2^3$ gilt $\nabla f(x) = (2x_1, 3x_2^2)$ und $H_f(x) = \begin{pmatrix} 2 & 0 \\ 0 & 6x_2 \end{pmatrix}$. Also ist $a = (0, 0)$ ein stationärer Punkt mit $H_f(a)$ positiv semidefinit. Der Punkt a ist aber kein lokales Minimum.
- $f(x) = x_1^2 + x_2^4$ gilt $\nabla f(x) = (2x_1, 4x_2^3)$ und $H_f(x) = \begin{pmatrix} 2 & 0 \\ 0 & 12x_2^2 \end{pmatrix}$. Also ist $a = (0, 0)$ ein stationärer Punkt mit $H_f(a)$ positiv semidefinit. Der Punkt a ist globales isoliertes Minimum.



$$x \mapsto x_1^2 + x_2^3$$



$$x \mapsto x_1^2 + x_2^4$$

Kontrollfrage 8.5. Wir betrachten die Funktion

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x) := (x_1 - 1)^2 - (x_2 - 2)^2.$$

Welche der nachfolgenden Aussagen sind wahr?

(A) Der Graph von f ist ein hyperbolisches Paraboloid.

✓: f ist quadratisch und also identisch mit seiner Schmiegequadratik. Wenn man nun in Bemerkung 2.90 das a auf $(1, -2)$ setzt, erhält die Behauptung mit

$$H_f(a) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

(Denn die Eigenwerte sind -2 und 2 .)

(B) $(1, 2)$ ist ein stationärer Punkt von f .

✓: Der Wert von

$$\nabla f(x) = 2 \begin{pmatrix} x_1 - 1 \\ -x_2 + 2 \end{pmatrix}$$

ist $(0, 0)$ für $x = (1, 2)$.

(C) Wegen $\nabla f(1, 2) = 0$ hat f im Punkt $(1, 2)$ ein lokales Extremum.

✗: Satz 2.96 a) zusammen mit (D).

(D) $H_f(1, 2)$ ist indefinit.

✓: Die Eigenwerte sind -2 und 2 .

Kontrollfrage 8.6. Sei $f \in C^2(\mathbb{R}^2)$ mit $\nabla f(0) = 0$ und $f(0) = 0$. Welche der nachfolgenden Aussagen sind wahr?

(A) Das Skalarfeld $f(x) := \|x\|_2^2$ erfüllt die Voraussetzungen.

✓: $\nabla f(x) = 2x$ für alle x .

(B) Der Tangentialraum an die Nullniveaumenge von f (im Punkt $x = 0$) ist $\mathbb{R}^2$.

✓: Für $x = 0$ ist $f(x) = 0$ und $\forall v \in \mathbb{R}^2 : v \cdot \nabla f(x) = v \cdot 0 = 0$.

(C) Der Ursprung liegt im Schnitt von $\text{Graph}(f)$ mit der Hyperebene $H = \{(t, 0) : t \in \mathbb{R}\}$.

✓: Für den Ursprung gilt $0 \in H$ und außerdem $0 \in \text{Graph}(f)$ wegen $f(0) = 0$. (Man beachte, dass wir mit "0" sowohl die Null in $\mathbb{R}$ als auch den Nullvektor bezeichnen).

(D) Die Hyperebene $H = \{(x, 0) : x \in \mathbb{R}^2\} \subseteq \mathbb{R}^3$ schneidet $\text{Graph}(f)$ nur im Ursprung.

✗: Gegenbeispiel: $f(x) = 0$ für alle $x \in \mathbb{R}$.

(E) Der Punkt $(0, 0)$ ist ein lokales Extremum von f .

✗: Gegenbeispiel: $f(x) = 0$ für alle $x \in \mathbb{R}$.

Kontrollfrage 8.7. Für

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x) = x_1^2 + x_1 + x_2^2 + x_2$$

gilt

$$Df(0, 0) = (1, 1) \text{ und } H_f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Welche der nachfolgenden Aussagen sind wahr?

(A) $H_f(0, 0)$ ist positiv definit.

✓: Siehe Bemerkung 2.95.

(B) f hat in $(0, 0)$ ein isoliertes, lokales Minimum.

✗: Siehe Satz 2.93.

(C) f hat in $(0, 0)$ ein lokales Maximum.

✗: Siehe Satz 2.93.

(D) f hat in $(0, 0)$ kein lokales Extremum.

✓: Siehe Satz 2.93.

Kontrollfrage 8.8. Für

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x) = x_2^2$$

gilt

$$Df(0,0) = (0,0) \text{ und } H_f(0,0) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}.$$

Welche der nachfolgenden Aussagen sind wahr?

(A) $H_f(0,0)$ ist positiv definit.

×: Siehe Bemerkung 2.95.

(B) $H_f(0,0)$ ist negativ semidefinit.

×: Siehe Bemerkung 2.95.

(C) $H_f(0,0)$ ist indefinit.

×: Siehe Bemerkung 2.95.

(D) f hat in $(0,0)$ ein lokales Minimum.

$$\checkmark: f(0,0) = 0 \leq x_2^2 = f(x) \text{ für alle } x \in \mathbb{R}^2.$$

(E) f hat in $(0,0)$ ein isoliertes lokales Minimum.

×: Jedes Element der Menge $\{x \in \mathbb{R}^2 : x_2 = 0\}$ ist ein lokales Minimum.

(F) f hat in $(0,0)$ kein lokales Extremum.

×: Siehe (D).

24.5.2020

#28

8.6.2020

2.7 Vektorfelder, Divergenz und Rotation

Definition 2.99 (Vektorfeld, Divergenz und Rotation). Sei $\Omega \subset \mathbb{R}^n$ offen.

(a) $F : \Omega \rightarrow \mathbb{R}^n$ heißt Vektorfeld auf Ω .

(b) $F : \Omega \rightarrow \mathbb{R}^n$ heißt Gradientenfeld, falls ein skalar Feld $f : \Omega \rightarrow \mathbb{R}$ existiert mit $F = \nabla f$. In diesem Fall heißt f Potential von F .

(c) Ist $F : \Omega \rightarrow \mathbb{R}^n$ partiell differenzierbar, dann heißt

$$\operatorname{div} F : \Omega \rightarrow \mathbb{R}, \quad \operatorname{div} F(x) := \sum_{j=1}^n \frac{\partial}{\partial x_j} F_j(x)$$

Divergenz von F .

(d) Für $n = 3$ heißt

$$\begin{aligned} \operatorname{rot} F : \Omega &\rightarrow \mathbb{R}^3, \\ \operatorname{rot} F(x) &:= \left(\frac{\partial}{\partial x_2} F_3(x) - \frac{\partial}{\partial x_3} F_2(x), \frac{\partial}{\partial x_3} F_1(x) - \frac{\partial}{\partial x_1} F_3(x), \frac{\partial}{\partial x_1} F_2(x) - \frac{\partial}{\partial x_2} F_1(x) \right). \end{aligned}$$

Rotation von F .

Beispiel 2.100 (Strömungsfeld). Sei $n = 3$. Wir nehmen an, dass eine Flüssigkeit ein Testvolumen $\Omega \subseteq \mathbb{R}^3$ durchfließt. An jedem Raumpunkt $x \in \Omega$ sei die Geschwindigkeit der Flüssigkeit zeitlich konstant (‘stationäre Strömung’) und durch einen Geschwindigkeitsvektor $F(x) \in \mathbb{R}^3$ beschrieben. Ein Partikel, den wir uns ohne Masse vorstellen, wird durch eine Flüssigkeit transportiert. Wir beschreiben die Position des Partikels mit einer Kurve $f : [a, b] \rightarrow \Omega$. Dann gilt

$$(2.18) \quad f'(t) = F(f(t)).$$

So eine Gleichung nennt man eine gewöhnliche Differentialgleichung. Wie man diese löst, lernen Sie im nächsten Semester.

Beispiel: Ein ‘Wirbel um die Achse $\{x = (0, 0, x_3) : x_3 \in \mathbb{R}\}$ ’.

$$F(x) = g(x_3)(-x_2, x_1, 0)$$

Dann erfüllt beispielsweise

$$f(t) := (\cos(g(c)t), \sin(g(c)t), c)$$

die Gleichung (2.18).

Beispiel: Eine ‘Quelle bei $x = 0$ ’

$$F(x) = x.$$

Dann erfüllt beispielsweise für $\xi \in \mathbb{R}^3$

$$f(t) := \exp(t)\xi$$

die Gleichung (2.18).

Beispiel: Eine ‘Senke bei $x = 0$ ’

$$F(x) = -x.$$

Dann erfüllt beispielsweise für $\xi \in \mathbb{R}^3$

$$f(t) := \exp(-t)\xi$$

die Gleichung (2.18).

Beispiel 2.101 (Physikalische Felder). Elektrische, magnetische Felder und Gravitationsfelder sind Beispiele für Vektorfelder im $\mathbb{R}^3$ in physikalischen Modellen.

Durch das Vektorfeld

$$g(x) = -Gm_0 \frac{x - x_0}{\|x - x_0\|_2^3}, \quad x \in \mathbb{R}^3 \setminus \{x_0\}$$

wird beispielsweise das Gravitationsfeld eines Massenpunkts im Punkt x_0 und Masse m_0 beschrieben. Auf einen weiteren Massenpunkt im Punkt $x \neq x_0$ mit Masse m wirkt dann die Gravitationskraft

$$F_G(x) = mg(x).$$

Das Vektorfeld $g : \mathbb{R}^3 \setminus \{x_0\} \rightarrow \mathbb{R}^3$ ist ein Gradientenfeld mit Potential

$$\phi(x) = -\frac{Gm_0}{\|x - x_0\|_2} + C,$$

wobei $C \in \mathbb{R}$ beliebig gewählt werden kann. (Nachrechnen!)

Es bezeichnet $u(t) \in \mathbb{R}^3 \setminus \{x_0\}$ die Position eines Massenpunktes im Gravitationsfeld g zum Zeitpunkt $t \in \mathbb{R}$. Auf ihn wirkt die Gravitationskraft

$$mg(u(t)) = -Gm_0m \frac{u(t) - x_0}{\|u(t) - x_0\|_2^3}.$$

Wir fordern nun, dass das zweite Newtonsche Gesetz gilt, also Masse $\times$ Beschleunigung = Kraft:

$$mu''(t) = mg(u(t)) \quad \text{für alle } t \in \mathbb{R}.$$

Das ist eine gewöhnliche Differentialgleichung, die beispielsweise durch die Kreisbewegung

$$u(t) = x_0 + r(\sin(vt), \cos(vt))$$

mit geeigneten Parametern $r > 0$ und $v \in \mathbb{R}$ erfüllt ist (siehe Übung).

Beispiel 2.102. Sei $\Omega \subseteq \mathbb{R}^n$ offen, seien $f : \Omega \rightarrow \mathbb{R}$, $F : \Omega \rightarrow \mathbb{R}^n$ partiell differenzierbar, dann gilt

$$\operatorname{div}(fF) = \nabla f \cdot F + f \operatorname{div} F.$$

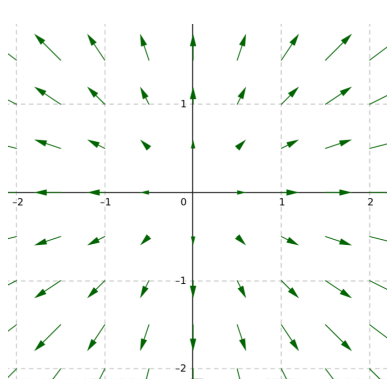
Beispiele 2.103 (Divergenz- und rotationsfreie Felder). Sei $n = 3$. Wir nennen ein Vektorfeld F divergenz- / rotationsfrei, falls $\operatorname{div} F = 0$ bzw. $\operatorname{rot} F = 0$ gilt.

(a) Sei $F(x) = Ax$ ein lineares Vektorfeld.

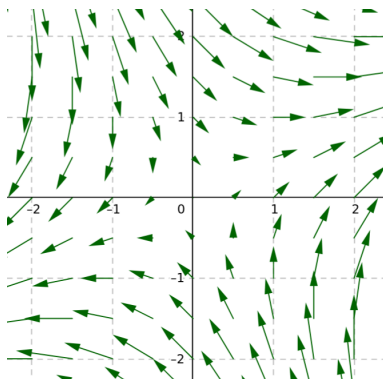
Es gilt $\frac{\partial}{\partial x_j} F_i(x) = A_{ij}$, sowie

$$F \text{ divergenzfrei} \quad \Leftrightarrow \quad \sum_{i=1}^n A_{ii} = 0 \quad (\text{spurfrei})$$

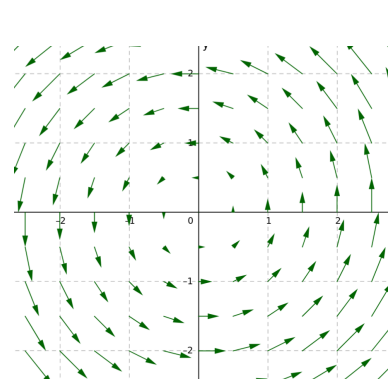
$$F \text{ rotationsfrei} \quad \Leftrightarrow \quad A^\top - A = 0 \quad (\text{symmetrisch})$$



$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$



$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$



$$A = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

(b) F Gradientenfeld mit zweimal stetig differenzierbarem Potential $\Rightarrow F$ rotationsfrei.

Anmerkung: Man kann jedes glatte Vektorfeld als Summe eines rotations- und divergenzfreien Vektorfeldes schreiben (siehe Helmholtz Zerlegung).

Divergenzfreie Gradientenfelder sind besonders schön und heißen harmonisch:

Definition 2.104. Sei $\Omega \subseteq \mathbb{R}^n$ offen. Für $u \in C^2(\Omega)$ setze

$$\Delta u(x) := \operatorname{div} \nabla u(x) := \sum_{j=1}^n \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_j} u(x).$$

Gilt $\Delta u(x) = 0$ in Ω , so heißt u harmonisch in Ω . Der Operator Δ heißt Laplace-Operator.

Beispiele 2.105. (a) Affine Funktionen $u(x) = a \cdot x + c$ sind harmonisch.

(b) Die Funktion $u(x) = x_1^2 - x_2^2$ ist harmonisch.

(c) Das Gravitationspotential $g(x) = -\frac{Gm_0}{\|x-x_0\|_2}$ ist harmonisch in $\mathbb{R}^3 \setminus \{x_0\}$.

Bemerkung 2.106. Harmonische Funktionen spielen eine fundamentale Rolle in der Theorie partieller Differentialgleichungen und in physikalischen Modellen.

Kontrollfrage 9.1. Sei

$$F : \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad F(x) := Ax \quad \text{mit } A := \begin{pmatrix} 2 & 1 & 2 \\ 1 & -1 & 2 \\ 2 & -2 & -1 \end{pmatrix}.$$

Welche der nachfolgenden Aussagen sind richtig?

(A) F ist divergenzfrei

✓: Siehe Bsp. 2.103 a).

(B) $\operatorname{rot} F(x) = 0$ für alle $x \in \mathbb{R}^3$.

✗: Siehe Bsp. 2.103 a).

(C) F ist ein Gradientenfeld

✗: Siehe Bsp. 2.103 b).

Kontrollfrage 9.2. Sei $g : \mathbb{R}^3 \rightarrow \mathbb{R}$ partiell differenzierbar und $a \in \mathbb{R}^3$. Betrachte das Vektorfeld

$$F(x) = g(x)a.$$

Welche der nachfolgenden Aussagen ist richtig?

(A) F ist divergenzfrei $\Leftrightarrow \nabla g(x) \cdot a = 0$ für alle $x \in \mathbb{R}^3$.

✓

(B) F ist rotationsfrei $\Leftrightarrow \nabla g(x) \cdot a = 0$ für alle $x \in \mathbb{R}^3$.

$\times$: Gegenbeispiel: $g : \mathbb{R}^3 \rightarrow \mathbb{R}, x \mapsto x_3$ und $a = (1, 1, 0)$. Dann $\operatorname{rot} F = (-1, 1, 0)$ und $\nabla g = (0, 0, 1)$.

(C) F ist divergenzfrei $\Leftrightarrow \nabla g(x) \times a = 0$ für alle $x \in \mathbb{R}^3$. Hierbei bezeichnet $a \times b$ das Kreuzprodukt

$$a \times b = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}$$

$\times$: Gegenbeispiel: $g : \mathbb{R}^3 \rightarrow \mathbb{R}, x \mapsto x_3$ und $a = (1, 1, 0)$. Dann $\operatorname{div} F = 0$ und $\nabla g = (0, 0, 1)$.

(D) Keine der Aussagen ist richtig.

$\times$: Siehe A.

Kontrollfrage 9.3. Welche der nachfolgenden Aussagen sind richtig?

(A) $u : \mathbb{R}^3 \rightarrow \mathbb{R}, u(x) := x_1^2 - x_2^2 + x_3^2$ ist harmonisch in $\mathbb{R}^3$.

$\times$: $\Delta u = 2$.

(B) Ist $u : \mathbb{R}^3 \rightarrow \mathbb{R}$ harmonisch, dann ist das Vektorfeld $F = \nabla u$ rotations- und divergenzfrei.

$\checkmark$: Rotationsfrei wegen Beispiel 2.103 b) und divergenzfrei wegen Definition von Δ .

(C) $u : \mathbb{R}^2 \rightarrow \mathbb{R}, u(x) := \exp(x_1) \cos(x_2)$ ist harmonisch in $\mathbb{R}^2$.

$\checkmark$: $\Delta u = 0$.

Kontrollfrage 9.4. Sei $F : \mathbb{R}^n \rightarrow \mathbb{R}^n, F(x) := x$. Welche der nachfolgenden Aussagen sind richtig?

(A) $\operatorname{div} F(x) = n$.

$\checkmark$

(B) $\operatorname{rot} F(x) = 0$.

$\checkmark$

(C) F ist ein Gradientenfeld.

$\checkmark$: Potential ist $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}, x \mapsto \frac{1}{2} \|x\|_2^2$.

8.6.2020

#29

8.6.2020

Bemerkung 2.107 (Vielfältige Schreibweisen). In der Literatur werden viele verschiedene Schreibweisen für partielle Ableitungen verwendet:

- $\frac{\partial}{\partial x_i} u(x), \frac{\partial u}{\partial x_i}(x), \partial_i u(x)$

- $\frac{\partial}{\partial x_i} u_j(x)$, $u_{j,i}(x)$
- $D^\alpha u(x)$, $\partial^\alpha u(x)$
- In der Physik oder Ingenieurmathematik werden Komponenten eines Vektors in $\mathbb{R}^2$ oder $\mathbb{R}^3$ häufig mit den Buchstaben x, y, z und partielle Ableitungen mit ∂_x , $\partial_y, \partial_{xy}$ usw. bezeichnet. Beispiel:

$$u(x, y, z) = x^3 + y^2 + xz, \quad \partial_x u(x, y, z) = 3x^2 + z, \quad \partial_{xz} = 1.$$

- In der mathematischen Modellierung treten häufig Funktionen der Form $u(t, x)$ mit $t \in \mathbb{R}$ (Zeitvariable) und $x \in \mathbb{R}^3$ (Ortsvariable) auf. Man verwendet dann gerne folgende Bezeichnungen:

$$\partial_t u(t, x), \quad \Delta u(t, x) = \Delta_x u(t, x) = \sum_{i=1}^3 \partial_{x_i}^2 u(t, x).$$

Bemerkung 2.108 (partielle Differentialgleichung). Eine Gleichung, in der eine Funktion (mehrerer Veränderlicher) sowie deren partielle Ableitungen vorkommen heißt partielle Differentialgleichung. Diese werden im 3. Studienjahr besprochen. Prominenten Beispiele für partielle Differentialgleichungen sind

- die Stationäre Wärmeleitungsgleichung

$$-\Delta u(x) = f(x), \quad x \in \Omega \subseteq \mathbb{R}^d.$$

- die Wärmeleitungsgleichung

$$\partial_t u(t, x) - \Delta u(t, x) = f(t, x), \quad t \geq 0, x \in \Omega \subseteq \mathbb{R}^d$$

- die Wellengleichung

$$\partial_t^2 u(t, x) - \Delta u(t, x) = f(t, x), \quad t \geq 0, x \in \Omega \subseteq \mathbb{R}^d$$

Bemerkung 2.109. Ein einfaches Modell für eine (in einer Ebene) schwingende Saite ist durch die Wellengleichung gegeben:

- Wir beschreiben die Deformation einer schwingenden Saite mit Länge ℓ durch ein Skalarfeld $u : \mathbb{R}_{\geq 0} \times I \rightarrow \mathbb{R}$ mit $I := [0, \ell]$.
- Hierbei bezeichnet $u(t, x)$ die Auslenkung Saite im Punkt $x \in I$ zum Zeitpunkt t . Genauer gesagt nimmt man an, dass die Saite im Ruhezustand das Gebiet $I \times \{0\}$ einnimmt. Jedes $x \in I$ steht also für einen Punkt auf der Saite; Zum Zeitpunkt $t \geq 0$ befindet sich dieser Saitenpunkt an der Position $(x, 0) + (x, u(t, x))$; d.h., das Skalarfeld u hat die Bedeutung einer Verschiebung in Normalenrichtung.
- Das Modell besteht nun aus der partiellen Differentialgleichung

$$(2.19a) \quad \partial_{tt} u(t, x) - K \partial_{xx} u(t, x) = 0, \quad t > 0, x \in I$$

den Randbedingungen

$$(2.19b) \quad u(t, 0) = u(t, \ell) = 0, \quad t \geq 0.$$

(die Saite ist eingespannt), und den Anfangswertbedingungen

$$(2.19c) \quad u(0, x) = u_0(x) \text{ und } \partial_t u(0, x) = v_0(x), \quad x \in (0, \ell).$$

Hierbei ist $K > 0$ ein Parameter, der von den Materialeigenschaften (Masse, elastischer Modul,...) abhängig ist, $u_0 : I \rightarrow \mathbb{R}$ und $v_0 : I \rightarrow \mathbb{R}$ beschreiben die Auslenkung und (Normalen)geschwindigkeit der Saite zum Zeitpunkt $t = 0$.

- In der Theorie partieller Differentialgleichungen zeigt man, dass dieses Problem (für hinreichend schöne Daten u_0, v_0) eine eindeutige Lösung besitzt. In der Numerik partieller Differentialgleichungen lernt Verfahren zur Berechnung von Lösungen mit dem Computer.

Ein einfache Möglichkeit eine Lösung zu finden, ist einen geeigneten Ansatz zu betrachten. Sei

$$u(t, x) := \cos(ct) \sin(kx),$$

mit zu bestimmenden Parametern $c > 0$ und $k > 0$.

1. Wann löst u (2.19a)? Es gilt

$$\partial_{tt} u(t, x) = -c^2 u(t, x), \quad \partial_{xx} u(t, x) = -k^2 u(t, x).$$

Durch einsetzen in (2.19a) sehen wir: u löst (2.19a), falls

$$(2.20) \quad c^2 - Kk^2 = 0.$$

2. Wann erfüllt u (2.19b)? (2.19b) ist äquivalent zu

$$\sin(kx) = 0 \quad \text{für } x \in \{0, \ell\}.$$

Dies gilt genau dann, wenn $kx \in \pi\mathbb{N}$. Wir folgern:

$$(2.21) \quad k \in \left\{ n \frac{\pi}{\ell} : n \in \mathbb{N} \right\}.$$

3. Fazit: Für alle $n \in \mathbb{N}$ löst $u_n : \mathbb{R}_{\geq 0} \times [0, \ell] \rightarrow \mathbb{R}$,

$$u_n(t, x) := \cos\left(n \frac{\sqrt{K}\pi}{\ell} t\right) \sin\left(n \frac{\pi}{\ell} x\right)$$

(2.19a) und (2.19b).

Wir können die Funktionen u_n als Film visualisieren, aber auch anhören. Als Töne entspricht die Sequence $u_1, u_2, u_3, \dots$ einer Obertonreihe zum Grundton u_1 . Der Zweiklang u_1, u_2 ist eine Oktave, u_2, u_3 ein Quinte usw.

4. Beobachtung: Durch einsetzen sieht man, dass auch beliebige Linearkombinationen von $\{u_n\}_{n \in \mathbb{N}}$ (2.19a) und (2.19b) lösen. Beispiel:

$$u(t, x) := \sum_{n=1}^N \frac{1}{n} u_n(t, x).$$

Dieses u hat folgende Anfangswerte:

$$\begin{aligned} u(0, x) &:= \sum_{n=1}^N \frac{1}{n} u_n(0, x) = \sum_{n=1}^N \frac{1}{n} \sin(n \frac{\pi}{\ell} x) \\ \partial_t u(0, x) &:= 0 \end{aligned}$$

5. In der Funktionalanalysis zeigt man, dass eine große Klasse von Funktionen u_0 (mit $u_0(0) = u_0(\ell) = 0$) sich als Funktionenreihe (sogenannte Fourier-Reihe) der Form $u_0(x) = \sum_{n=1}^{\infty} a_n \sin(n \frac{\pi}{\ell} x)$ darstellen lassen. In diesem Fall ist eine Lösung zu (2.19a) - (2.19c) (mit $v_0 = 0$) durch die Funktion $u = \sum_{n=1}^{\infty} a_n u_n$ gegeben.

Kontrollfrage 9.5. Wir betrachten die Wellengleichung

$$(2.22) \quad \partial_{tt} u(t, x) - 2 \partial_{xx} u(t, x) = 0 \quad \text{für } (t, x) \in \mathbb{R} \times \mathbb{R}.$$

Welche der nachfolgenden Aussagen sind richtig?

- (A) Sei $g \in C^2(\mathbb{R})$ und $u(t, x) := g(x/\sqrt{2} + t)$. Dann löst u die Wellengleichung (2.22).

$$\checkmark: \partial_{tt} u(t, x) = g''(x/\sqrt{2} + t) \text{ und } \partial_{xx} u(t, x) = (1/\sqrt{2})^2 g''(x/\sqrt{2} + t).$$

- (B) Sei $g \in C^2(\mathbb{R})$ und $u(t, x) := g(x + \sqrt{2}t)$. Dann löst u die Wellengleichung (2.22).

$$\checkmark: \text{Zurückführbar auf A mit } u(t, x) = g(\sqrt{2}(x/\sqrt{2} + t)).$$

- (C) Sei $g \in C^2(\mathbb{R})$ und $u(t, x) := g(x - \sqrt{2}t)$. Dann löst u die Wellengleichung (2.22).

$$\checkmark: \partial_{tt} u(t, x) = (-\sqrt{2})^2 g''(x - \sqrt{2}t) \text{ und } \partial_{xx} u(t, x) = g''(x - \sqrt{2}t).$$

- (D) Die Abbildung $u : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, $u(t, x) := \cos(\sqrt{2}t) \sin(x)$ löst (2.22).

$$\checkmark: \text{Nachrechnen.}$$

Kontrollfrage 9.6. Betrachte

$$u : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}, \quad u(x) := \ln(\|x\|_2).$$

Welche der nachfolgenden Aussagen ist richtig?

- (A) u ist harmonisch in $\mathbb{R}^2 \setminus \{0\}$.

$$\checkmark: \Delta u = 0 \text{ nachrechnen.}$$

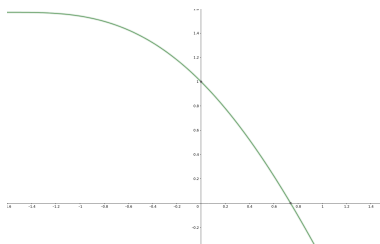
- (B) u ist nicht harmonisch in $\mathbb{R}^2 \setminus \{0\}$.

×

3 Banachscher Fixpunktsatz und Anwendungen

Wir betrachten in diesem Abschnitt eine Methode um Lösungen für nichtlineare Gleichungen zu berechnen. Zur Motivation betrachten wir das Nullstellenproblem für die Funktion

$$g : \mathbb{R} \rightarrow \mathbb{R}, \quad g(x) = \cos(x) - x.$$



Wegen $g(0) = 1$ und $g(\pi/2) = -\pi/2$ können wir die Existenz einer Nullstelle im Intervall $(0, \pi/2)$ aus dem Zwischenwertsatz folgern. Aber wie kann man die Nullstelle berechnen?

Mit der Funktion

$$f : [0, \pi/2] \rightarrow [0, \pi/2], \quad f(x) = \cos(x),$$

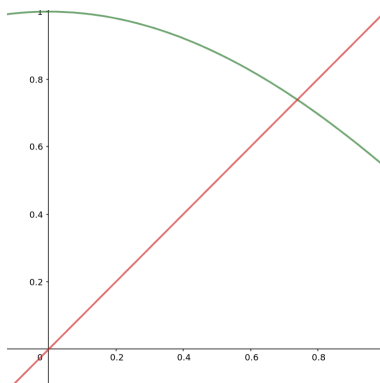
lässt sich das Problem als Fixpunktproblem umformulieren, d.h. wir suchen ein $x \in [0, \pi/2]$ mit $f(x) = x$.

Wir definieren nun iterativ eine Folge durch

$$x_0 \in [0, \frac{\pi}{2}], \quad x_k = f(x_{k-1}).$$

Es stellt sich heraus, dass diese Folge gegen einen Grenzwert y konvergiert. Dieser muss dann ein Fixpunkt von f sein, denn wegen der Stetigkeit von f gilt

$$y = \lim_{k \rightarrow \infty} x_k = \lim_{k \rightarrow \infty} f(x_{k-1}) = f(\lim_{k \rightarrow \infty} x_{k-1}) = f(y).$$



Satz 3.1 (Banachscher Fixpunktsatz). *Sei (M, d) ein vollständiger metrischer, nichtleerer Raum und sein $f : M \rightarrow M$ eine strikte Kontraktion, d.h. es existiert $L < 1$ mit*

$$d(f(x), f(y)) \leq Ld(x, y) \quad \text{für alle } x, y \in M.$$

Dann besitzt f genau einen Fixpunkt.

Beweis. • Wähle $x_0 \in M$.

- Fixpunktiteration

$$x_k := f(x_{k-1}) \quad k \in \mathbb{N}_+.$$

Es gilt für alle $k \in \mathbb{N}$:

$$\begin{aligned} d(x_{k+1}, x_k) &= d(f(x_k), f(x_{k-1})) \leq Ld(x_k, x_{k-1}) \\ &\leq L^k d(x_1, x_0). \end{aligned}$$

Es folgt (Teleskopsumme)

$$\begin{aligned} d(x_m, x_n) &\leq \sum_{k=n}^{m-1} d(x_{k+1}, x_k) \leq \left(\sum_{k=n}^{m-1} L^k \right) d(x_1, x_0) \\ &= \frac{L^n}{1-L} d(x_1, x_0) \rightarrow 0 \quad (m, n \rightarrow \infty). \end{aligned}$$

- Damit ist (x_k) eine Cauchy-Folge und da M vollständig ist, existiert $x = \lim_{k \rightarrow \infty} x_k$. Aus der Stetigkeit von f folgt nun

$$d(f(x), x) = \lim_{k \rightarrow \infty} d(f(x_k), x_k) = \lim_{k \rightarrow \infty} d(x_{k+1}, x_k) = 0.$$

□

Beispiel 3.2. Wir betrachten $M = [0, 1]$ mit der Abstandsmetrik.

(a) Die Funktion $M \rightarrow M$, $x \mapsto \cos(x)$ ist eine strikte Kontraktion: Für alle $x, y \in M$ gilt

$$\exists \xi \in M : \quad |\cos(x) - \cos(y)| = |\sin(\xi)| |x - y| = |\sin(\xi)| |x - y| \leq \sin(1) |x - y|$$

für alle $x, y \in M$. Der Banachsche Fixpunktsatz ist also anwendbar.

(b) Die Funktion $M \rightarrow M$, $x \mapsto \sin(x)$ ist keine strikte Kontraktion, denn es gilt

$$\lim_{\substack{h \rightarrow 0 \\ h > 0}} \frac{|\sin(h) - \sin(0)|}{|h - 0|} = \cos(0) = 1,$$

also kann kein $L < 1$ existieren mit $|\sin(x) - \sin(y)| \leq L|x - y|$ für alle $x, y \in M$.

Beispiel 3.3. Wir betrachten $M = \mathbb{R}^d$ mit dem euklidischen Abstand. Sei

$$f(x) = x + \omega Ax - \omega c$$

mit $A \in \mathbb{R}^{d \times d}$, $\omega \neq 0$ und $c \in \mathbb{R}^d$. Ein Punkt x ist ein Fixpunkt von f genau dann, wenn

$$Ax = c.$$

Es gilt

$$\|f(x) - f(y)\|_2 \leq \|I + \omega A\| \|x - y\|_2,$$

wobei $\|F\| := \sup_{\|x\|_2=1} \|Fx\|_2$ die natürliche Matrixnorm bezeichnet.

In der linearen Algebra zeigt man: $\|F\|^2 = \mu_{\max}$, wobei $\mu_{\max}$ den maximalen Singulärwert von F bezeichnet (=Eigenwert von $F^\top F$).

Im Fall $\|I + \omega A\| < 1$ folgt aus dem Banachschen Fixpunktsatz, dass die Folge

$$x_0 = 0, \quad x_n = f(x_{n-1}) = x_{n-1} + \omega Ax_{n-1} - \omega c,$$

gegen einen Fixpunkt von f und damit gegen eine Lösung der Gleichung $Ax = c$ konvergiert. Bei diesem Vorgehen, muss die Matrix A nicht invertiert werden!

Beispiele:

- Sei A eine Diagonalmatrix mit Diagonaleinträgen $a_1, \dots, a_n \in \mathbb{R}$. Sind alle Einträge echt positiv, dann existiert $\omega < 0$, so dass $\|I + \omega A\| < 1$. Sind alle Einträge von A echt negativ, dann existiert $\omega > 0$, so dass $\|I + \omega A\| < 1$. Andernfalls gilt $\|I + \omega A\| \geq 1$ für alle $\omega \neq 0$.
- Sei A symmetrisch. In der lineare Algebra zeigt man, dass $Q \in O(d)$ und eine Diagonalmatrix D (deren Einträge die Eigenwerte von A sind) existieren, so dass $A = Q^\top D Q$. Nachrechnen zeigt nun

$$\|I + \omega A\| = \|Q^\top (I + \omega D) Q\| = \|I + \omega D\|.$$

Wir folgern: Sind alle Eigenwerte von A echt positiv oder echt negativ, dann existiert $\omega \neq 0$, so dass $\|I + \omega A\| < 1$.

Kontrollfrage 9.7. Sei $I := [0, \pi]$, $f : I \rightarrow I$, $f(x) := \sin(x)$. Welche der nachfolgenden Aussagen sind richtig?

(A) f hat einen Fixpunkt, weil f eine strikte Kontraktion ist.

×: Fixpunkt ja (0), strikte Kontraktion nein ($f'(0) = 1$).

(B) f hat einen Fixpunkt, obwohl f keine strikte Kontraktion ist.

✓: Siehe A.

(C) f hat keinen Fixpunkt, weil f keine strikte Kontraktion ist.

×: Siehe A.

(D) f hat keinen Fixpunkt, obwohl f keine strikte Kontraktion ist.

×: Siehe A.

Kontrollfrage 9.8. Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $n \geq 2$ eine strikte Kontraktion auf $(\mathbb{R}^n, \|\cdot\|_2)$. Welche der nachfolgenden Aussagen ist dann richtig?

(A) Sei $\|\cdot\|$ eine beliebige Norm auf $\mathbb{R}^n$. Dann ist f eine strikte Kontraktion auf $(\mathbb{R}^n, \|\cdot\|)$

×: Siehe B.

(B) f ist eine strikte Kontraktion auf $(\mathbb{R}^n, \|\cdot\|_\infty)$, da $\|\cdot\|_2$ stärker als $\|\cdot\|_\infty$ ist.

$\times$: Eine Idee fürs Gegenbeispiel lässt sich u.U. finden, wenn man Abbildung 1 im Kopf hat. Es funktioniert z.B.

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^2 : x \mapsto \frac{99}{100} \|x\|_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Denn dann ist f eine strikte Kontraktion bzgl. $\|\cdot\|_2$, aber

$$\|f \begin{pmatrix} 1 \\ 1 \end{pmatrix} - f(0)\|_\infty = \left\| \frac{99}{100} \sqrt{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\|_\infty = \frac{99}{100} \sqrt{2} > 1 = \left\| \begin{pmatrix} 1 \\ 1 \end{pmatrix} - 0 \right\|_\infty.$$

(C) f ist niemals eine strikte Kontraktion auf $(\mathbb{R}^n, \|\cdot\|_\infty)$.

$\times$: Gegenbeispiel: $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2 : x \mapsto 0$.

(D) Keine der Aussagen ist richtig.

✓: (Nur diese hier.)

8.6.2020

#31

15.6.2020

3.1 Diffeomorphismen

Seien U, V Mengen und $\Phi : U \rightarrow V$ eine Bijektion. Es bezeichne $\Psi : V \rightarrow U$ die Umkehrabbildung von Φ .

Ist $f : V \rightarrow \mathbb{R}$ gegeben, so wird durch $\tilde{f} := f \circ \Phi$ eine Abbildung von U nach $\mathbb{R}$ definiert. Die Umkehrung gilt auch: Ist $\tilde{f} : U \rightarrow \mathbb{R}$ gegeben, dann wird durch $f := \tilde{f} \circ \Psi$ eine Abbildung von V nach $\mathbb{R}$ definiert.

Seien nun U, V Teilmengen metrischer Räume. Wir möchten, dass für alle $f : V \rightarrow \mathbb{R}$ und $\tilde{f} : U \rightarrow \mathbb{R}$ mit $\tilde{f} := f \circ \Phi$ die Äquivalenz

$$f \text{ stetig} \quad \Leftrightarrow \quad \tilde{f} \text{ stetig.}$$

Dies ist genau dann erfüllt, wenn Φ und Ψ stetig sind, d.h. wenn Φ ein Homöomorphismus ist.

Seien nun U, V offene Teilmengen des $\mathbb{R}^n$. Wir möchten, dass für alle $f : V \rightarrow \mathbb{R}$ und $\tilde{f} : U \rightarrow \mathbb{R}$ mit $\tilde{f} := f \circ \Phi$ die Äquivalenz

$$f \text{ stetig differenzierbar} \quad \Leftrightarrow \quad \tilde{f} \text{ stetig differenzierbar,}$$

gilt. Dies ist genau dann erfüllt, wenn Φ ein Homöomorphismus ist und sowohl Φ als auch Ψ stetig differenzierbar sind.

Definition 3.4 (Diffeomorphismus). Seien $U, V \subseteq \mathbb{R}^n$ offen. Eine Abbildung $\Phi : U \rightarrow V$ heißt Diffeomorphismus (oder Koordinatentransformation) falls Φ bijektiv ist und sowohl Φ als auch die Umkehrfunktion Φ^{-1} stetig differenzierbar sind.

Kontrollfrage 10.1. Seien $U, V \subseteq \mathbb{R}^n$ offen. Sei $\Phi : U \rightarrow V$ ein Diffeomorphismus und $f, g \in C^\infty(\mathbb{R}^n)$. Welche der nachfolgenden Aussagen sind richtig?

(A) Durch $u : U \rightarrow \mathbb{R}$, $u(x) := f(\Phi(x))$ wird eine Funktion in $C^1(V)$ definiert.

$\times$: Mit $C^1(U)$ statt $C^1(V)$ wäre es richtig.

(B) Durch $u : U \rightarrow \mathbb{R}$, $u(x) := f(\Phi(x))$ wird eine Funktion in $C^2(U)$ definiert.

$\times$: Kann i.A. nur C^1 erwarten.

(C) Sei $u : U \rightarrow \mathbb{R}$, $u(x) := f(\Phi(x))$ und $v : U \rightarrow \mathbb{R}$, $v(x) := g(\Phi(x))$. Dann gilt $u = v$ in U genau dann, wenn $g = f$ in V .

✓: Nutzt Bijektivität von Φ .

(D) Φ ist ein Homöomorphismus.

✓: Folgt aus Definition 3.4 und daraus, dass Stetigkeit durch Differenzierbarkeit impliziert wird.

Beispiele 3.5. (a) Die Identität auf $\mathbb{R}^n$,

$$\text{id} : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad x \mapsto x$$

ist ein Diffeomorphismus und es gilt $\text{Did}(x) = \text{id}$ für alle $x \in \mathbb{R}^n$.

(b) Umsortieren der Koordinaten des $\mathbb{R}^n$.

Sei $\pi : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ eine Permutation. Dann ist

$$\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad \Phi(x) := (x_{\pi(1)}, \dots, x_{\pi(n)}),$$

eine Koordinatentransformation.

(c) Affine Koordinatentransformation.

Sei $A \in \mathbb{R}^{n \times n}$ eine invertierbare Matrix, $c \in \mathbb{R}^n$. Dann ist

$$\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad \Phi(x) := Ax + c$$

eine affine Koordinatentransformation. Es gilt

$$D\Phi(x)(h) = Ah, \quad \Phi^{-1}(y) = A^{-1}(y - c), \quad D\Phi^{-1}(y)(h) = A^{-1}h.$$

Man kann zeigen: Eine stetig differenzierbare Abbildung $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ ist genau dann eine affine Koordinatentransformation, wenn $D\Phi(x)$ unabhängig von $x \in \mathbb{R}^d$ und invertierbar ist.

(d) Polarkoordinaten in der Ebene.

Wir betrachten die Abbildung

$$\Phi : (0, \infty) \times (0, 2\pi) \rightarrow \mathbb{R}^2 \setminus ([0, \infty) \times \{0\}), \quad \Phi(r, \alpha) := (r \cos(\alpha), r \sin(\alpha))$$

(r, α) heißen Polarkoordinaten des Punktes $x = \Phi(r, \alpha)$. Eine direkte Rechnung zeigt: $x = \Phi(r, \alpha) \Rightarrow r = \|x\|_2$.

Die Bijektivität folgt aus der geometrischen Konstruktion. Die Differenzierbarkeit der Umkehrabbildung diskutieren wir gleich.

(e) Kugelkoordinaten im $\mathbb{R}^3$.

Wir betrachten die Abbildung

$$\begin{aligned}\Phi : (0, \infty) \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \times (0, 2\pi) &\rightarrow \mathbb{R}^3 \setminus ([0, \infty) \times \{0\} \times \mathbb{R}), \\ \Phi(r, \alpha, \beta) &:= (r \cos(\alpha) \cos(\beta), r \cos(\alpha) \sin(\beta), r \sin(\alpha))\end{aligned}$$

(r, α, β) heißen Kugelkoordinaten des Punktes $x = \Phi(r, \alpha, \beta)$. Eine direkte Rechnung zeigt: $x = \Phi(r, \alpha, \beta) \Rightarrow r = \|x\|_2$.

Die Bijektivität folgt aus der geometrischen Konstruktion. Die Differenzierbarkeit der Umkehrabbildung diskutieren wir gleich.

(f) Zylinderkoordinaten im $\mathbb{R}^3$. Wir betrachten die Abbildung

$$\begin{aligned}\Phi : (0, \infty) \times (0, 2\pi) \times \mathbb{R} &\rightarrow (\mathbb{R}^2 \setminus ([0, \infty) \times \{0\})) \times \mathbb{R}, \\ \Phi(r, \alpha, z) &:= (r \cos(\alpha), r \sin(\alpha), z)\end{aligned}$$

(r, α, z) heißen Zylinderkoordinaten des Punktes $x = \Phi(r, \alpha, z)$.

Die Bijektivität folgt aus der geometrischen Konstruktion. Die Differenzierbarkeit der Umkehrabbildung diskutieren wir gleich.

Kontrollfrage 10.2. Es bezeichne Φ die Koordinatentransformation aus Beispiel 3.5 (f) (Zylinderkoordinaten). Welche der nachfolgenden Aussagen sind wahr?

(A) Es gilt $\{\Phi(r, \alpha, z) : r > 0, 0 < \alpha < 2\pi, z \in \mathbb{R}\} = \mathbb{R}^3$

$\times : [0, \infty) \times \{0\} \times \mathbb{R}$ fehlt.

(B) Der Abschluss von $\{\Phi(r, \alpha, z) : r > 0, 0 < \alpha < 2\pi, z \in \mathbb{R}\}$ ist gleich $\mathbb{R}^3$

✓: Siehe A.

(C) Die Menge $\{\Phi(1, \alpha, 0) : 0 < \alpha < 2\pi\}$ liegt im Schnitt der Einheitssphäre im $\mathbb{R}^3$ mit der x_1 - x_2 -Ebene.

✓

(D) Die Menge $\{\Phi(r, 0, 0) : r > 0\}$ ist eine Gerade.

$\times$: Das Objekt ist gar nicht wohldefiniert, da „ $(r, 0, 0)$ “ nicht im Definitionsbereich wäre. Aber selbst, wenn man die Rechenvorschrift intuitiv ausweiten würde, wäre das Objekt nur ein Strahl und keine vollständige Gerade.

Kontrollfrage 10.3. Sei

$$U := (0, \infty) \times [0, 2\pi], \quad \Phi : U \rightarrow \mathbb{R}^2, \quad \Phi(r, \alpha) := r(\cos \alpha, \sin \alpha)$$

Welche der nachfolgenden Aussagen sind wahr?

(A) $\Phi(U) = \mathbb{R}^2$.

$\times$: Die Null fehlt.

(B) $\Phi : U \rightarrow \mathbb{R}^2 \setminus \{0\}$ ist surjektiv

✓

(C) $\Phi : U \rightarrow \mathbb{R}^2 \setminus \{0\}$ ist injektiv

$\times$: Beachte zum Beispiel, dass $U(1, 0) = U(1, 2\pi)$.

Beispiel 3.6. Häufig ist es einfacher Abbildungen in geeigneten “Koordinaten” anzugeben. Beispiele:

(a) Das Gravitationspotential eines radialsymmetrischen Körpers ist (bis auf Konstanten) durch das Skalarfeld

$$(3.1) \quad \tilde{u}(r, \alpha, \beta) := -\frac{1}{r} \quad \text{in Kugelkoordinaten}$$

gegeben.

In euklidischen Koordinaten erhalten wir das Skalarfeld

$$u(x) = \tilde{u}(\Phi^{-1}(x)) \quad \text{für alle } x \in \text{Bild}(\Phi),$$

wobei Φ die Koordinatentransformation auf die Kugelkoordinaten bezeichnet.

(b) Wir betrachten das Skalarfeld

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = \|(x, y)\|_2^{\frac{1}{2}} \left(1 + \frac{1}{2} \cos(10 \cos^{-1}(\frac{x}{\|(x, y)\|_2})) \right).$$

In Polarkoordinaten gilt

$$\tilde{f}(r, \alpha) = \sqrt{r}(1 + \cos(10\alpha)) = f(\Phi(r, \alpha))$$

Es ist einfacher die partiellen Ableitungen von $\tilde{f}$ als von f zu berechnen.

Mit der Kettenregel gilt aber mit $x = \Phi(r, \alpha)$

$$\nabla \tilde{f}(r, \alpha) = \begin{pmatrix} \partial_r \tilde{f}(r, \alpha) \\ \partial_\alpha \tilde{f}(r, \alpha) \end{pmatrix} = J_\Phi(r, \alpha) \nabla f(x)$$

$$\nabla f(x) = J_\Psi(x) \nabla \tilde{f}(r, \alpha),$$

wobei $\Psi := \Phi^{-1}$ die Umkehrabbildung bezeichnet.

Um die totale Ableitung der Umkehrabbildungen der Polar- und Kugelkoordinaten zu bestimmen, müssen wir die Umkehrabbildung nicht explizit kennen:

Lemma 3.7. Sei $\Phi : U \rightarrow V$ ein differenzierbarer Homöomorphismus zwischen offenen Mengen $U, V \subseteq \mathbb{R}^n$. Die Umkehrabbildung $\Psi : V \rightarrow U$ sei im Punkt $y \in V$ differenzierbar. Dann gilt

$$(3.2) \quad D\Psi(y) = (D\Phi(x))^{-1}, \quad J_\Psi(y) = (J_\Phi(x))^{-1}, \quad \text{mit } x := \Psi(y).$$

Beweis. Aus $\Psi \circ \Phi = \text{id}$ und der Kettenregel (Proposition 2.44) folgt mit den linearen Abbildungen $T = D\Phi(x)$ und $S = D\Psi(\Phi(x))$:

$$\begin{aligned}\text{id} &= D(\Psi \circ \Phi)(x) = D\Psi(\Phi(x)) \circ D\Phi(x) = S \circ T \\ \text{id} &= D(\Phi \circ \Psi)(\Phi(x)) = D\Phi(x) \circ D\Psi(\Phi(x)) = T \circ S,\end{aligned}$$

also $S = T^{-1}$. Die Identität für die Jacobimatrizen folgt durch Übergang von lineare Abbildungen zu Darstellungsmatrizen. $\square$

Kontrollfrage 10.4. Welche der nachfolgenden Aussagen ist wahr? Die Abbildung $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $f(x, y) = (x^3, y^3)$ ist kein Diffeomorphismus, denn

(A) sie ist nicht mal bijektiv.

$\times$: Doch, ist sie.

(B) sie ist nicht mal ein Homöomorphismus.

$\times$: Doch, die Umkehrabbildung ist auch stetig.

(C) ihre Umkehrabbildung ist nicht differenzierbar.

$\checkmark$: Sie ist in $(0, 0)$ nicht diffbar.

(D) ihre Jacobi-Matrix im Punkt $(0, 0)$ ist nicht invertierbar.

$\checkmark$: Für Diffeomorphismen muss die Jakobimatrix laut Lemma 3.7 überall invertierbar sein.

Der folgende Satz gibt ein Kriterium, wann stetig differenzierbare Homöomorphismen bereits Diffeomorphismen sind.

Satz 3.8. Sei $\Phi : U \rightarrow V$ ein Homöomorphismus zwischen offenen Mengen $U, V \subseteq \mathbb{R}^n$ (d.h. Φ ist bijektiv und Φ, Φ^{-1} sind stetig). Sei Φ in U stetig differenzierbar und für alle $x \in U$ sei $J_\Phi(x)$ invertierbar. Dann ist auch die Umkehrabbildung $\Psi := \Phi^{-1} : V \rightarrow U$ stetig differenzierbar.

Beweis. 1. Transformation auf eine einfachere Situation: Sei $x_0 \in U$ und $y_0 \in V$. Es gelte $y_0 = \Phi(x_0)$. Setze

$$\tilde{U} := U - x_0, \quad \tilde{V} := J_\Phi(x_0)^{-1}(V - y_0).$$

Dann wird durch

$$\tilde{\Phi}(x) := J_\Phi(x_0)^{-1}(\Phi(x + x_0) - y_0)$$

ein stetig differenzierbarer Homöomorphismus $\tilde{\Phi} : \tilde{U} \rightarrow \tilde{V}$ definiert und es gilt

$$\tilde{\Phi}(0) = 0, \quad D\tilde{\Phi}(0) = \text{id}.$$

Nach Konstruktion gilt weiter: Die Umkehrabbildung $\tilde{\Psi} := \tilde{\Phi}^{-1} : \tilde{V} \rightarrow \tilde{U}$ ist genau dann (stetig) differenzierbar in 0, wenn Ψ (stetig) differenzierbar in y_0 ist.

2. Behauptung: Ψ ist differenzierbar in V . Wir müssen Differenzierbarkeit von Ψ in einem Punkt $y \in V$ nachweisen. Nach Schritt 1 dürfen o.B.d.A. annehmen, dass $0 \in U$, $0 \in V$ und

$$(3.3) \quad y = 0, \quad \Phi(0) = 0, \quad D\Phi(0) = \text{id}.$$

Da Φ differenzierbar in 0 ist folgt mit der Weierstraßschen Zerlegungsformel,

$$\forall h \in U : \Phi(h) - (\Phi(0) + D\phi(0)(h)) = \|h\|\rho(h)$$

mit $\rho(h) \rightarrow 0$ für $h \rightarrow 0$. Mit (3.3) folgt

$$(3.4) \quad \Phi(h) = h + \|h\|\rho(h).$$

Wir betrachten nun die Weierstraßsche Zerlegungsformel für Ψ : Sei $k \in V$. Setze $h := \Psi(k)$. Dann gilt

$$\Psi(k) = h = \Phi(h) - \underbrace{\|h\|\rho(h)}_{=\|\Psi(k)\|\rho(\Psi(k))} = k + \|k\|\tilde{\rho}(k),$$

mit

$$\tilde{\rho}(k) := \begin{cases} \frac{\|\psi(k)\|}{\|k\|} \rho(\psi(k)) & k \neq 0 \\ 0 & \text{sonst.} \end{cases}$$

Wir zeigen, dass $\tilde{\rho}(k) \rightarrow 0$ für $k \rightarrow 0$ (denn dann ist Ψ differenzierbar im Punkt 0). Sei $(k_j)_{j \in \mathbb{N}}$ eine Folge in V mit $k_j \rightarrow 0$ und $k_j \neq 0$. Es gilt:

- (i) $h_j := \Psi(k_j) \rightarrow \Psi(0) = 0$, da Ψ stetig ist.
- (ii) $\rho(\Psi(k_j)) = \rho(h_j) \rightarrow 0$.
- (iii) Durch $c_j := \frac{\|\Psi(k_j)\|}{\|k_j\|}$ wird eine beschränkte Folge definiert. Argument: Wähle j_0 , sodass

$$\|\rho(h_j)\| \leq \frac{1}{2} \quad \text{für alle } j \geq j_0.$$

Das ist möglich wegen (i). Für $j \geq j_0$ gilt nun wegen (3.4),

$$\|\Psi(k_j)\| = \|h_j\| = \left\| \underbrace{\Phi(h_j)}_{=k_j} - \underbrace{\|h_j\|\rho(h_j)}_{=\Psi(k_j)} \right\| \leq \|k_j\| + \frac{1}{2}\|\Psi(k_j)\|$$

und damit

$$\|\Psi(k_j)\| \leq 2\|k_j\|.$$

Aus (ii) und (iii) folgt nun $\tilde{\rho}(k_j) \rightarrow 0$.

3. Behauptung: Ψ ist stetig differenzierbar. Wir setzen

$$\begin{aligned} T : U &\rightarrow \text{GL}(\mathfrak{n}), & T(x) &:= J_\Phi(x), \\ S : V &\rightarrow \mathbb{R}^{n \times n}, & S(y) &:= J_\Psi(y), \end{aligned}$$

wobei $GL(n) \subseteq \mathbb{R}^{n \times n}$ die Menge aller invertierbaren $n \times n$ -Matrizen bezeichnet. Es genügt zu zeigen, dass S stetig ist. Nach Lemma 3.7 gilt für alle $y \in V$

$$J_{\Psi}(y) = \left(J_{\Phi}(\Psi(y)) \right)^{-1},$$

und damit

$$S = f \circ T \circ \Psi$$

wobei $f : GL(n) \rightarrow GL(n)$ die Matrixinversion $f(A) := A^{-1}$ bezeichnet. Die Abbildung f ist stetig. Damit ist S die Verknüpfung dreier stetiger Abbildungen und damit selber stetig. $\square$

Kontrollfrage 10.5. Seien $f, g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ Diffeomorphismen. Welche der nachfolgenden Aussagen sind dann wahr?

- $f \circ g$ ist ein Diffeomorphismus, nicht aber $f^{-1} \circ g^{-1}$.

$\times$: Siehe B.

- $g \circ f$ und $f^{-1} \circ g^{-1}$ sind Diffeomorphismen

✓: Die erste Behauptung lässt sich unter anderem mit der Kettenregel zeigen. Die zweite folgt dann direkt, weil die Umkehrabbildungen von Diffeomorphismen selbst Diffeomorphismen sind.

- Weder $g \circ f$ noch $f^{-1} \circ g^{-1}$ müssen Diffeomorphismen sein.

$\times$: Siehe B.

Bemerkung 3.9. Im eindimensionalen Fall gilt noch mehr: Ist $U \subseteq \mathbb{R}$ ein offenes Intervall und $\Phi : U \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion mit $\Phi'(x) \neq 0$ für alle $x \in U$, dann ist Φ bereits streng monoton und es folgt, dass $\Phi : U \rightarrow \Phi(U)$ bijektiv und die Umkehrfunktion Φ^{-1} stetig differenzierbar ist. Damit ist $\Phi : U \rightarrow \Phi(U)$ ein Diffeomorphismus.

Im $\mathbb{R}^n$ mit $n > 1$ gilt diese Aussage nicht mehr. Wie wir in der Übung zeigen, muss eine stetig differenzierbare Abbildung $\Phi : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$, deren Jacobi-Matrix in jedem Punkt invertierbar ist, nicht mal injektiv sein. Wir werden aber sehen, dass solche Abbildungen zumindest lokal invertierbar sind.

Definition 3.10 (lokaler Diffeomorphismus). Sei $U \subseteq \mathbb{R}^n$. Eine Abbildung $\Phi : U \rightarrow \mathbb{R}^n$ heißt lokaler Diffeomorphismus um $p \in U$, wenn es offene Umgebungen $\tilde{U} \subseteq U$ von p und $\tilde{V} \subseteq \mathbb{R}^n$ von $\Phi(p)$ gibt, sodass $\Phi|_{\tilde{U}} : \tilde{U} \rightarrow \tilde{V}$ ein Diffeomorphismus ist.

Satz 3.11 (Satz über den lokalen Diffeomorphismus). Sei $U \subseteq \mathbb{R}^n$ offen und $p \in U$. Sei $\Phi : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig differenzierbar in U und $J_{\Phi}(p)$ invertierbar. Dann ist Φ ein lokaler Diffeomorphismus um p .

Beweis. Mit dem Transformationsargument aus dem Beweis von Satz 3.8 (Schritt 1) genügt es den Spezialfall

$$p = 0, \quad \Phi(p) = 0, \quad D\Phi(0) = \text{id}$$

zu betrachten. Nach Satz 3.8 genügt es nun zu zeigen, dass offene Umgebungen $\tilde{U}, \tilde{V} \subseteq \mathbb{R}^n$ von 0 existieren, sodass:

(A1) $\Phi|_{\tilde{U}} : \tilde{U} \rightarrow \tilde{V}$ ist ein Homöomorphismus,

(A2) $J_\Phi(x)$ ist invertierbar für alle $x \in \tilde{U}$.

1. Behauptung:

$$(\star)_1 \quad \text{Es existiert } r > 0, \text{ so dass } K(0; r) \subseteq U \text{ und} \\ \|D\Phi(x) - \text{id}\| + |\det J_\Phi(x) - 1| \leq \frac{1}{2} \text{ für alle } x \in K(0; r).$$

Argument: Da U offen ist, gilt immer $K(0; r) \subseteq U$, falls $r > 0$ nur klein genug ist. Weiter ist die Funktion $f : U \rightarrow \mathbb{R}$, $f(x) := \|D\Phi(x) - \text{id}\| + |\det J_\Phi(x) - 1|$ stetig und es gilt nach Voraussetzung $f(0) = 0$. Daher existiert (ε - δ -Kriterium) ein $r > 0$ mit $f(x) \leq \frac{1}{2}$ für alle $x \in K(0; r)$.

Von nun an sei $r > 0$ wie in Schritt 1 gewählt.

2. Behauptung:

$$(\star)_2 \quad \text{Für alle } y \in B(0; r/2) \text{ existiert genau ein } x \in B(0; r) \text{ mit } \Phi(x) = y.$$

Die Begründung basiert auf dem Banachschem Fixpunktsatz und ist das Kernargument des Beweises.

Argument: Für $y \in K(0; r/2)$ und $x \in U$ setze

$$\varphi_y(x) := y + x - \Phi(x).$$

Dann gilt $\Phi(x) = y$ genau dann, wenn x ein Fixpunkt von φ_y ist. Es gilt $D\varphi_y(x) = \text{id} - D\Phi(x)$. Mit $(\star)_1$ folgt

$$\forall x \in K(0; r) : \|D\varphi_y(x)\| \leq \frac{1}{2}$$

Aus dem Schrankensatz folgt nun,

$$(3.5) \quad \|\varphi_y(x_2) - \varphi_y(x_1)\| \leq \frac{1}{2} \|x_2 - x_1\| \quad \text{für alle } x_1, x_2 \in K(0; r).$$

Für $x \in K(0; r)$ gilt $\|\varphi_y(x)\| \leq \|\varphi_y(x) - \varphi_y(0)\| + \|y\| \leq \frac{r}{2} + \|y\|$, also

$$\begin{aligned} \varphi_y(K(0; r)) &\subseteq K(0; r) & \text{für } y \in K(0; \tfrac{r}{2}) \\ \varphi_y(B(0; r)) &\subseteq B(0; r) & \text{für } y \in B(0; \tfrac{r}{2}). \end{aligned}$$

Also ist für $y \in K(0; \frac{r}{2})$ die Abbildung $\varphi_y : K(0; r) \rightarrow K(0; r)$ eine strikte Kontraktion. Aus dem Banachschem Fixpunktsatz folgt die Existenz eines Fixpunktes und damit $(\star)_2$.

3. Setze $\tilde{V} := B(0; r/2)$, $\tilde{U} := B(0; r) \cap \Phi^{-1}(\tilde{V})$. Dann sind $\tilde{U}, \tilde{V}$ offene Umgebungen der 0 und $\Phi|_{\tilde{U}} : \tilde{U} \rightarrow \tilde{V}$ ist wegen $(\star)_2$ bijektiv. Wir müssen nur noch nachfolgende Behauptung zeigen.

Behauptung: Die Umkehrabbildung $\Psi := \Phi^{-1} : \tilde{V} \rightarrow \tilde{U}$ ist stetig.

Argument: Seien $y_1, y_2 \in \tilde{V}$. Setze $x_i = \Psi(y_i)$. Dann folgt aus (3.5) (mit $y = 0$),

$$\begin{aligned} \|\Psi(y_2) - \Psi(y_1)\| &= \|x_2 - x_1\| \leq \|\varphi_0(x_2) - \varphi_0(x_1)\| + \|\Phi(x_2) - \Phi(x_1)\| \\ &\leq \frac{1}{2} \|x_2 - x_1\| + \|y_2 - y_1\|. \end{aligned}$$

Wir subtrahieren auf beiden Seiten $\frac{1}{2}\|\Psi(y_2) - \Psi(y_1)\| = \frac{1}{2}\|x_2 - x_1\|$ und erhalten

$$(3.6) \quad \|\Psi(y_2) - \Psi(y_1)\| \leq 2\|y_2 - y_1\|.$$

Damit ist Ψ Lipschitzstetig.

4. Konklusion: $\Phi|_{\tilde{U}} : \tilde{U} \rightarrow \tilde{V}$ erfüllt (A1) (vgl. Schritt 3). Es gilt auch (A2), denn $(\star)_1$ impliziert, dass $\det J_\Phi(x) \geq \frac{1}{2}$. Damit ist für alle $x \in \tilde{U}$ die Matrix $J_\Phi(x)$ invertierbar. Damit ist (da $\mathbb{R}^n$ endliche Dimension hat) $D\Phi(x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ein Isomorphismus.

□

Satz 3.12 (Globaler Umkehrsatz). *Sei $U \subseteq \mathbb{R}^n$ offen. Sei $\Phi : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig differenzierbar. Für alle $x \in U$ sei $J_\Phi(x)$ invertierbar. Dann gilt:*

(a) (Offenheitssatz). $\text{Bild}(\Phi) \subseteq \mathbb{R}^n$ ist offen.

(b) Ist Φ zusätzlich injektiv, so ist $\Phi : U \rightarrow \text{Bild}(\Phi)$ ein Diffeomorphismus.

Beweis. (a) Nach Satz 3.11 besitzt jeder Punkt $x \in U$ eine offene Umgebung U_x mit $\Phi(U_x)$ offen. Also ist auch $\text{Bild}(\Phi) = \bigcup_{x \in U} \Phi(U_x)$ offen.

(b) Die Umkehrabbildung $\Psi : \text{Bild}(\Phi) \rightarrow U$ ist stetig, denn für jede offene Menge $U' \subseteq U$ ist nach (a) die Urbildmenge $\Psi^{-1}(U') = \Phi(U')$ offen. Aus Satz 3.8 folgt nun die Behauptung.

□

Mit dem globalen Umkehrsatz kann man nun einfach nachweisen, dass die Polar-, Kugel- und Zylinderkoordinaten Diffeomorphismen sind.

Kontrollfrage 10.6. *Sei $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ein lokaler Diffeomorphismus um den Punkt 0. Welche der nachfolgenden Aussagen gilt dann?*

(A) Die Abbildung $D\phi(0) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ist invertierbar.

✓: Siehe Lemma 3.7. (Die lineare Abbildung $D\phi(0)$ ist invertierbar genau dann, wenn die zugehörige Abbildungsmatrix, nämlich die Jakobimatrix, invertierbar ist.)

(B) Die Abbildung $D\phi(x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ist für alle $x \in \mathbb{R}^n$ invertierbar.

✗: $D\phi(\cdot)$ muss nicht einmal auf ganz $\mathbb{R}^n$ definiert sein.

(C) Es existiert eine offene Umgebung U von 0, so dass Abbildung $D\phi(x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ für alle $x \in U$ invertierbar ist.

✓: Siehe Lemma 3.7.

(D) $\det J_\Phi(0) \neq 0$.

✓: Siehe Lemma 3.7. (Invertierbar $\Leftrightarrow \det = 0$.)

3.2 Satz über implizite Funktionen

Jede Gerade $G \subseteq \mathbb{R}^2$ kann als Menge aller Lösungen $(x, y) \in \mathbb{R}^2$ einer Gleichung der Form

$$f(x, y) = Ax + By + C = 0$$

geschrieben werden, wobei $A, B, C \in \mathbb{R}$ mit $(A, B) \neq (0, 0)$. Wir nennen dies eine **implizite Beschreibung** der Geraden.

Durch Auflösen nach x oder y erhalten wir eine **explizite Beschreibung**:

- Ist $B \neq 0$, dann gilt

$$f(x, \varphi(x)) = 0, \quad \varphi : \mathbb{R} \rightarrow \mathbb{R}, \quad \varphi(x) = -\frac{A}{B}x - \frac{C}{B}$$

Also ist G der Graph von φ über der x -Achse.

- Ist $A \neq 0$, dann ist G der Graph der Funktion

$$f(\varphi(y), y) = 0, \quad \varphi : \mathbb{R} \rightarrow \mathbb{R}, \quad \varphi(y) = -\frac{B}{A}y - \frac{C}{A}$$

Also ist G der Graph von φ über der y -Achse.

Da $(A, B) \neq (0, 0)$ muss mindestens einer der Fälle gelten.

Wir stellen uns nun die gleiche Frage für eine nichtlineare Gleichung:

Beispiel 3.13. Die Lösungsmenge von $f(x, y) := x^2 + y^2 - 1 = 0$ ist die Einheitssphäre $S^1 \subseteq \mathbb{R}^2$.

- Können wir S^1 als Graph einer Funktion über der x -Achse darstellen?
Für x mit $|x| \leq 1$ kann die Gleichung nach y aufgelöst werden:

$$\Rightarrow \text{zwei Lösungen:} \quad \pm \sqrt{1 - x^2}.$$

Mit den Lösungsfunktionen $\varphi_{\pm} : [-1, 1] \rightarrow \mathbb{R}$, $\varphi_{\pm}(x) = \pm \sqrt{1 - x^2}$ gilt

$$\text{Graph}(\varphi_{\pm}) \subsetneq S^1, \quad S^1 = \text{Graph}(\varphi_+) \cup \text{Graph}(\varphi_-).$$

- Sei (x_*, y_*) eine Lösung der Gleichung. Können wir S^1 lokal um (x_*, y_*) als Graph einer Funktion über der x -Achse darstellen? Das heißt: Existiert eine offene Umgebung $U = U' \times U''$ von (x_*, y_*) und eine Abbildung $\varphi : U' \rightarrow U''$ mit

$$(\star) \quad \forall (x, y) \in U : \left(f(x, y) = 0 \Leftrightarrow y = \varphi(x) \right) \quad ?$$

- Im Fall $|x_*| = 1$ ist das nicht möglich.
- Im Fall $|x_*| < 1$ existiert ein offenes Intervall $U' \subseteq (-1, 1)$ mit $x_* \in U'$.
Setze

$$U'' := \begin{cases} (0, \infty) & \text{falls } y_* > 0, \\ (0, -\infty) & \text{falls } y_* < 0, \end{cases} \quad \varphi : U' \rightarrow U'', \quad \varphi(x) := \begin{cases} \varphi_+(x) & \text{falls } y_* > 0, \\ \varphi_-(x) & \text{falls } y_* < 0. \end{cases}$$

Dann gilt $(\star)$.

- Was können wir über die Ableitung von φ sagen?

Aus $\varphi \in C^1(U')$, der Identität $f(x, \varphi(x)) = 0$ und der Kettenregel folgt:

$$0 = \frac{\partial}{\partial x} (f(x, \varphi(x))) = \frac{\partial f}{\partial x}(x, \varphi(x)) + \frac{\partial f}{\partial y}(x, \varphi(x)) \frac{\partial \varphi}{\partial x}(x).$$

Falls $\frac{\partial f}{\partial y}(x, y) \neq 0$, dann erhalten wir

$$\frac{\partial \varphi}{\partial x}(x) = - \frac{\frac{\partial f}{\partial x}(x, \varphi(x))}{\frac{\partial f}{\partial y}(x, \varphi(x))}.$$

D.h. wir können die Ableitung der Lösungsfunktion φ berechnen, ohne diese explizit zu kennen!

Als nächstes betrachten wir nichtlineare Gleichungssysteme bestehend aus

- $N = n + m$ Unbekannten $(x, y) = (x_1, \dots, x_n, y_1, \dots, y_m)$, und
- m Gleichungen

$$\begin{aligned} f_1(x_1, \dots, x_n, y_1, \dots, y_m) &= 0, \\ &\vdots \\ f_m(x_1, \dots, x_n, y_1, \dots, y_m) &= 0. \end{aligned}$$

Dieses Gleichungssystem lässt sich kompakt aufschreiben:

$$f(x, y) = 0 \quad \text{mit } x = (x_1, \dots, x_n), y = (y_1, \dots, y_m) \text{ und } f = (f_1, \dots, f_m).$$

Hierbei wird f auf einer offenen Menge $U \subseteq \mathbb{R}^n \times \mathbb{R}^m$ definiert sein.

Die Jacobi-Matrix von f in (x, y) ist eine $m \times (n + m)$ -Matrix und lässt sich horizontal in zwei Blöcke zerlegen:

$$J_f(x, y) = \left(\underbrace{\partial_x f(x, y)}_{\in \mathbb{R}^{m \times n}} \mid \underbrace{\partial_y f(x, y)}_{\in \mathbb{R}^{m \times m}} \right)$$

mit

$$\begin{aligned} \partial_x f(x, y) &:= \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x, y) & \cdots & \frac{\partial f_1}{\partial x_n}(x, y) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x, y) & \cdots & \frac{\partial f_m}{\partial x_n}(x, y) \end{pmatrix}, \\ \partial_y f(x, y) &:= \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(x, y) & \cdots & \frac{\partial f_1}{\partial y_m}(x, y) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial y_1}(x, y) & \cdots & \frac{\partial f_m}{\partial y_m}(x, y) \end{pmatrix}. \end{aligned}$$

Satz 3.14 (Satz über implizite Funktionen). Sei $U \subseteq \mathbb{R}^n \times \mathbb{R}^m$ offen und $f : U \rightarrow \mathbb{R}^m$ stetig differenzierbar. Es gelte:

- $(x_*, y_*) \in U$ ist eine Lösung von $f(x_*, y_*) = 0$.
- $\partial_y f(x_*, y_*)$ ist invertierbar.

Dann lässt sich die Gleichung $f(x, y) = 0$ lokal bei (x_*, y_*) nach y auflösen, d.h.: Es existieren eine offene Umgebung $U' \subseteq \mathbb{R}^n$ von x_* und eine offene Umgebung $U'' \subseteq \mathbb{R}^m$ von y_* sowie eine stetig differenzierbare Funktion $\varphi : U' \rightarrow U''$, so dass

(a) $(x_*, y_*) \in U' \times U'' \subseteq U$, und

(b) für alle $(x, y) \in U' \times U''$ gilt $f(x, y) = 0$ genau dann, wenn $y = \varphi(x)$.

Für alle $x \in U'$ gilt weiter

$$(3.7) \quad J_\varphi(x) = -\left(\partial_y f(x, \varphi(x))\right)^{-1} \partial_x f(x, \varphi(x))$$

Beweis. Setze

$$\Phi : U \rightarrow \mathbb{R}^n \times \mathbb{R}^m, \quad (x, y) \mapsto (x, f(x, y)).$$

1. Behauptung: Es existiert eine offene Umgebung $\tilde{U}$ von (x_*, y_*) und eine offene Menge $\tilde{V} \subseteq \mathbb{R}^n \times \mathbb{R}^m$, sodass $\Phi|_{\tilde{U}} : \tilde{U} \rightarrow \tilde{V}$ ein Diffeomorphismus ist.

Da Φ stetig differenzierbar ist, genügt es zu zeigen, dass $J_\Phi(x_*, y_*)$ invertierbar ist. Die Behauptung folgt dann aus dem Satz über den lokalen Diffeomorphismus. Die Jacobi-Matrix von Φ hat folgende Blockgestalt:

$$J_\Phi(x_*, y_*) = \begin{pmatrix} I & 0 \\ A & B \end{pmatrix}$$

wobei I die Einheitsmatrix in $\mathbb{R}^{n \times n}$ bezeichnet, und

$$A = \partial_x f(x_*, y_*) \in \mathbb{R}^{m \times n}, \quad B = \partial_y f(x_*, y_*) \in \mathbb{R}^{m \times m}.$$

Die inverse Matrix zu $J_\Phi(x_*, y_*)$ ist durch

$$\begin{pmatrix} I & 0 \\ -B^{-1}A & B^{-1} \end{pmatrix}$$

gegeben.

2. Es bezeichne $\Psi : \tilde{V} \rightarrow \tilde{U}$ die Umkehrabbildung von $\Phi|_{\tilde{U}}$. Die Umkehrabbildung Ψ hat die gleiche Bauart wie Φ : Es gilt

$$\Psi(\xi, \eta) = (\xi, g(\xi, \eta)),$$

mit einer stetig differenzierbaren Funktion $g(\xi, \eta) : \tilde{V} \rightarrow \mathbb{R}^m$.

Für alle $(x, y) \in \tilde{U}$ gilt

$$(\star) \quad f(x, y) = 0 \quad \Leftrightarrow \quad \Phi(x, y) = (x, 0) \quad \Leftrightarrow \quad y = g(x, 0).$$

Insbesondere gilt $y_* = g(x_*, 0)$. Wir wählen nun eine offene Umgebung $U' \subseteq \mathbb{R}^n$ von x_* und eine offene Umgebung $U'' \subseteq \mathbb{R}^m$ von y_* , so dass

$$U' \times U'' \subseteq \tilde{U} \quad \text{und} \quad \{g(x, 0) : x \in U'\} \subseteq U''.$$

Dies ist möglich, da g stetig in $(x_*, 0)$ ist. Durch $\varphi : U' \rightarrow U''$, $\varphi(x) = g(x, 0)$ wird nun ein stetig differenzierbare Funktion definiert. Behauptung (b) folgt aus $(\star)$.

3. Aus $f(x, \varphi(x)) = 0$ und der Kettenregel folgt (3.7). □

Beispiel 3.15 (Niveaulinie). Sei $U \subseteq \mathbb{R}^2$ offen und $f \in C^1(U)$. Sei $(x_*, y_*) \in U$ und $c := f(x_*, y_*)$. Frage: Wie sieht $N_c(f)$ in der Nähe von (x_*, y_*) aus?

- (a) Falls $\frac{\partial}{\partial y}f(x_*, y_*) \neq 0$, dann existiert nach Satz 3.14 ein offenes Intervall $I = (x_* - \delta, x_* + \delta)$ und $\varphi \in C^1(I)$ mit $\varphi(x_*) = y_*$ und $f(x, \varphi(x)) = c$ für alle $x \in I$. Dieser Fall tritt für $f(x, y) = x^2 + y^2 - 1 = 0$ für alle $(x_*, y_*) \in S^1 \setminus \{(1, 0), (-1, 0)\}$ ein.
- (b) Falls $\frac{\partial}{\partial x}f(x_*, y_*) \neq 0$, dann existiert nach Satz 3.14 ein offenes Intervall $J = (y_* - \delta, y_* + \delta)$ und $\psi \in C^1(J)$ mit $\psi(y_*) = x_*$ und $f(\psi(y), y) = c$ für alle $y \in J$. Dieser Fall tritt für $f(x, y) = x^2 + y^2 - 1 = 0$ für alle $(x_*, y_*) \in S^1 \setminus \{(0, 1), (0, -1)\}$ ein.
- (c) Im singulären Fall $\nabla f(x_*, y_*) = 0$ können verschiedene Phänomene auftreten:
- $f(x, y) = x^3 - y^2 = 0$ ist singulär in $(0, 0)$. Auflösen nach x führt auf $\psi(y) = \sqrt[3]{y^2}$. ψ ist nicht differenzierbar für $y = 0$ und $N_0(f)$ hat eine Spitze in $(0, 0)$.
 - $f(x, y) = x^2 - y^2 = 0$ ist singulär in $(0, 0)$. $N_0(f)$ besteht aus den beiden Winkelhalbierenden.
 - $f(x, y) = r^6 \sin(r^{-2}) = 0$ (in Polarkoordinaten, d.h. $r = \sqrt{x^2 + y^2}$), $f(0) = 0$ und f ist singulär in $(0, 0)$. $N_0(f)$ in einer Umgebung des Ursprungs besteht aus dem Ursprung und unendlich vielen Kreisen um 0.

Kontrollfrage 10.7. Sei $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, $(x, y, z) \mapsto f(x, y, z)$ eine stetig differenzierbare Funktion mit $f(0, 0, 0) = 0$. Welche der nachfolgenden Aussagen sind wahr?

- (A) Gilt $\frac{\partial}{\partial x}f(0, 0, 0) \neq 0$, dann existiert eine offene Umgebung $U \subseteq \mathbb{R}^2$ von $(0, 0)$ und eine Abbildung $\varphi : U \rightarrow \mathbb{R}$ mit $f(\varphi(y, z), y, z) = 0$ für alle $(y, z) \in U$.

✓: Satz 3.14.

- (B) Gilt $\frac{\partial}{\partial z}f(0, 0, 0) \neq 0$, dann existiert eine offene Umgebung $U \subseteq \mathbb{R}^2$ von $(0, 0)$ und eine Abbildung $\varphi : U \rightarrow \mathbb{R}$ mit $f(x, y, \varphi(x, y)) = 0$ für alle $(x, y) \in U$.

✓: Satz 3.14.

Kontrollfrage 10.8. Es bezeichne $S \subseteq \mathbb{R}^2$ die Lösungsmenge der Gleichung

$$x^2 + y^3 + y = 1.$$

Welche der nachfolgenden Aussagen ist dann wahr?

- (A) Für alle $(x_*, y_*) \in S$ existiert eine offene Umgebung O von (x_*, y_*) , so dass $S \cap O$ als Graph einer Funktion über y geschrieben werden kann.

×: Sei y^* so gewählt, dass $(y^*)^3 + y^* = 1$. (Das existiert wegen des Zwischenwertsatzes.) Dann ist $(0, y^*) \in S$. Aber für jede Umgebung von $(0, y^*)$ gilt:

$$\exists (x, y) \in \mathbb{R}^2 : (x, y) \in S \text{ und } (-x, y) \in S.$$

(B) Für alle $(x_*, y_*) \in S$ existiert eine offene Umgebung O von (x_*, y_*) , so dass $S \cap O$ als Graph einer Funktion über x geschrieben werden kann.

✓: Wegen $\frac{\partial}{\partial y_*} f(x^*, y^*) = 3(y^*)^2 + 1 \neq 0$ für beliebiges $(x^*, y^*) \in S$ gibt es nach Satz 3.14 eine Umgebung U_{x^*} von x^* und eine Umgebung U_{y^*} von y^* so, dass $S \cap (U_{x^*} \times U_{y^*})$ als Graph („über der ersten Komponente, also der x -Komponente“) dargestellt werden kann. Und $O := U_{x^*} \times U_{y^*}$ ist auch eine Umgebung von (x^*, y^*) .

15.6.2020

#33

22.6.2020

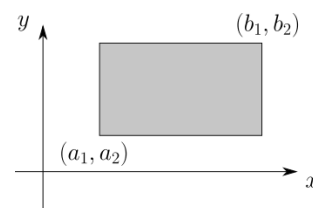
4 Das Riemannintegral

4.1 Motivation

Sei $A \subseteq \mathbb{R}^2$ eine beschränkte Menge. Wir möchten den Flächeninhalt von A bestimmen. Wir nennen diesen auch das zweidimensionale Volumen von A und bezeichnen es mit $\text{Vol}_2(A)$.

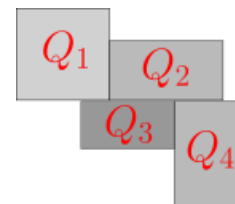
- Falls $A = [a_1, b_1] \times [a_2, b_2]$ ein Quader ist, dann gilt

$$\text{Vol}_2(A) = (b_1 - a_1)(b_2 - a_2).$$



- Falls A sich als Vereinigung endlich vieler paarweise disjunkter Quader $Q_1, \dots, Q_N$ schreiben lässt, dann gilt

$$(4.1) \quad \text{Vol}_2(A) = \sum_{j=1}^N \text{Vol}_2(Q_j).$$



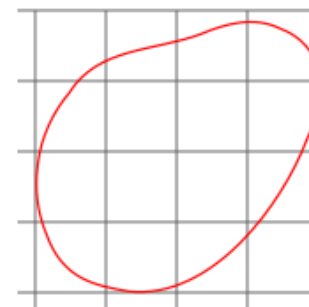
(Die Formel gilt nicht, falls die Quader nicht disjunkt sind. Aus einer Vereinigung endlich vieler, sich möglicherweise überlappender Quader können wir aber immer eine endliche Vereinigung disjunkter Quader machen.)

Wie kann man $\text{Vol}_2(A)$ für kompliziertere Mengen bestimmen / definieren? Zwei Strategien:

- “Zerlegung in Quadrate”: Für $j \in \mathbb{N}$ bezeichne

$$\mathcal{Q}_j = \left\{ 2^{-j} \left(\left[k_1 - \frac{1}{2}, k_1 + \frac{1}{2} \right) \times \left[k_2 - \frac{1}{2}, k_2 + \frac{1}{2} \right) \right) : k = (k_1, k_2) \in \mathbb{Z}^2 \right\}$$

die Menge aller (halb-offenen) Quadrate mit Kantenlänge 2^{-j} und Mittelpunkte in $2^{-j}\mathbb{Z}^2$. (Man stelle sich ein unendlich ausgedehntes Schachbrett vor, dessen quadratische Felder die Kantenlänge 2^{-j} und den Flächeninhalt 2^{-2j} besitzen.)



Für alle $j \in \mathbb{N}$ gilt

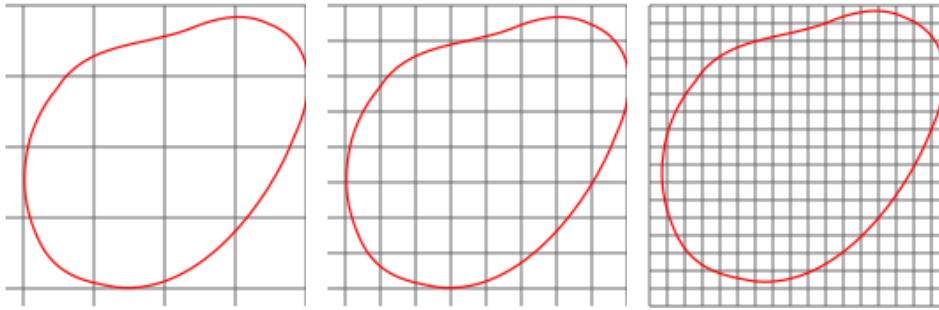
$$\underline{A}_j := \bigcup_{\substack{Q \in \mathcal{Q}_j \\ Q \subseteq A}} Q \subseteq A \subseteq \bigcup_{\substack{Q \in \mathcal{Q}_j \\ Q \cap A \neq \emptyset}} Q =: \bar{A}_j.$$

Das Volumen ist monoton im Sinne von $A \subseteq A' \Rightarrow \text{Vol}_2(A) \leq \text{Vol}_2(A')$. Daher gilt

$$\text{Vol}_2(\underline{A}_j) \leq \text{Vol}_2(A) \leq \text{Vol}_2(\bar{A}_j)$$

Die approximativen Flächeninhalte links und rechts können wir mittels (4.1) berechnen. Wir müssen nur die Quadrate in $\underline{A}_j$ (bzw. $\bar{A}_j$) zählen und die Anzahl mit 2^{-2j} multiplizieren.

Wir werden sehen, dass für “schöne” A die approximativen Flächeninhalte $\text{Vol}_2(\underline{A}_j)$, $\text{Vol}_2(\bar{A}_j)$ für $j \rightarrow \infty$ gegen den selben Grenzwert konvergieren. Wir definieren $\text{Vol}_2(A)$ als diesen Grenzwert.

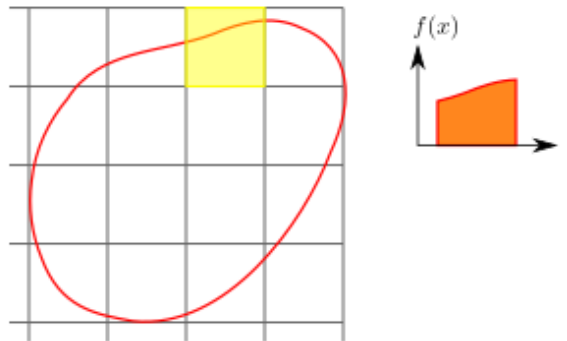


Dieses Verfahren kann auch auf höherdimensionale Mengen A ausgedehnt werden.

- “Zerlegung in Quadrate und Hypographe”.²¹ und Wir können A auch in Quadrate zerlegen, die vollständig in A liegen und Quadrate die ∂A schneiden. Für eine “schöne” Menge A kann der Schnitt eines Quadrates Q mit A (falls Q klein genug ist) als Hypograph einer Funktion $f : [a, b] \rightarrow \mathbb{R}$ (geschnitten mit dem oberen Halbraum) dargestellt werden, d.h.

$$Q \cap A = A(f) := \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, 0 \leq y \leq f(x)\}.$$

Hierbei bezeichnet $f : [a, b] \rightarrow \mathbb{R}_{\geq}$ eine lokale Parametrisierung des Randes von ∂A (in Q) über der x -Achse.



²¹Korrektur: “Epigraph” und die Schreibweise $\text{epi}(f)$ wurde durch “Hypograph” und $A(f)$ ersetzt.

Die Fläche der Menge $A(f)$ ist dann durch das Integral

$$\int_a^b f(x) dx$$

gegeben – ein Objekt, dass wir in diesem Kapitel einführen werden.

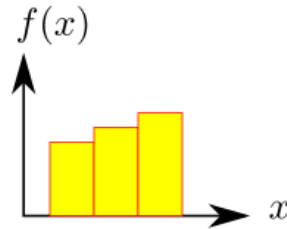
Wie können wir den Wert des Integrals bestimmen?

(i) Ist f eine konstante Funktion, z.B. $f(x) = c > 0$, dann ist $A(f)$ ein Quader der Form $[a, b] \times [0, c]$ und es soll daher gelten

$$\int_a^b f(x) dx = \text{Vol}_2([a, b] \times [0, c]) = (b - a)c.$$

(ii) Sei f stückweise konstant mit endlich vielen Stücken, d.h. es gibt disjunkte Intervalle $I_1, \dots, I_M \subseteq [a, b]$ und $c_1, \dots, c_M \geq 0$, sodass

$$f(x) = \sum_{j=1}^M c_j \mathbf{1}_{I_j}(x), \quad \mathbf{1}_{I_j}(x) := \begin{cases} 1 & \text{falls } x \in I_j, \\ 0 & \text{sonst,} \end{cases}$$

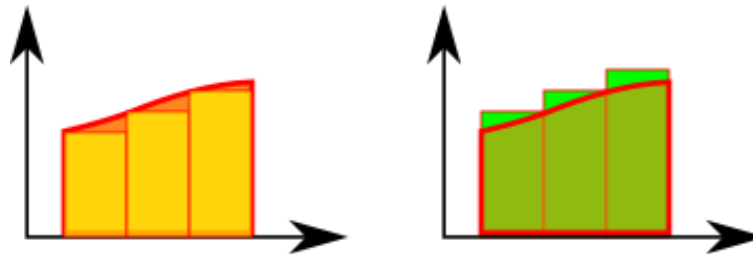


Wir können uns dann $A(f)$ als endliche Vereinigung der disjunkten Quader $I_j \times [0, c_j]$ vorstellen. Daher soll gelten:

$$\int_a^b f(x) dx = \sum_{j=1}^M c_j \text{Vol}_1(I_j),$$

wobei $\text{Vol}_1(I_j)$ die Länge des Intervalls I_j bezeichnet.

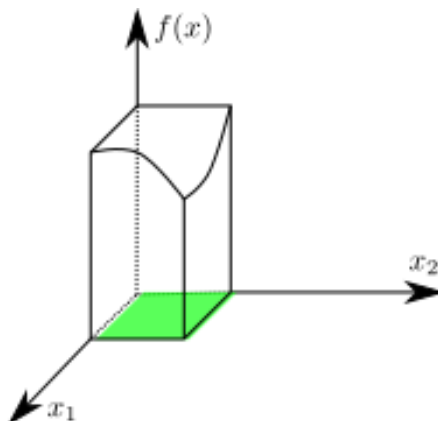
(iii) Für kompliziertere Funktionen f werden wir f von oben und unten durch Funktionen wie in (ii) annähern. Dies führt auf den Begriff des Riemannintegrals.



Wir werden das Riemannintegral nicht nur für Funktionen auf einem Intervall einführen, sondern gleich für Funktionen auf $\mathbb{R}^n$. Wir wählen dabei den Zugang von Darboux.

Mehrdimensionale Integrale haben viele Anwendungen:

- (Geometrie). Sei $n = 2$, $Q \subseteq \mathbb{R}^2$ ein Quader, $f : Q \rightarrow \mathbb{R}_{\geq}$ eine Funktion. Dann hat $\int_Q f(x) dx$ die Bedeutung des (dreidimensionalen) Volumes von $A(f)$.



- (Physik). Sei $n = 3$. Wir stelle uns vor, dass ein Quader $Q \subseteq \mathbb{R}^3$ von einem Material ausgefüllt sei, dessen Massendichte in jedem Punkt $x \in Q$ durch $f(x) \geq 0$ beschrieben sei. Dann hat $\int_Q f(x) dx$ die Bedeutung der Gesamtmasse des Quaders.
- (Wahrscheinlichkeitstheorie). Bei einem Glücksrad sei die Wahrscheinlichkeit, dass ein Winkel $\alpha \in [0, 2\pi)$ erspielt wird, durch eine Wahrscheinlichkeitsdichte $\rho : [0, 2\pi) \rightarrow [0, 1)$ gegeben, d.h. die Wahrscheinlichkeit, dass der Winkel α in im Intervall $I \subseteq [0, 2\pi)$ landet ist gerade $\int_I \rho(\alpha) d\alpha$. Es muss $\int_{[0, 2\pi)} \rho(\alpha) d\alpha = 1$ gelten. Im Falle $\rho = \frac{1}{2\pi}$ tritt jeder Winkel mit der gleichen Wahrscheinlichkeit auf.

Sei nun $n \in \mathbb{N}_+$. Wir drehen n -mal (unabhängig) das Glücksrad und erhalten einen Vektor aus Winkeln $x := (\alpha_1, \dots, \alpha_n)$. Sei $A \subseteq [0, 2\pi)^n$. Die Wahrscheinlichkeit, dass ein x in A erspielt wird, ist

$$\int_A \rho(x) dx \quad \text{mit} \quad x = (x_1, \dots, x_n), \quad \rho(x) := \rho(x_1) \cdot \dots \cdot \rho(x_n).$$

Es gibt verschiedene Integral-Begriffe. Wir betrachten hier das Riemann- (oder auch Darboux-) Integral. Typischerweise stimmen diese für “schöne” Mengen und Funktionen überein.

In der Vorlesung “Maß- und Integrationstheorie” wird das Lebesgue-Integrals eingeführt.

Dieser moderne Integralbegriff bildet die Grundlage für die moderne Wahrscheinlichkeitstheorie und Funktionalanalysis.

22.6.2020

#34

22.6.2020

4.2 Grundlegende Eigenschaften

Ziele:

- Begriffe: Quader, Treppenfunktion, Integral einer Treppenfunktion
- Riemannintegral und grundlegende Eigenschaften
- Riemannsches Integrierbarkeitskriterium
- Identität von Fubini.

Treppenfunktionen

Definition 4.1 (Quader und Treppenfunktion). • Ein Quader in $\mathbb{R}^n$ ist eine Menge der Form

$$Q = [a, b] := [a_1, b_1] \times \dots \times [a_n, b_n]$$

mit $a, b \in \mathbb{R}^n$ wobei $a_i \leq b_i$ für alle $i = 1, \dots, n$. Wir bezeichnen mit $\mathcal{Q}(\mathbb{R}^n)$ die Menge aller Quader in $\mathbb{R}^n$. Für $\alpha, \beta \in \mathbb{R}$ setzen wir

$$[\alpha, \beta]^n := \underbrace{[\alpha, \beta] \times \dots \times [\alpha, \beta]}_{=n\text{-mal}}$$

- Das n -dimensionale Volumen eines Quaders $Q = [a, b]$ ist durch

$$\text{Vol}_n(Q) := (b_1 - a_1) \cdot \dots \cdot (b_n - a_n)$$

definiert.

- Für eine Menge $A \subseteq \mathbb{R}^n$ bezeichnet

$$\mathbf{1}_A : \mathbb{R}^n \rightarrow \mathbb{R}, \quad \mathbf{1}_A(x) := \begin{cases} 1 & x \in A, \\ 0 & x \notin A, \end{cases}$$

die Indikatorfunktion (oder auch charakterisistische Funktion) der Menge A .

- Eine Treppenfunktion (auf $\mathbb{R}^n$) ist eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ der Form

$$(4.2) \quad f = \sum_{j=1}^N c_j \mathbf{1}_{Q_j},$$

mit $N \in \mathbb{N}_+$, $c_1, \dots, c_N \in \mathbb{R}$ und $Q_1, \dots, Q_N \in \mathcal{Q}(\mathbb{R}^n)$. Wir bezeichnen mit

$$T(\mathbb{R}^n) := \{ f : f \text{ ist eine Treppenfunktion auf } \mathbb{R}^n \}$$

den Vektorraum aller Treppenfunktionen.

Bemerkung 4.2. • $\mathcal{Q}(\mathbb{R}^n)$ enthält auch den leeren Quader $\emptyset$.

- $T(\mathbb{R}^n)$ ist gerade die lineare Hülle von $\{\mathbf{1}_Q : Q \in \mathcal{Q}(\mathbb{R}^n)\} \subseteq \text{Abb}(\mathbb{R}^n; \mathbb{R})$ und damit ein Unterraum von $\text{Abb}(\mathbb{R}^n; \mathbb{R})$.

- Jede Treppenfunktion besitzt einen kompakten Träger²². Der Wertebereich einer Treppenfunktion besteht aus endlich vielen reellen Zahlen. Jede Treppenfunktion ist beschränkt.
- Die Quader in der Darstellung (4.2) müssen weder disjunkt noch eindeutig sein. Beispiel: Mit

$$Q := [0, 1) \times \cdots \times [0, 1),$$

$$Q_1 := [0, \frac{1}{2}) \times [0, 1) \times \cdots \times [0, 1), \quad Q_2 := [\frac{1}{2}, 1) \times [0, 1) \times \cdots \times [0, 1),$$

gilt

$$\mathbf{1}_Q = \sum_{j=1}^2 \mathbf{1}_{Q_j}.$$

- In (4.2) treten nur endlich viele Terme auf. Aber: Setze $x_j := 1 - 2^{-j}$ für $j \in \mathbb{N}$. Dann werden durch

$$Q_j := [x_j, x_{j+1}) \times [0, 1) \times \cdots \times [0, 1) \subseteq \mathbb{R}^n$$

abzählbar unendlich viele, disjunkte Quader definiert und es gilt für

$$f := \sum_{j=0}^{\infty} \mathbf{1}_{Q_j}, \quad Q := [0, 1) \times \cdots \times [0, 1) :$$

f ist mit $\mathbf{1}_Q$ identisch und damit eine Treppenfunktion,

$$g := \sum_{j=0}^{\infty} 2^{-j} \mathbf{1}_{Q_j}$$

(punktweiser Grenzwert) aber nicht.

- Sei $1 \leq m < n$ und $Q = [a, b) \in \mathcal{Q}(\mathbb{R}^n)$. Setze

$$Q' := [a_1, b_1) \times \cdots \times [a_m, b_m) \in \mathcal{Q}(\mathbb{R}^m),$$

$$Q'' := [a_{m+1}, b_{m+1}) \times \cdots \times [a_n, b_n) \in \mathcal{Q}(\mathbb{R}^{n-m}).$$

Dann gilt

$$Q = Q' \times Q'', \quad \text{Vol}_n(Q) = \text{Vol}_m(Q') \cdot \text{Vol}_{n-m}(Q'').$$

Lemma 4.3. Für beliebige Quader $Q, Q' \in \mathcal{Q}(\mathbb{R}^n)$ gilt

(a) $Q \cap Q' \in \mathcal{Q}(\mathbb{R}^n)$.

(b) Es existieren paarweise disjunkte (möglicherweise leere) Quader $Q_1, \dots, Q_{3^n-1}$ mit

$$Q \setminus Q' = \bigcup_{j=1}^{3^n-1} Q_j$$

²²Für eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ist der Träger als Abschluss der Menge $\{x \in \mathbb{R}^n : f(x) \neq 0\}$ definiert.

(c) Es existieren paarweise disjunkte (möglicherweise leere) Quader $Q_1, \dots, Q_{2 \cdot 3^n - 1}$ mit

$$Q = \bigcup_{j=1}^{3^n} Q_j, \quad Q' = \bigcup_{j=3^n}^{2 \cdot 3^n - 1} Q_j, \quad Q \cap Q' = Q_{3^n}.$$

Beweis. (a) Sei $Q = [a, b]$ und $Q' = [a', b']$. Definiere $a'', b'' \in \mathbb{R}^n$ durch

$$a''_j := \max\{a_j, a'_j\}, \quad b''_j := \min\{b_j, b'_j\}.$$

Dann gilt

$$\begin{aligned} x \in Q \cap Q' &\Leftrightarrow \forall j = 1, \dots, n : (a_j \leq x_j < b_j) \text{ und } (a'_j \leq x_j < b'_j) \\ &\Leftrightarrow \forall j = 1, \dots, n : (a''_j \leq x_j < b''_j) \\ &\Leftrightarrow x \in [a'', b'']. \end{aligned}$$

(b) Da $Q \setminus Q' = Q \setminus (Q \cap Q')$ und da $Q \cap Q' \in \mathcal{Q}(n)$ dürfen wir O.B.d.A. annehmen, dass $Q' \subseteq Q$, d.h. für alle $j = 1, \dots, n$ gilt $a_j \leq a'_j \leq b'_j \leq b_j$. Setze

$$I_{j1} := [a_j, a'_j), \quad I_{j2} := [b'_j, b_j), \quad I_{j3} := [a'_j, b'_j).$$

Für jedes n -Tupel $\sigma = (\sigma_1, \dots, \sigma_n) \in \{1, 2, 3\}^n$ setzen wir

$$Q_\sigma := I_{1\sigma_1} \times \dots \times I_{n\sigma_n}.$$

Hierdurch werden 3^n disjunkte (möglicherweise leer) Quader definiert. Es gilt $Q = \bigcup_{\sigma \in \{1, 2, 3\}^n} Q_\sigma$, denn $x \in Q$ genau dann, wenn $x_j \in I_{j\sigma_j}$ für ein $\sigma_j \in \{1, 2, 3\}$ und damit $x \in Q_\sigma$. Weiter gilt nach Konstruktion $Q' = Q_{\sigma'}$ mit $\sigma' = \{3, 3, \dots, 3\}$ und damit ist $Q \setminus Q'$ die Vereinigung aller Quader Q_σ mit $\sigma \neq \sigma'$.

(c) Das folgt direkt aus (b) und der Zerlegung $Q \cup Q' = (Q \setminus Q') \cup (Q \cap Q') \cup (Q' \setminus Q)$. □

Lemma 4.4.

(a) Für alle $f, g \in T(\mathbb{R}^n)$, $\lambda \in \mathbb{R}$ gilt $|f|, f + \lambda g, fg \in T(\mathbb{R}^n)$.

(b) Sei f eine Treppenfunktion auf $\mathbb{R}^n$ und $1 \leq m < n$. Dann ist

- für alle $x'' := (x_{m+1}, \dots, x_n) \in \mathbb{R}^{n-m}$ die Abbildung

$$f(\cdot, x'') : \mathbb{R}^m \rightarrow \mathbb{R}, \quad x' \mapsto f(x', x'')$$

eine Treppenfunktion auf $\mathbb{R}^m$.

- für alle $x' := (x_1, \dots, x_m) \in \mathbb{R}^m$ die Abbildung

$$f(x', \cdot) : \mathbb{R}^{n-m} \rightarrow \mathbb{R}, \quad x'' \mapsto f(x', x'')$$

eine Treppenfunktion auf $\mathbb{R}^{n-m}$.

Beweis. (a) $f + \lambda g, |f| \in T(\mathbb{R}^n)$ folgt direkt aus der Definition. Die Behauptung $fg \in T(\mathbb{R}^n)$ sieht man wie folgt: Sei $f = \sum_{j=1}^N c_j \mathbf{1}_{Q_j}$, $g = \sum_{j=1}^{N'} c'_j \mathbf{1}_{Q'_j}$. Dann gilt

$$fg = \sum_{j=1}^N \sum_{k=1}^{N'} c_j c'_k \mathbf{1}_{Q_{jk}}$$

mit $Q_{jk} = Q_j \cap Q'_k \in \mathcal{Q}(\mathbb{R}^n)$.

(b) Da Treppenfunktionen Linearkombinationen von Indikatorfunktionen von Quadern sind, genügt es die Behauptung für $f = \mathbf{1}_Q$ zu zeigen. Sei $Q = [a, b)$. Setze

$$Q' := [a_1, b_1) \times \cdots \times [a_m, b_m), \quad Q'' := [a_{m+1}, b_{m+1}) \times \cdots \times [a_n, b_n).$$

Dann gilt $Q = Q' \times Q''$ und

$$f(x', x'') = \mathbf{1}_{Q'}(x') \mathbf{1}_{Q''}(x'').$$

Also sind $f(x', \cdot)$ und $f(\cdot, x'')$ Treppenfunktionen. □

Definition 4.5 (Integral einer Treppenfunktion). Sei $f \in T(\mathbb{R}^n)$ eine Treppenfunktion mit $f = \sum_{j=1}^N c_j \mathbf{1}_{Q_j}$. Dann heißt

$$\int_{\mathbb{R}^n} f := \int_{\mathbb{R}^n} f(x) dx := \sum_{j=1}^N c_j \text{Vol}_n(Q_j)$$

das Integral von f .

Bemerkung 4.6. (a) Sei $Q \subseteq \mathbb{R}^n$ ein Quader, $c \geq 0$. Dann ist

$$\int_{\mathbb{R}^n} c \mathbf{1}_Q(x) dx = c \text{Vol}_n(Q) = \text{Vol}_{n+1}(Q \times [0, c))$$

gerade das $n+1$ -dimensionale Volumen des Quaders $Q \times [0, c) \subseteq \mathbb{R}^{n+1}$.

Für $n = 1$ ist $Q = [a, b)$ ein Intervall und $Q \times [0, c)$ ein Rechteck in der Ebene mit Kantenlängen $(b - a) \times c$.

Für $n = 2$ ist $Q = [a_1, b_1) \times [a_2, b_2)$ ein Rechteck in der Ebene und $Q \times [0, c)$ ein Quader im dreidimensionalen Raum mit Kantenlängen $(b_1 - a_1) \times (b_2 - a_2) \times c$.

(b) Seien die Quader Q_j in der Darstellung (4.2) von f disjunkt und die Koeffizienten $c_j \geq 0$. Dann ist $\int_{\mathbb{R}^n} f(x) dx$ gerade die Summe der $n+1$ -dimensionalen Volumen der N -Würfel $Q_j \times [0, c_j)$.

(c) Wir betrachten N disjunkte Würfel $Q_1, \dots, Q_N \in \mathcal{Q}(\mathbb{R}^3)$ und nehmen an, dass die Quader Q_j mit einem Material mit Dichte c_j gefüllt sind. Dann beschreibt

$$c_j \text{Vol}_3(Q_j)$$

die Masse des j ten Würfels und

$$\int_{\mathbb{R}^3} f(x) dx$$

die Gesamtmasse aller Würfel.

Schreibweise: ²³ Für Funktionen $f, g : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ schreiben wir

$$\begin{aligned} f < g &\Leftrightarrow f(x) < g(x) \text{ für alle } x \in \Omega, \\ f \leq g &\Leftrightarrow f(x) \leq g(x) \text{ für alle } x \in \Omega, \\ f = g &\Leftrightarrow f(x) = g(x) \text{ für alle } x \in \Omega. \end{aligned}$$

Proposition 4.7 (Eigenschaften des Integrals einer Treppenfunktion).

(a) (Wohldefiniertheit). Das Integral einer Treppenfunktion ist unabhängig von der speziellen Darstellung, d.h. für $f, \hat{f} \in T(\mathbb{R}^n)$ mit

$$f = \sum_{j=1}^N c_j \mathbf{1}_{Q_j}, \quad \hat{f} = \sum_{k=1}^{\hat{N}} \hat{c}_k \mathbf{1}_{\hat{Q}_k}$$

gilt

$$f = \hat{f} \quad \Rightarrow \quad \sum_{j=1}^N c_j \operatorname{Vol}_n(Q_j) = \sum_{k=1}^{\hat{N}} \hat{c}_k \operatorname{Vol}_n(\hat{Q}_k).$$

(b) (Linearität). Die Abbildung

$$T(\mathbb{R}^n) \rightarrow \mathbb{R}, \quad f \mapsto \int_{\mathbb{R}^n} f(x) dx.$$

ist linear.

(c) (Monotonie / Positivität). Für alle $f, g \in T(\mathbb{R}^n)$ gilt

$$f \leq g \quad \Rightarrow \quad \int_{\mathbb{R}^n} f \leq \int_{\mathbb{R}^n} g.$$

(d) (Dreiecksungleichung). Für alle $f \in T(\mathbb{R}^n)$ gilt $|f| \in T(\mathbb{R}^n)$ und

$$\left| \int_{\mathbb{R}^n} f \right| \leq \int_{\mathbb{R}^n} |f|.$$

(e) (Fubini). Sei $1 \leq m < n$ und $f \in T(\mathbb{R}^n)$. Dann sind die Funktionen

$$\mathbb{R}^m \rightarrow \mathbb{R}, \quad x' \mapsto \int_{\mathbb{R}^{n-m}} f(x', x'') dx''$$

und

$$\mathbb{R}^{n-m} \rightarrow \mathbb{R}, \quad x'' \mapsto \int_{\mathbb{R}^m} f(x', x'') dx'$$

Treppenfunktionen und es gilt

$$\begin{aligned} \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^{n-m}} f(x', x'') dx'' \right) dx' &= \int_{\mathbb{R}^{n-m}} \left(\int_{\mathbb{R}^m} f(x', x'') dx' \right) dx'' \\ &= \int_{\mathbb{R}^n} f(x) dx. \end{aligned}$$

²³Hinzugefügt am 1.7.2020

Im Beweis verwenden wir folgendes technisches Lemma:

Lemma 4.8. Seien $f, \hat{f} \in T(\mathbb{R}^n)$ durch

$$f = \sum_{j=1}^N c_j \mathbf{1}_{Q_j}, \quad \hat{f} = \sum_{k=1}^{\hat{N}} \hat{c}_k \mathbf{1}_{\hat{Q}_k}$$

gegeben. Dann existieren endlich viele, paarweise disjunkte Quader $\tilde{Q}_1, \dots, \tilde{Q}_M$ und Koeffizienten $\gamma_1, \dots, \gamma_M$ sowie $\hat{\gamma}_1, \dots, \hat{\gamma}_M$ mit

$$f = \sum_{\ell=1}^M \gamma_\ell \mathbf{1}_{\tilde{Q}_\ell} \quad \text{und} \quad \hat{f} = \sum_{\ell=1}^M \hat{\gamma}_\ell \mathbf{1}_{\tilde{Q}_\ell},$$

und es gilt

$$\sum_{j=1}^N c_j \operatorname{Vol}_n(Q_j) = \sum_{\ell=1}^M \gamma_\ell \operatorname{Vol}_n(\tilde{Q}_\ell), \quad \sum_{k=1}^{\hat{N}} \hat{c}_k \operatorname{Vol}_n(\hat{Q}_k) = \sum_{\ell=1}^M \hat{\gamma}_\ell \operatorname{Vol}_n(\tilde{Q}_\ell).$$

Beweis. 1. Behauptung: Es existiert eine Menge $\mathcal{A} = \{\tilde{Q}_1, \dots, \tilde{Q}_M\}$ bestehend aus endlich vielen, disjunkten Quadern in $\mathcal{Q}(\mathbb{R}^n)$, so dass sich jeder Quader $Q \in \mathcal{Q} := \{Q_1, \dots, Q_N, \hat{Q}_1, \dots, \hat{Q}_{\hat{N}}\}$ als Vereinigung von Quadern in $\mathcal{A}$ schreiben lässt.

Argument: Für jede Koordinatenrichtung $i = 1, \dots, n$ wähle reelle Zahlen $\tau_{i,1}, \dots, \tau_{i,M}$ (mit $M = 2(N + \hat{N})$), so dass

$$\{\tau_{i,1} \leq \dots \leq \tau_{i,M}\} := \left\{ \inf_{x \in Q} x_i, \sup_{x \in Q} x_i : Q \in \mathcal{Q} \right\},$$

— das sind gerade alle i ten Koordinaten von Eckpunkten der Quader $Q \in \mathcal{Q}$. Insbesondere existieren für alle $Q \in \mathcal{Q}$ Indizes $1 \leq j_i \leq j'_i \leq M$ für $i = 1, \dots, n$ mit

$$Q = [\tau_{1,j_1}, \tau_{1,j'_1}) \times \dots \times [\tau_{n,j_n}, \tau_{n,j'_n}).$$

Durch

$$\mathcal{A} := \{\tilde{Q} := [\tau_{1,j_1}, \tau_{1,j_1+1}) \times \dots \times [\tau_{n,j_n}, \tau_{n,j_n+1}) : 1 \leq j_1, \dots, j_n < M\}.$$

wird nun eine endliche Menge, disjunkter Quader definiert, die die Behauptung erfüllt.

2. Für $1 \leq j \leq N$ und $1 \leq \ell \leq M$ setze

$$\sigma_{j,\ell} := \begin{cases} 1 & \text{falls } \tilde{Q}_\ell \subseteq Q_j \\ 0 & \text{sonst.} \end{cases}$$

Dann gilt für alle $1 \leq j \leq N$:

$$\mathbf{1}_{Q_j} = \sum_{\ell=1}^M \sigma_{j,\ell} \mathbf{1}_{\tilde{Q}_\ell} \quad \text{und} \quad \operatorname{Vol}_n(Q_j) = \sum_{\ell=1}^M \sigma_{j,\ell} \operatorname{Vol}_n(\tilde{Q}_\ell).$$

(Wir überlassen das Argument dem aufmerksamen Leser).

Hieraus folgern wir

$$(4.3) \quad \sum_{j=1}^N c_j \mathbf{1}_{Q_j} = \sum_{j=1}^N c_j \sum_{\ell=1}^M \sigma_{j,\ell} \mathbf{1}_{\tilde{Q}_\ell} = \sum_{\ell=1}^M \underbrace{\left(\sum_{j=1}^N c_j \sigma_{j,\ell} \right)}_{=:\gamma_\ell} \mathbf{1}_{\tilde{Q}_\ell},$$

sowie

$$(4.4) \quad \begin{aligned} \sum_{j=1}^N c_j \operatorname{Vol}_n(Q_j) &= \sum_{j=1}^N c_j \sum_{\ell=1}^M \sigma_{j,\ell} \operatorname{Vol}_n(\tilde{Q}_\ell) \\ &= \sum_{\ell=1}^M \gamma_\ell \operatorname{Vol}_n(\tilde{Q}_\ell). \end{aligned}$$

Analog setzen wir für $1 \leq k \leq \hat{N}$ und $1 \leq \ell \leq M$,

$$\hat{\sigma}_{k,\ell} := \begin{cases} 1 & \text{falls } \tilde{Q}_\ell \subseteq \hat{Q}_k \\ 0 & \text{sonst.} \end{cases}$$

Dann gilt

$$(4.5) \quad \sum_{k=1}^{\hat{N}} \hat{c}_k \mathbf{1}_{\hat{Q}_k} = \sum_{\ell=1}^M \underbrace{\left(\sum_{k=1}^{\hat{N}} \hat{c}_k \hat{\sigma}_{k,\ell} \right)}_{=:\hat{\gamma}_\ell} \mathbf{1}_{\tilde{Q}_\ell},$$

sowie

$$(4.6) \quad \sum_{k=1}^{\hat{N}} \hat{c}_k \operatorname{Vol}_n(\hat{Q}_k) = \sum_{\ell=1}^M \hat{\gamma}_\ell \operatorname{Vol}_n(\tilde{Q}_\ell).$$

□

Beweis von Proposition 4.7. (a) Wir wenden Lemma 4.8 auf f und $\hat{f}$ an. Aus $f = \hat{f}$, (4.3), (4.5) folgern wir nun $\gamma_\ell = \hat{\gamma}_\ell$. Also sind die linken Seiten von (4.4) und (4.6) und damit die Integrale identisch.

(b) (Linearität). Das folgt aus der Linearität von Summen.

(c) (Monotonie / Positivität). Wegen Lemma 4.8 dürfen wir o.B.d.A. annehmen, dass f, g die Gestalt

$$f = \sum_{\ell=1}^M c_\ell \mathbf{1}_{Q_\ell}, \quad g = \sum_{\ell=1}^M \hat{c}_\ell \mathbf{1}_{Q_\ell}$$

haben. Aus $f \leq g$ folgt nun $c_\ell \leq \hat{c}_\ell$ und damit

$$\int_{\mathbb{R}^n} f = \sum_{\ell=1}^M c_\ell \operatorname{Vol}_n(Q_\ell) \leq \sum_{\ell=1}^M \hat{c}_\ell \operatorname{Vol}_n(Q_\ell) \leq \int_{\mathbb{R}^n} g.$$

(d) (Drecksungleichung). Das folgt aus der Drecksungleichung für Summen.

(e) (Fubini). Wegen der Linearität genügt es die Aussage für die Indikatorfunktion eines Quaders zu beweisen. Sei $f = \mathbf{1}_Q$ mit $Q \in \mathcal{Q}(\mathbb{R}^n)$. Nach Lemma 4.3 existieren $Q' \in \mathcal{Q}(\mathbb{R}^m)$ und $Q'' \in \mathcal{Q}(\mathbb{R}^{n-m})$ mit $Q = Q' \times Q''$. Es gilt für alle $x = (x', x'') \in \mathbb{R}^n$

$$\mathbf{1}_Q(x', x'') = \mathbf{1}_{Q'}(x') \mathbf{1}_{Q''}(x'')$$

sowie

$$\int_{\mathbb{R}^{n-m}} f(x', x'') dx'' = \mathbf{1}_{Q'}(x') \text{Vol}_{n-m}(Q'').$$

Damit ist $x' \mapsto \int_{\mathbb{R}^{n-m}} f(x', x'') dx'' = \text{Vol}_{n-m}(Q'') \mathbf{1}_{Q'}(x')$ eine Treppenfunktion. Weiter gilt

$$\int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^{n-m}} f(x', x'') dx'' \right) dx' = \text{Vol}_{n-m}(Q'') \text{Vol}_m(Q') = \text{Vol}_n(Q) = \int_{\mathbb{R}^n} \mathbf{1}_Q.$$

Die Aussage für $x'' \mapsto \int_{\mathbb{R}^m} f(x', x'') dx'$ zeigt man analog. □

Bemerkung 4.9. *Durch Iteration von Fubini erhalten wir*

$$\int_{\mathbb{R}^n} f(x) dx = \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x_1, \dots, x_n) dx_1 \dots dx_n,$$

weiterhin können wir auf der rechten Seite die Reihenfolge der Integration beliebig vertauschen.

22.6.2020

#35

22.6.2020

Riemann-integrierbare Funktionen

Beispiel 4.10 (Dyadische Zerlegung). Für $j \in \mathbb{N}$ setze

$$\mathcal{Q}_j := \{Q := 2^{-j}(z + [-\frac{1}{2}, \frac{1}{2})^n) : z \in \mathbb{Z}^n\}.$$

Die Quader in $\mathcal{Q}_j$ bilden eine uniforme Partitionierung von $\mathbb{R}^n$ in Quader mit Kantenlänge 2^{-j} .

Nach Konstruktion gilt: Jeder Quader in $\mathcal{Q}_{j+1}$ liegt in genau einem Quader in $\mathcal{Q}_j$.

Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine beschränkte Funktion mit Träger in der beschränkten Menge $\Omega \subseteq \mathbb{R}^n$. Für $j \in \mathbb{N}$ definiere

$$\bar{g}_j := \sum_{\substack{Q \in \mathcal{Q}_j \\ Q \cap \Omega \neq \emptyset}} \left(\sup_Q f \right) \mathbf{1}_Q, \quad \underline{g}_j := \sum_{\substack{Q \in \mathcal{Q}_j \\ Q \cap \Omega \neq \emptyset}} \left(\inf_Q f \right) \mathbf{1}_Q.$$

(Es tauchen nur endlich viele Terme auf!).

Für alle $j \in \mathbb{N}$ gilt:

- $\bar{g}_j$ und $\underline{g}_j$ sind Treppenfunktionen und es gilt

$$\underline{g}_j \leq f \leq \bar{g}_j.$$

- $(\bar{g}_j)_{j \in \mathbb{N}}$ ist monoton fallend und nach unten durch f beschränkt.
Argument: Seien $x \in \mathbb{R}^n$ und $j \in \mathbb{N}$. Dann existiert $Q \in \mathcal{Q}_{j+1}$ mit $x \in Q$ und es existiert $Q' \in \mathcal{Q}_j$ mit $Q \subseteq Q'$. Es gilt

$$\bar{g}_{j+1}(x) = \sup_Q f \leq \sup_{Q'} f = \bar{g}_j(x).$$

- $(\underline{g}_j)_{j \in \mathbb{N}}$ ist monoton wachsend und nach oben (durch f) beschränkt.
- Aufgrund der Monotonie des Integrals für Treppenfunktionen gilt

$$L_j := \int_{\mathbb{R}^n} \underline{g}_j \leq \int_{\mathbb{R}^n} \bar{g}_j =: U_j.$$

Weiter ist (L_j) (bzw. (U_j)) eine beschränkte monoton wachsende (bzw. fallende) Folge und damit konvergent.

- Sei f gleichmäßig stetig und $\varepsilon > 0$. Dann existiert $j_0 \in \mathbb{N}$, so dass

$$\forall j \geq j_0, \forall Q \in \mathcal{Q}_j, \forall x, y \in Q : |f(x) - f(y)| < \varepsilon.$$

Wir folgern, dass

$$\forall j \geq j_0, \forall Q \in \mathcal{Q}_j, \forall x, y \in Q : \sup_Q f - \inf_Q f < \varepsilon,$$

und damit

$$\forall j \geq j_0 : \|\bar{g}_j - \underline{g}_j\|_\infty < \varepsilon,$$

Also konvergieren $(\bar{g}_j)$ und $(\underline{g}_j)$ gleichmäßig gegen f für $j \rightarrow \infty$.

- Da $\Omega \subset \mathbb{R}^n$ beschränkt ist, existiert eine Konstante C mit

$$\forall j \in \mathbb{N} : \#\{Q \in \mathcal{Q}_j : Q \cap \Omega \neq \emptyset\} 2^{-nj} \leq C.$$

Wir folgern nun

$$0 \leq U_j - L_j = \sum_{\substack{Q \in \mathcal{Q}_j \\ Q \cap \Omega \neq \emptyset}} (\sup_Q f - \inf_Q f) \text{Vol}_n(Q) \leq C\varepsilon.$$

Damit konvergieren U_j und L_j gegen den gleichen Grenzwert. Diesen Grenzwert werden wir das Riemann-Integral von f nennen.

- Beim Zugang von Darboux zum Riemannintegral legt man keine Partition von $\mathbb{R}^n$ in Quader zugrunde, sondern betrachtet einfach alle Treppenfunktionen die oberhalb und unterhalb der Funktion liegen.

Definition 4.11 (Riemann-Integral). Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine Funktion. Dann heißen

$$\begin{aligned} \int_{\mathbb{R}^n}^- f &:= \int_{\mathbb{R}^n}^- f(x) dx := \inf \left\{ \int_{\mathbb{R}^n} g : g \in T(\mathbb{R}^n) \text{ und } f \leq g \right\} \\ \int_{\mathbb{R}^n}^+ f &:= \int_{\mathbb{R}^n}^+ f(x) dx := \sup \left\{ \int_{\mathbb{R}^n} g : g \in T(\mathbb{R}^n) \text{ und } f \geq g \right\} \end{aligned}$$

das Ober- und Unterintegral von f (mit der Konvention, dass $\inf \emptyset = +\infty$ und $\sup \emptyset = -\infty$).
 f heißt Riemann-integrierbar, falls das Ober- und Unterintegral übereinstimmen und

$$-\infty < \int_{\mathbb{R}^n} f = \int_{\mathbb{R}^n}^{\bar{}} f < +\infty.$$

In diesem Fall heißt

$$\int_{\mathbb{R}^n} f := \int_{\mathbb{R}^n} f(x) dx = \int_{\mathbb{R}^n}^{\bar{}} f(x) dx$$

das Riemannintegral von f . Wir bezeichnen mit $R(\mathbb{R}^n)$ die Menge aller Riemann-integrierbaren Funktionen auf $\mathbb{R}^n$.

Sei $\Omega \subset \mathbb{R}^n$ eine nichtleere Menge. Eine Funktion $f : \Omega \rightarrow \mathbb{R}$ heißt Riemann-integrierbar auf Ω , falls die Fortsetzung

$$f^* : \mathbb{R}^n \rightarrow \mathbb{R}, \quad f^*(x) := \begin{cases} f(x) & \text{falls } x \in \Omega \\ 0 & \text{sonst.} \end{cases}$$

Riemann-integrierbar ist. Wir definieren in diesem Fall das Riemannintegral von f auf Ω durch

$$\int_{\Omega} f := \int_{\Omega} f(x) dx := \int_{\mathbb{R}^n} f^*(x) dx,$$

und bezeichnen mit $R(\Omega)$ die Menge aller auf Ω Riemann-integrierbaren Funktionen.²⁴

Bemerkung 4.12. (a) Bei dem in Definition 4.11 eingeführten Begriff handelt es sich um die Darboux'sche Konstruktion des Riemann-Integrals. In der Literatur werden solche Funktionen auch Darboux-integrierbar genannt.

(b) Jede Treppenfunktion ist Riemann-integrierbar und das Riemannintegral für Treppenfunktionen stimmt mit dem Integral für Treppenfunktionen überein - das folgt direkt aus der Monotonie des Integrals für Treppenfunktionen und der Definition des Ober- und Unterintegrals.

Satz 4.13. Die Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ sei gleichmässig stetig und besitze einen kompakten Träger. Dann ist f Riemann-integrierbar und $\int_{\mathbb{R}^n} f dx$ kann wie in Beispiel 4.10 beschrieben berechnet werden.

Beweis. Mit der Notation von Beispiel 4.10 gilt

$$L_j \leq \int_{\mathbb{R}^n} f \leq \int_{\mathbb{R}^n}^{\bar{}} f \leq U_j$$

Da $|U_j - L_j| \rightarrow 0$ für $j \rightarrow \infty$ folgt die Behauptung. □

Proposition 4.14 (Riemannsches Integrierbarkeitskriterium). Eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ist genau dann Riemann-Integrierbar, wenn

$$\forall \varepsilon > 0 \exists \underline{g}, \bar{g} \in T(\mathbb{R}^n) : \underline{g} \leq f \leq \bar{g} \text{ und } \int_{\mathbb{R}^n} (\bar{g} - \underline{g}) \leq \varepsilon.$$

²⁴Korrektur 26.6.2020

Beweis. 1. Argument für $\Rightarrow$: Ist f Riemann-integrierbar, dann besitzen die Mengen

$$\left\{ \int_{\mathbb{R}^n} g : g \in T(\mathbb{R}^n) \text{ und } f \leq g \right\} \quad \text{und} \quad \left\{ \int_{\mathbb{R}^n} g : g \in T(\mathbb{R}^n) \text{ und } f \geq g \right\}$$

ein Infimum s_* bzw. ein Supremum s^* in $\mathbb{R}$. Also existieren für $\varepsilon > 0$ Treppenfunktionen $\bar{g}, \underline{g}$ mit

$$\int_{\mathbb{R}^n} \bar{g} < s_* + \frac{\varepsilon}{2} \quad \int_{\mathbb{R}^n} \underline{g} < s^* - \frac{\varepsilon}{2}$$

Da f Riemann-integrierbar ist, gilt $s^* = s_*$ und damit

$$\int_{\mathbb{R}^n} \bar{g} - \underline{g} = \int_{\mathbb{R}^n} \bar{g} - s_* + s^* - \int_{\mathbb{R}^n} \underline{g} < \varepsilon.$$

2. Argument für $\Leftarrow$: Sei $\varepsilon > 0$. Wähle $\underline{g}, \bar{g} \in T(\mathbb{R}^n)$ mit $\int_{\mathbb{R}^n} \bar{g} - \underline{g} < \varepsilon$ und

$$(4.7) \quad \underline{g} \leq f \leq \bar{g}.$$

Es gilt

$$\int_{\mathbb{R}^n} \underline{g} \stackrel{(4.7)}{\leq} \inf \left\{ \int_{\mathbb{R}^n} g : g \in T(\mathbb{R}^n) \text{ und } f \leq g \right\} \leq \int_{\mathbb{R}^n} \bar{g}$$

sowie

$$\int_{\mathbb{R}^n} \underline{g} \leq \sup \left\{ \int_{\mathbb{R}^n} g : g \in T(\mathbb{R}^n) \text{ und } f \geq g \right\} \stackrel{(4.7)}{\leq} \int_{\mathbb{R}^n} \bar{g}$$

Subtraktion der Ungleichungen ergibt

$$-\varepsilon \leq - \int_{\mathbb{R}^n} \bar{g} - \underline{g} dx \leq \overline{\int_{\mathbb{R}^n} f} - \underline{\int_{\mathbb{R}^n} f} \leq \int_{\mathbb{R}^n} \bar{g} - \underline{g} dx \leq \varepsilon.$$

Da $\varepsilon > 0$ beliebig klein gewählt werden kann, folgt

$$\overline{\int_{\mathbb{R}^n} f} = \underline{\int_{\mathbb{R}^n} f}.$$

□

Beispiel 4.15. Betrachte

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) := \begin{cases} 1+x & \text{falls } x \in [-1, 0) \\ x & \text{falls } x \in [0, 1) \\ 0 & \text{sonst.} \end{cases}$$

Für $j \in \mathbb{N}_+$ setze

$$\underline{g}_j(x) := \begin{cases} \frac{k}{j} & \text{falls } x \in [-1 + \frac{k}{j}, -1 + \frac{k+1}{j}) \text{ für ein } k = 0, \frac{1}{j}, \dots, \frac{j-1}{j} \\ \frac{k}{j} & \text{falls } x \in [\frac{k}{j}, \frac{k+1}{j}) \text{ für ein } k = 0, \frac{1}{j}, \dots, \frac{j-1}{j} \\ 0 & \text{sonst,} \end{cases}$$

$$\bar{g}_j(x) := \begin{cases} \frac{k+1}{j} & \text{falls } x \in [-1 + \frac{k}{j}, -1 + \frac{k+1}{j}) \text{ für ein } k = 0, \frac{1}{j}, \dots, \frac{j-1}{j} \\ \frac{k+1}{j} & \text{falls } x \in [\frac{k}{j}, \frac{k+1}{j}) \text{ für ein } k = 0, \frac{1}{j}, \dots, \frac{j-1}{j} \\ 0 & \text{sonst.} \end{cases}$$

Dann gilt $\underline{g}_j \leq f \leq \bar{g}_j$ sowie $\bar{g}_j(x) - \underline{g}_j(x) \leq \begin{cases} \frac{2}{j} & \text{für } x \in [-1, 1) \\ 0 & \text{sonst.} \end{cases}$

$$\Rightarrow \int_{\mathbb{R}} \bar{g}_j - \underline{g}_j \leq \frac{4}{j} \quad \Rightarrow \quad f \in R(\mathbb{R}).$$

Korollar 4.16. Jede Riemann-integrierbare Funktion f ist beschränkt und besitzt einen kompakten Träger.

Beweis. Das folgt direkt aus dem Riemannschen Integrierbarkeitskriterium und der Tatsache, dass Treppenfunktion beschränkt sind und einen kompakten Träger besitzen. $\square$

Lemma 4.17. Seien $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ Funktionen. Dann gilt:

(a) Für alle $\lambda \geq 0$ ist

$$\int_{\mathbb{R}^n}^{\bar{\cdot}} \lambda f(x) dx = \lambda \int_{\mathbb{R}^n}^{\bar{\cdot}} f(x) dx.$$

(b) Falls $\int_{\mathbb{R}^n}^{\bar{\cdot}} f(x) dx > -\infty$ und $\int_{\mathbb{R}^n}^{\bar{\cdot}} g(x) dx > -\infty$ (oder beide $< +\infty$), dann ist

$$\int_{\mathbb{R}^n}^{\bar{\cdot}} (f(x) + g(x)) dx \leq \int_{\mathbb{R}^n}^{\bar{\cdot}} f(x) dx + \int_{\mathbb{R}^n}^{\bar{\cdot}} g(x) dx.$$

(c)

$$\int_{\mathbb{R}^n}^{\bar{\cdot}} (-f(x)) dx = - \int_{\mathbb{R}^n}^{\bar{\cdot}} f(x) dx$$

Aussagen (a) - (e) gelten auch für Riemann-integrierbare Funktionen auf $\Omega \subseteq \mathbb{R}^n$.

Beweis. (a) Für $\lambda = 0$ ist nichts zu zeigen. Sei $\lambda > 0$ und $g \in T(\mathbb{R}^n)$ mit $f \leq g$. Dann gilt $\lambda f \leq \lambda g$ und damit

$$\int_{\mathbb{R}^n}^{\bar{\cdot}} \lambda f dx \leq \int_{\mathbb{R}^n}^{\bar{\cdot}} \lambda g dx = \lambda \int_{\mathbb{R}^n}^{\bar{\cdot}} g dx$$

Übergang zum Infimum über alle $g \in T(\mathbb{R}^n)$ mit $f \leq g$ führt auf

$$\int_{\mathbb{R}^n}^{\bar{\cdot}} \lambda f dx \leq \lambda \int_{\mathbb{R}^n}^{\bar{\cdot}} f dx$$

Durch betrachten von $\tilde{f} = \lambda f$ erhalte wir

$$\int_{\mathbb{R}^n}^{\bar{\cdot}} f dx \leq \int_{\mathbb{R}^n}^{\bar{\cdot}} \frac{1}{\lambda} \tilde{f} dx \leq \frac{1}{\lambda} \int_{\mathbb{R}^n}^{\bar{\cdot}} \lambda f dx,$$

also die umgekehrte Abschätzung und die Behauptung folgt.

- (b) O.B.d.A. seine beide Integrale von $\pm\infty$ verschieden. Seien $g, \tilde{g} \in T(\mathbb{R}^n)$ mit $f \leq g$ und $\tilde{f} \leq \tilde{g}$. Dann gilt $f + \tilde{f} \leq g + \tilde{g}$ und

$$\int_{\mathbb{R}^n}^{\bar{}} f + \tilde{f} dx \leq \int_{\mathbb{R}^n} g + \tilde{g} dx = \int_{\mathbb{R}^n} g + \int_{\mathbb{R}^n} \tilde{g} dx$$

Durch Übergang zum Infimum über alle $g, \tilde{g} \in T(\mathbb{R}^n)$ folgt die Behauptung.

- (c) Dies folgt aus der Definition des Ober- und Unterintegrals und der Implikation

$$f \leq g \quad \Rightarrow \quad -f \geq -g.$$

□

Proposition 4.18 (Eigenschaften des Riemann-Integrals). (a) (Linearität). $R(\mathbb{R}^n)$ ist ein Vektorraum und die Abbildung $R(\mathbb{R}^n) \rightarrow \mathbb{R}, f \mapsto \int_{\mathbb{R}^n} f(x) dx$ ist linear.

- (b) (Monotonie / Positivität). Für alle $f, g \in R(\mathbb{R}^n)$ gilt

$$f \leq g \quad \Rightarrow \quad \int_{\mathbb{R}^n} f \leq \int_{\mathbb{R}^n} g.$$

- (c) (Dreiecksungleichung). Für alle $f \in R(\mathbb{R}^n)$ folgt $|f| \in R(\mathbb{R}^n)$ und

$$\left| \int_{\mathbb{R}^n} f \right| \leq \int_{\mathbb{R}^n} |f|.$$

- (d) (Fubini). Sei $1 \leq m < n$ und $f \in R(\mathbb{R}^n)$. Dann ist die Funktionen

$$\mathbb{R}^m \rightarrow \mathbb{R}, \quad x' \mapsto \int_{\mathbb{R}^{n-m}}^{\bar{}} f(x', x'') dx''$$

Riemann-integrierbar und es gilt

$$\int_{\mathbb{R}^n} f(x) dx = \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^{n-m}}^{\bar{}} f(x', x'') dx'' \right) dx'.$$

Die selbe Aussage gilt auch für das Unterintegral, sowie für vertauschte Rollen von x' und x'' .

Beweis. (a) Seien $f_1, f_2 \in R(\mathbb{R}^n)$, $\lambda \in \mathbb{R}$ und $\varepsilon > 0$. Wähle

$$\underline{g}_i \leq f_i \leq \bar{g}_i \quad (i = 1, 2)$$

mit

$$\sum_{i=1}^2 \int_{\mathbb{R}^n} \bar{g}_i - \underline{g}_i < \varepsilon.$$

Dann gilt

$$\underline{g} := \underline{g}_1 + \lambda \underline{g}_2 \leq f_1 + \lambda f_2 \leq \bar{g}_1 + \lambda \bar{g}_2 =: \bar{g}$$

und

$$\int_{\mathbb{R}^n} \bar{g} - \underline{g} < (1 + \lambda)\varepsilon.$$

Da $\varepsilon > 0$ beliebig ist, folgt die Gleichheit des Ober- und Unterintegrals und damit die Riemann-integrierbarkeit von $f_1 + \lambda f_2$. Die Linearität der Abbildung $f \mapsto \int_{\mathbb{R}^n} f$ ist eine Übung.

- (b) Das folgt aus der Monotonie des Integrals für Treppenfunktionen durch Übergang zu Infimum und Supremum.
- (c) Wir zeigen zunächst $|f| \in R(\mathbb{R}^n)$ mit Hilfe des Riemann Integrierbarkeitskriteriums. Sei $\varepsilon > 0$ wähle $\underline{g}, \bar{g} \in T(\mathbb{R}^n)$ mit

$$\underline{g} \leq f \leq \bar{g}, \quad \int_{\mathbb{R}^n} \bar{g} - \underline{g} < \varepsilon.$$

Wir dürfen annehmen, dass $\underline{g}$ und $\bar{g}$ folgende Gestalt haben:

$$\underline{g} = \sum_{j=1}^M (\inf_{Q_j} f) \mathbf{1}_{Q_j} \text{ sowie } \bar{g} = \sum_{j=1}^M (\sup_{Q_j} f) \mathbf{1}_{Q_j}$$

mit $Q_1, \dots, Q_M \in \mathcal{Q}(\mathbb{R}^n)$. Wir setzen

$$\underline{h} = \sum_{j=1}^M (\inf_{Q_j} |f|) \mathbf{1}_{Q_j} \text{ sowie } \bar{h} = \sum_{j=1}^M (\sup_{Q_j} |f|) \mathbf{1}_{Q_j}.$$

Dann gilt $\underline{h} \leq |f| \leq \bar{h}$.

Es gilt (wegen der inversen Dreiecksungleichung)

$$\forall x, y \in Q_j : |f(x)| - |f(y)| \leq |f(x) - f(y)|,$$

also

$$\sup_{Q_j} |f| - \inf_{Q_j} |f| \leq \sup_{Q_j} f - \inf_{Q_j} f,$$

und damit

$$\begin{aligned} \int_{\mathbb{R}^n} \bar{h} - \underline{h} &= \sum_{j=1}^M (\sup_{Q_j} |f| - \inf_{Q_j} |f|) \text{Vol}_n(Q_j) \leq \sum_{j=1}^M (\sup_{Q_j} f - \inf_{Q_j} f) \text{Vol}_n(Q_j) \\ &= \int_{\mathbb{R}^n} \bar{g} - \underline{g} < \varepsilon. \end{aligned}$$

Also ist $|f|$ Riemann-integrierbar. Die Dreiecksungleichung gilt, denn

$$\begin{aligned} |f| \geq f &\Rightarrow \int_{\mathbb{R}^n} |f| \geq \int_{\mathbb{R}^n} f, \\ |f| \geq -f &\Rightarrow \int_{\mathbb{R}^n} |f| \geq - \int_{\mathbb{R}^n} f. \end{aligned}$$

- (d) Sei $\varepsilon > 0$ wähle $\underline{g}, \bar{g} \in T(\mathbb{R}^n)$ mit

$$\underline{g} \leq f \leq \bar{g}, \quad \int_{\mathbb{R}^n} \bar{g} - \underline{g} < \varepsilon.$$

Nach Proposition 4.7 sind

$$\underline{h}(x') := \int_{\mathbb{R}^{n-m}} \underline{g}(x', x'') dx'', \quad \bar{h}(x') := \int_{\mathbb{R}^{n-m}} \bar{g}(x', x'') dx''$$

Treppenfunktionen und es gilt

$$\underline{h}(x') \leq \int_{\mathbb{R}^{n-m}}^{\bar{\cdot}} f(x', x'') dx'' \leq \bar{h}(x'),$$

sowie

$$\int_{\mathbb{R}^m} \bar{h} - \underline{h} dx' = \int_{\mathbb{R}^n} \bar{g} - \underline{g} < \varepsilon.$$

Die Behauptung folgt nun aus dem Riemannschen Integrierbarkeitskriterium. □

Bemerkung 4.19. • *Iteration von Fubini führt auf*

$$\int_{\mathbb{R}^n} f(x) dx = \int_{\mathbb{R}}^{\bar{\cdot}} \dots \int_{\mathbb{R}}^{\bar{\cdot}} f(x_1, \dots, x_n) dx_1 \dots dx_n,$$

wobei die Reihenfolge auf der rechten Seite beliebig vertauscht werden darf und die Oberintegral durch Unterintegrale ersetzt werden dürfen.

• *Falls*

$$(4.8) \quad \mathbb{R}^m \mapsto \mathbb{R}, \quad x' \mapsto f(x', x''), \quad x'' \in \mathbb{R}^{n-m} \text{ fest gewählt}$$

Riemann-integrierbar ist, dann kann das Oberintegral im Fubini durch das Riemannintegral ersetzt werden.

Wir werden später sehen, dass für stetige Funktionen für alle x'' die Funktion in (4.8) glm. stetig und damit Riemann-integrierbar ist.

Wir werden weiter sehen, dass für Riemann-integrierbare Funktionen die Funktion in (4.8) nur für alle x'' bis auf eine $\mathbb{R}^{n-m}$ -Nullmenge Riemann-integrierbar ist. Was eine Nullmenge ist erklären wir später.

Kontrollfrage 11.1. Betrachte

$$Q_1 := [0, 1) \times [0, 2), \quad Q_2 := [0, 2) \times [1, 3), \quad Q_3 := [1, 1) \times [0, 2).$$

Welche der folgenden Aussagen sind wahr?

(A) Q_1, Q_2, Q_3 sind paarweise disjunkt.

×: Zum Beispiel ist $(-\frac{1}{2}, \frac{3}{2}) \in Q_1 \cap Q_2$.

(B) $Q_3 = \emptyset$.

✓

(C) Q_1, Q_3 sind disjunkt.

✓: Aus $(x, y) \in Q_1 \cap Q_3$ folgt $x < 1$ wegen Q_1 und $x \geq 1$ wegen Q_3 .

(D) $A := Q_1 \cup Q_2 \cup Q_3$, $\text{Vol}_2(A) = \text{Vol}_2(Q_1) + \text{Vol}_2(Q_2) + \text{Vol}_2(Q_3)$.

$$\times: \text{Vol}_2(Q_1) = 2, \text{Vol}_2(Q_2) = 4, \text{Vol}_2(A) = 5.$$

$$(E) \ A := Q_1 \cup Q_2 \cup Q_3, \text{Vol}_2(A) = 5$$

✓

$$(F) \ Q_2 \setminus Q_1 \text{ ist ein Quader.}$$

$\times: Q_2 \setminus Q_1$ kann aber als Vereinigung von zwei Quadern dargestellt werden.

Kontrollfrage 11.2. Sei

$$(4.9) \quad Q_1 := [1, 2) \times [1, 2), \quad Q_2 := [0, 3) \times [0, 3), \quad f := \mathbf{1}_{Q_1} - \frac{1}{9}\mathbf{1}_{Q_2}.$$

Welche der nachfolgenden Aussagen sind richtig?

$$(A) \ f \text{ besitzt eine Darstellung der Form } \sum_{j=1}^N c_j \mathbf{1}_{\tilde{Q}_j} \text{ mit } N = 42 \text{ paarweise disjunkten Quadern } \tilde{Q}_1, \dots, \tilde{Q}_N$$

✓: $f = \frac{8}{9}\mathbf{1}_{Q_1 \cap Q_2} + \mathbf{1}_{Q_1 \setminus Q_2} - \frac{1}{9}\mathbf{1}_{Q_2 \setminus Q_1}$. Nach Lemma 4.3 ist $Q_1 \cap Q_2$ ein Quader, $Q_1 \setminus Q_2$ durch 8 disjunkte Quader und $Q_2 \setminus Q_1$ ebenfalls durch 8 disjunkte Quader darstellbar. Das macht insgesamt 17 disjunkte Quader und mit $42 - 17$ mal der leeren Menge kommt man auch auf 42.

$$(B) \ f \text{ besitzt eine Darstellung der Form } \sum_{j=1}^N c_j \mathbf{1}_{\tilde{Q}_j} \text{ mit } N = 5 \text{ paarweise disjunkten Quadern } \tilde{Q}_1, \dots, \tilde{Q}_N$$

✓: Wir nutzen, dass $Q_1 \setminus Q_2 = \emptyset$. Und $Q_2 \setminus Q_1$ ist durch $[0, 1) \times [0, 3) \cup [2, 3) \times [0, 3) \cup [1, 2) \times [0, 1) \cup [1, 2) \times [2, 3)$ darstellbar. Das macht zusammen mit $Q_1 \cap Q_2$ insgesamt 5 Quader.

$$(C) \ f \text{ besitzt eine Darstellung der Form } \sum_{j=1}^N c_j \mathbf{1}_{\tilde{Q}_j} \text{ mit } N = 4 \text{ paarweise disjunkten Quadern } \tilde{Q}_1, \dots, \tilde{Q}_N$$

$\times: B$ lässt sich nicht mehr verbessern, da man mindestens 4 Quader für $Q_2 \setminus Q_1$ braucht.

Kontrollfrage 11.3. Sei f wie in (4.9) definiert. Welche der nachfolgenden Aussagen sind richtig?

$$(A) \ f(1, 1) = f(2, 2) = \frac{8}{9}.$$

$$\times: f(2, 2) = -\frac{1}{9} \text{ und } f(1, 1) = \frac{8}{9}.$$

$$(B) \ f(3, 3) = 0.$$

✓

$$(C) \ \int_{\mathbb{R}^2} f(x) dx = 0$$

✓

$$(D) \ \int_{\mathbb{R}^2} f(x) dx = \frac{8}{9}$$

×

$$(E) \int_{\mathbb{R}^2} |f(x)| dx = 2$$

$\times$: Ergibt $\frac{16}{9}$.

Kontrollfrage 11.4. Sei f wie in (4.9) definiert. Welche der nachfolgenden Aussagen sind richtig?

$$(A) f(2, \cdot) = -\frac{1}{9}\mathbf{1}_I \text{ mit } I := [0, 3).$$

✓

$$(B) \int_{\mathbb{R}} f(x_1, x_2) dx_2 = \frac{2}{3} \quad \Rightarrow \quad 1 \leq x_1 < 2.$$

✓: Für $2 \leq x_1 < 3$ ist das Integral gleich $-\frac{1}{9}3$. Für $0 \leq x_1 < 1$ ebenso. Und für $x_1 < 0$ oder $x_1 \geq 3$ ist es Null.

$$(C) f(\cdot, 1) = -\frac{1}{9}\mathbf{1}_I \text{ mit } I := [0, 3).$$

$\times$: Gegenbeispiel $f(1, 1) = \frac{8}{9}$.

$$(D) f(\frac{3}{2}, \cdot) = -\frac{1}{9}\mathbf{1}_I + \frac{8}{9}\mathbf{1}_J \text{ mit } I := [0, 1) \cup [2, 3), J := [1, 2).$$

✓

Kontrollfrage 11.5. Betrachte

$$Q_k := [k, k+1), \quad k \in \mathbb{N}.$$

Für $k \in \mathbb{N}$ sei $c_k \in \mathbb{R}$. Setze

$$f := \sum_{k \in \mathbb{N}} c_k \mathbf{1}_{Q_k}.$$

(Die Funktionefolge ist punktweise konvergent). Welche der folgenden Aussagen sind wahr?

$$(A) \{Q_k : k \in \mathbb{Z}\} \text{ sind paarweise disjunkte Quadern in } \mathcal{Q}(\mathbb{R}).$$

✓

$$(B) \text{ Für } c_k := \frac{1}{k+1} \text{ ist } (c_k) \text{ eine Nullfolge, also ist } f \text{ eine beschränkte Treppenfunktion.}$$

$\times$: f ist zwar beschränkt, aber nicht mit endlich vielen Quadern darstellbar.

$$(C) \text{ Sei } c_k := k \text{ für } k \leq 100 \text{ und } c_k := 0 \text{ für } k > 100. \text{ Dann ist } f \text{ eine Treppenfunktion.}$$

✓:

$$(D) \text{ Sei } (c_k) \text{ eine konvergente Folge nicht-negativer Zahlen. Dann ist } f \text{ Riemann-integrierbar.}$$

$\times$: Im Allgemeinen nicht. Siehe Korollar 4.16.

Kontrollfrage 11.6. Welche der nachfolgenden Aussagen sind wahr?

$$(A) f : \mathbb{R}^n \rightarrow \mathbb{R}, f(x) := \max\{0, 1 - \|x\|_2^2\} \text{ ist Riemann-integrierbar.}$$

✓: Siehe Satz 4.13.

(B) $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) := \begin{cases} \frac{1}{x^2} & 0 < x < 1 \\ 0 & \text{sonst.} \end{cases}$ ist Riemann-integrierbar.

$\times$: Siehe Korollar 4.16.

(C) $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) := \exp(-x^2)$ ist Riemann-integrierbar.

$\times$: Siehe Korollar 4.16.

Kontrollfrage 11.7. Sei

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) := \begin{cases} 1 & \text{falls } x \in \mathbb{Q} \cap [0, 1) \\ 0 & \text{sonst.} \end{cases}$$

Welche der folgenden Aussagen sind richtig?

(A) f ist beschränkt und besitzt einen kompakten Träger.

✓: Beschränktheit ist klar. Siehe Fußnote der Bemerkungen 4.2 für die Definition des Trägers.

(B) Für jede Treppenfunktion $\underline{g}$ gilt die Implikation

$$\underline{g} \leq f \quad \Rightarrow \quad \underline{g} \leq 0.$$

✓: Denn jedes nichtleere Intervall, also auch jeder nichtleere Quader, enthält rationale Zahlen.

(C) Für jede Treppenfunktion $\bar{g}$ gilt die Implikation

$$\bar{g} \geq f \quad \Rightarrow \quad \bar{g} \geq \mathbf{1}_{[0,1)}.$$

✓: Jedes nichtleere Intervall enthält rationale Zahlen.

(D) f ist Riemann-integrierbar.

$\times$: Siehe B, C und Proposition 4.14

(E) f ist nicht Riemann-integrierbar.

✓: Siehe D.

Kontrollfrage 11.8. Seien $g, h \in R(\mathbb{R})$ mit

$$\int_{\mathbb{R}} g = 1, \quad \int_{\mathbb{R}} h = 2.$$

(A) Sei $f := g + \frac{1}{2}h$. Dann gilt $f \in R(\mathbb{R})$ und $\int_{\mathbb{R}} f = 3$.

$\times$: $f \in R(\mathbb{R})$ ja, aber $\int f = 2$. (Siehe Proposition 4.18 a.)

(B) Sei $f := g + \frac{1}{2}h$. Dann gilt $f \in R(\mathbb{R})$ und $\int_{\mathbb{R}} f = 2$.

✓: Siehe Proposition 4.18.

(C) Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x_1, x_2) := g(x_1)h(x_2)$. Dann gilt $f \in R(\mathbb{R}^2)$ und $\int_{\mathbb{R}^2} f = 2$.

✓: Für $g, h \geq 0$ folgt $f \in R(\mathbb{R}^2)$ mit Hilfe des Riemannsches Integrierbarkeitskriteriums. Für allgemeine g, h folgt dies durch Zerlegung in den Positiv- und Negativteil. Die Gleichung $\int_{\mathbb{R}^2} f = 2$ folgt nun mit der Fubini-Identität.

(D) Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x_1, x_2) := g(x_1)h(x_2)$. Dann gilt $f \in R(\mathbb{R}^2)$ und $\int_{\mathbb{R}^2} f = 3$.

×: Siehe C.

	22.6.2020
#36	22.6.2020
	22.6.2020
#37	29.6.2020

4.3 Das eindimensionale Riemannintegral

Ziele:

- Hinreichende Kriterien für Riemann-integrierbarkeit.
- Stammfunktion, Hauptsatz der Integral und Differentialrechnung
- Methoden zur Berechnung des Integrals: partielle Integration, Substitution
- Längenintegral einer C^1 -Kurve
- Uneigentliches Riemannintegral

Mit dem Satz von Fubini können Integrale im Mehrdimensionalen auf das eindimensionale Riemannintegral zurückgeführt werden. Wir untersuchen in diesem Abschnitt das eindimensionale Riemannintegral.

In diesem Kapitel gilt immer $a, b \in \mathbb{R}$ mit $a < b$, außer wir geben etwas anderes vor.

Wir betrachten das Riemann-integral für Funktionen auf kompakten Intervallen, d.h. $f : [a, b] \rightarrow \mathbb{R}$. Für $f \in R([a, b])$ benutzen wir folgende Notation:

$$\int_a^b f(x) dx := \int_{[a, b]} f(x) dx, \quad \int_b^a f(x) dx := - \int_{[a, b]} f(x) dx.$$

Um die Integrierbarkeit von $f : [a, b] \rightarrow \mathbb{R}$ zu überprüfen, genügt es Treppenfunktionen mit Quadern in $[a, b]$ (d.h., Intervalle der Form $[s, t] \subseteq [a, b]$) zu betrachten:

Lemma 4.20. Für eine beschränkte Funktion $f : [a, b] \rightarrow \mathbb{R}$ sind äquivalent:

(a) $f \in R([a, b])$

(b) Für alle $\varepsilon > 0$ existiert eine Zerlegung $Z = \{a = t_0 < \dots < t_N = b\} \in \mathcal{Z}([a, b])$, so dass

$$\sum_{k=1}^N \left(\sup_{[t_{k-1}, t_k)} f - \inf_{[t_{k-1}, t_k)} f \right) (t_k - t_{k-1}) < \varepsilon.$$

Weiterhin existiert für $f \in R([a, b])$ und $\varepsilon > 0$ eine Zerlegung $Z = \{a = t_0 < \dots < t_N = b\} \in \mathcal{Z}([a, b])$, deren Feinheit wir beliebig klein wählen können, so dass für beliebige Zwischenstellen $\xi_1, \dots, \xi_N$ mit $\xi_k \in [t_{k-1}, t_k)$ für $k = 1, \dots, N$ die Abschätzung

$$(4.10) \quad \left| \int_a^b f(x) dx - \sum_{k=1}^N f(\xi_k)(t_k - t_{k-1}) \right| < \varepsilon$$

gilt.

Bemerkung 4.21. Eine Summe der Form $\sum_{k=1}^N f(\xi_k)(t_k - t_{k-1})$ nennt man eine *Riemann-Summe* von f . Man kann zeigen: $f : [a, b] \rightarrow \mathbb{R}$ ist genau dann Riemann-integrierbar, wenn jede Folge von Riemann-Summen zu Zerlegungen mit verschwindender Feinheit konvergiert.

Beweis. Die Äquivalenz von (a) und (b) sieht man einfach mit dem Riemannschen Integrierbarkeitskriterium.

Wir zeigen hier nur (4.10). Wähle Z gemäß (b) und setze

$$\underline{g} := \sum_{k=1}^N \left(\inf_{[t_{k-1}, t_k)} f \right) \mathbf{1}_{[t_{k-1}, t_k)}, \quad \bar{g} := \sum_{k=1}^N \left(\sup_{[t_{k-1}, t_k)} f \right) \mathbf{1}_{[t_{k-1}, t_k)}, \quad g := \sum_{k=1}^N f(\xi_k) \mathbf{1}_{[t_{k-1}, t_k)}.$$

Es gilt $\underline{g} \leq g \leq \bar{g}$, sowie $\underline{g} \leq f \leq \bar{g}$ (auf $[a, b]$). Aus der Monotonie des Integrals folgt:

$$\int_{\mathbb{R}} \underline{g} \leq \int_{\mathbb{R}} g \leq \int_{\mathbb{R}} \bar{g}, \quad \int_{\mathbb{R}} \underline{g} \leq \int_{\mathbb{R}} f \leq \int_{\mathbb{R}} \bar{g},$$

also

$$\int_{\mathbb{R}} \underline{g} - \bar{g} dx \leq \int_{\mathbb{R}} f - g dx \leq \int_{\mathbb{R}} \bar{g} - \underline{g} dx.$$

Es folgt $|\int_{\mathbb{R}} f - g dx| < \varepsilon$ und damit (4.10). $\square$

Satz 4.22 (Hinreichende Bedingungen). (a) Ist $f : [a, b] \rightarrow \mathbb{R}$ stetig, dann ist f Riemann-integrierbar in $[a, b]$.

(b) Ist $f : [a, b] \rightarrow \mathbb{R}$ monoton, dann ist f Riemann-integrierbar in $[a, b]$.

Beweis. (a) Der Beweis ist analog zur Argumentation in Beispiel 4.10 und wird dem Leser überlassen.

(b) Wir beweisen nur den Fall, in dem f monoton wachsend ist.

Es gilt $f(a) \leq f \leq f(b)$, also ist f beschränkt.

Sei $a = t_0 < \dots < t_N = b$ eine equidistante Zerlegung von $[a, b]$, d.h. $t_k - t_{k-1} = \frac{b-a}{N}$. Dann gilt

$$\begin{aligned} & \sum_{k=1}^N \left(\sup_{[t_{k-1}, t_k)} f - \inf_{[t_{k-1}, t_k)} f \right) (t_k - t_{k-1}) \\ &= \sum_{k=1}^N \left(f(t_k) - f(t_{k-1}) \right) \frac{b-a}{N} \quad \text{Monotonie} \\ &= \frac{(f(b) - f(a))(b-a)}{N} \quad \text{Teleskopsumme.} \end{aligned}$$

Für große N wird die rechte Seite beliebig klein. Damit folgt die Behauptung folgt aus Lemma 4.20. $\square$

Lemma 4.23. Sei $f : [a, b] \rightarrow \mathbb{R}$ und $c \in (a, b)$.

- (a) Ist $f \in R([a, b])$, dann sind $f|_{[a, c]}$ und $f|_{[c, b]}$ auf $[a, c]$ und $[c, b]$ Riemann-integrierbar.
- (b) Ist $f|_{[a, c]}$ auf $[a, c]$ und $f|_{[c, b]}$ auf $[c, b]$ Riemann-integrierbar, dann ist $f \in R([a, b])$ und es gilt

$$\int_{[a, b]} f(x) dx = \int_{[a, c]} f(x) dx + \int_{[c, b]} f(x) dx.$$

Beweis. Beide Aussagen folgen direkt aus Lemma 4.20. $\square$

Lemma 4.24. (a) Seien $f, g \in R([a, b])$. Dann gilt $fg \in R([a, b])$.

(b) Sei $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ stetig und $f \in R([a, b])$. Dann gilt $\varphi \circ f \in R([a, b])$.

Beweis. Wir beweisen dies im nächsten Kapitel gleich im mehrdimensionalen Fall. $\square$

Definition 4.25 (Stammfunktion, unbestimmtes Integral). Sei $I \subseteq \mathbb{R}$ ein Intervall, $f : I \rightarrow \mathbb{R}$ eine Funktion. Eine Funktion

$$F : I \rightarrow \mathbb{R}$$

heißt Stammfunktion von f , falls F differenzierbar in I ist und $F' = f$ gilt.

Die Menge der Staummfunktionen heißt unbestimmtes Integral von f und wird mit $\int f$ bezeichnet.

Bemerkung 4.26. Ist F eine Stammfunktion von f , dann ist auch $F + C$ mit $C \in \mathbb{R}$ eine Stammfunktion von f . Sind F, G Stammfunktionen einer Funktion f , dann ist die Differenz $F - G$ konstant.

Beispiele 4.27 (Stammfunktionen).

f	Repräsentant von $\int f$	
x^α	$\frac{1}{\alpha+1}x^{\alpha+1}$	für $\alpha \in \mathbb{R} \setminus \{-1\}$
$\frac{1}{x}$	$\ln(x)$	
$\exp(x)$	$\exp(x)$	
$\cos(x)$	$\sin(x)$	
$\frac{1}{\cos^2(x)}$	$\tan(x)$	
$\frac{1}{1+x^2}$	$\arctan(x)$	
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin(x)$	auf $(-1, 1)$
$\frac{1}{\sqrt{x^2-1}}$	$\operatorname{arccosh}(x)$	auf $(1, \infty)$
$\frac{1}{\sqrt{x^2+1}}$	$\operatorname{arsinh}(x)$	

Satz 4.28 (Hauptsatz der Differential- und Integralrechnung). (a) Sei $f \in R([a, b])$ und sei $F : [a, b] \rightarrow \mathbb{R}$ eine Stammfunktion von f . Dann gilt:

$$\int_a^b f(x) dx = F(b) - F(a) =: [F]_a^b.$$

(b) Sei $I \subseteq \mathbb{R}$ ein Intervall, $f : I \rightarrow \mathbb{R}$ stetig, $x_0 \in I$, und sei

$$F : I \rightarrow \mathbb{R}, \quad F(x) := \int_{x_0}^x f(y) dy.$$

Dann ist F eine Stammfunktion von f .

Beweis. (a) Sei $Z = \{a = t_0 < \dots < t_N = b\}$ eine Zerlegung. Dann gilt

$$\begin{aligned} F(b) - F(a) &= \sum_{k=1}^N \left(F(t_k) - F(t_{k-1}) \right) && \text{(Teleskopsumme)} \\ &= \sum_{k=1}^N F'(\xi_k)(t_k - t_{k-1}) && \text{mit } \xi_k \in (t_{k-1}, t_k) \text{ (Mittelwertsatz)} \\ &= \sum_{k=1}^N f(\xi_k)(t_k - t_{k-1}) && \text{mit } \xi_k \in (t_{k-1}, t_k) \text{ (Stammfunktion)} \end{aligned}$$

Nach Lemma 4.20 existiert für beliebige $\varepsilon > 0$ eine Zerlegung Z , so dass

$$\left| \sum_{k=1}^N f(\xi_k)(t_k - t_{k-1}) - \int_a^b f(x) dx \right| < \varepsilon,$$

woraus die Behauptung folgt.

(b) Sei $x \in I$. Wähle ein kompaktes Intervall $J \subseteq I$ mit $x \in J$. f ist auf J gleichmässig stetig. Sei $\varepsilon > 0$. Wähle $\delta > 0$ mit $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$ für alle $x, y \in J$. Für $x, x + h \in J$, $h \neq 0$, $|h| < \delta$ gilt

$$\begin{aligned} \left| \frac{F(x+h) - F(x)}{h} - f(x) \right| &= \left| \frac{1}{h} \int_x^{x+h} f(y) dy - f(x) \right| && \text{(Lemma 4.23)} \\ &= \left| \frac{1}{h} \int_x^{x+h} (f(y) - f(x)) dy \right| && \left(\frac{1}{h} \int_x^{x+h} dy = 1 \right) \\ &\leq \frac{1}{h} \int_x^{x+h} |f(y) - f(x)| dy \\ &\leq \frac{1}{h} \int_x^{x+h} \varepsilon dy \\ &\leq \varepsilon. \end{aligned}$$

Also ist F in x differenzierbar mit $F'(x) = f(x)$.

□

Beispiel 4.29. • *Nicht jede Riemann-integrierbare Funktion besitzt eine Stammfunktion:*

$$f : [-1, 1] \rightarrow \mathbb{R}, \quad f(x) := \begin{cases} 1 & \text{falls } [-1, 0], \\ 2 & \text{falls } (0, 1], \end{cases}$$

ist Riemann-integrierbar. Annahme: F ist eine Stammfunktion von f . Dann ist $F : [-1, 1] \rightarrow \mathbb{R}$ stetig und $F'(x) = f(x)$ für $x \neq 0$. Solche Funktionen besitzen aber immer die Form

$$F(x) = \begin{cases} x + c & \text{falls } [-1, 0], \\ 2x + c & \text{falls } (0, 1], \end{cases}$$

für ein $c \in \mathbb{R}$, und sind im Punkt $x = 0$ nicht differenzierbar!

- Es gibt auch Funktionen, die eine Stammfunktionen besitzen, aber nicht Riemann-integrierbar sind: Betrachte

$$F : [0, 1] \rightarrow \mathbb{R}, \quad F(x) := \begin{cases} x^2 \sin(x^{-2}) & x \in (0, 1], \\ 0 & x = 0. \end{cases}$$

Dann ist F differenzierbar auf $[0, 1]$, aber $f := F'$ ist nicht Riemann-integrierbar (siehe Übung).

29.6.2020

#38

29.6.2020

Korollar 4.30 (Mittelwertsatz für Integrale). Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig. Dann existiert $\eta \in (a, b)$ mit

$$\frac{1}{b-a} \int_a^b f(x) dx = f(\eta).$$

Beweis. Da f stetig ist, existiert eine stetig differenzierbare Stammfunktion F . Aus dem Mittelwertsatz der Differentialrechnung folgt

$$\exists \eta \in (a, b) : \quad \frac{1}{b-a} \int_a^b f(x) dx = \frac{F(b) - F(a)}{b-a} = F'(\eta) = f(\eta).$$

□

Satz 4.31 (Partielle Integration). Seien $F, G : [a, b] \rightarrow \mathbb{R}$ differenzierbar und seien $F', G', F'G, G'F \in R([a, b])$. Dann gilt

$$\int_a^b F'(x)G(x) dx = - \int_a^b G'(x)F(x) dx + [FG]_a^b$$

Beweis. FG ist Stammfunktion für $(FG)' = F'G + G'F$. □

Die partielle Integration ist insbesondere auf $F, G \in C^1([a, b])$ anwendbar, denn dann sind $F', G', F'G, G'F, F'G + G'F$ stetig und damit Riemann-integrierbar.

Satz 4.32 (Substitutionsregel). Sei $I \subseteq \mathbb{R}$ ein Intervall, $f \in C(I)$, $\varphi \in C^1([a, b])$, $\varphi([a, b]) \subseteq I$. Dann gilt

$$\int_a^b f(\varphi(x))\varphi'(x) dx = \int_{\varphi(a)}^{\varphi(b)} f(y) dy.$$

Beweis. Sei $F : I \rightarrow \mathbb{R}$ eine Stammfunktion von f . Dann ist $G : [a, b] \rightarrow \mathbb{R}$, $G(x) := F(\varphi(x))$ eine Stammfunktion der stetigen und damit Riemann-integrierbaren Funktion $G'(x) = f(\varphi(x))\varphi'(x)$. Dann gilt

$$\int_a^b f(\varphi(x))\varphi'(x) dx = G(b) - G(a) = F(\varphi(b)) - F(\varphi(a)) = \int_{\varphi(a)}^{\varphi(b)} f(y) dy.$$

□

Satz 4.33 (Länge einer Kurve). *Sei $f : [a, b] \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Kurve. Dann gilt für deren Länge*

$$L(f) = \int_a^b \|f'(t)\|_2 dt.$$

Beweis (optional für LA-Studierende). Sei $\varepsilon > 0$.

1. Behauptung: Es existiert $\delta > 0$, so dass für alle $(s, t) \subseteq [a, b]$:

$$\left| \|f(t) - f(s)\|_2 - \|f'(t)\|_2(t - s) \right| < \varepsilon(t - s).$$

Das folgt durch komponentenweise Anwendung des Mittelwertsatzes auf $f(t) - f(s)$ und der gleichmäßigen Stetigkeit der Ableitungen $f'_1, \dots, f'_n$ (Vgl. Beweis von Lemma 2.33, Schritt 1). Die Details überlassen wir dem Leser.

2. Sei nun eine Zerlegung $Z = \{a = t_0 < \dots < t_N = b\}$ so gewählt, dass

$$(i) \quad |Z| < \delta$$

$$(ii) \quad \left| L(f) - \sum_{k=1}^N \|f(t_k) - f(t_{k-1})\|_2 \right| < \varepsilon \quad (\text{Vgl. Satz 2.32})$$

$$(iii) \quad \left| \int_a^b \|f'(t)\|_2 dt - \sum_{k=1}^N \|f'(t_k)\|_2(t_k - t_{k-1}) \right| < \varepsilon \quad (\text{Vgl. Lemma 4.20}).$$

Zusammen mit Schritt 1 folgt nun

$$\left| L(f) - \int_a^b \|f'(t)\|_2 dt \right| < (2 + (b - a))\varepsilon,$$

und damit die Behauptung.

□

Korollar 4.34 (Parametrisierung nach Bogenlängen). *Sei $f : [a, b] \rightarrow \mathbb{R}^n$ eine reguläre Kurve mit Länge $\ell > 0$. Dann gilt:*

(a) *Durch*

$$\psi : [a, b] \rightarrow [0, \ell], \quad \psi(t) := \int_a^t \|f'(s)\|_2 ds$$

wird eine streng monoton wachsende, stetig differenzierbare Funktion definiert.

(b) Sei $\varphi : [0, \ell] \rightarrow [a, b]$ die Umkehrfunktion von ψ . Dann wird durch $\tilde{f} : [0, \ell] \rightarrow \mathbb{R}^n$, $\tilde{f}(s) = f(\varphi(s))$ eine reguläre Kurve mit

$$(4.11) \quad \|\tilde{f}'(s)\|_2 = 1 \quad \text{für alle } s \in [0, \ell]$$

definiert und es gilt

$$L(\tilde{f}|_{[0,s]}) = s.$$

Eine Kurve mit der Eigenschaft (4.11) heißt eine nach der Bogenlänge parametrisierte Kurve.

Beweis (optional für LA-Studierende). (a) Das folgt aus $\min_{t \in [a,b]} \|f'(t)\|_2 > 0$.

(b) Für die Umkehrfunktion φ gilt nach Lemma 3.7

$$\varphi'(s) = \frac{1}{\psi'(\varphi(s))} = \frac{1}{\|f'(\varphi(s))\|_2}.$$

Mit der Kettenregel folgt

$$\|\tilde{f}'(s)\|_2 = \left\| \frac{f'(\varphi(s))}{\|f'(\varphi(s))\|_2} \right\|_2 = 1,$$

also

$$\int_0^s \|\tilde{f}'(t)\|_2 dt = s$$

□

Beispiel 4.35. Sei $f : [0, 2\pi] \rightarrow \mathbb{R}^2$, $f(t) := (\sin(t), \cos(t))$. Dann gilt

$$\|f'(t)\|_2 = 1, \quad L(f) = \int_0^{2\pi} \|f'(t)\|_2 dt = 2\pi.$$

29.6.2020

#39

29.6.2020

Ist $(0, 1) \rightarrow \mathbb{R}$, $x \mapsto \frac{1}{\sqrt{x(1-x)}}$ Riemann-integrierbar?

Bisher haben wir nur Integrale auf kompakten Intervallen betrachtet.

Definition 4.36 (Uneigentliches Riemannintegral). (a) Seien $a \in \mathbb{R}$, $b \in \mathbb{R} \cup \{+\infty\}$ mit $b > a$. Eine Funktion $f : [a, b) \rightarrow \mathbb{R}$ heißt *uneigentlich Riemann-integrierbar*, falls für alle $r \in (a, b)$ die Restriktion $f|_{[a,r]}$ Riemann-integrierbar ist und der Grenzwert

$$\lim_{r \rightarrow b} \int_a^r f(x) dx =: \int_a^b f(x) dx$$

existiert. Wir nennen dann $\int_a^b f(x) dx$ das *uneigentliche Riemann-Integral* von f auf $[a, b)$.

(b) Seien $a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R}$ mit $a < b$. Eine Funktion $f : (a, b] \rightarrow \mathbb{R}$ heißt *uneigentlich Riemann-integrierbar*, falls für alle $r \in (a, b)$ die Restriktion $f|_{[r,b]}$ Riemann-integrierbar ist und der Grenzwert

$$\lim_{r \rightarrow a} \int_r^b f(x) dx =: \int_a^b f(x) dx$$

existiert. Wir nennen dann $\int_a^b f(x) dx$ das *uneigentlich Riemann-Integral* von f auf $(a, b]$.

(c) Seien $a, b \in \mathbb{R} \cup \{-\infty, \infty\}$ mit $a < b$. Eine Funktion $f : (a, b) \rightarrow \mathbb{R}$ heißt *uneigentlich Riemann-integrierbar auf $\mathbb{R}$* , falls $c \in \mathbb{R}$ existiert, sodass $f|_{(a,c]}$ und $f|_{[c,b)}$ *uneigentlich Riemann-integrierbar* sind. Wir nennen dann²⁵

$$\int_a^b f(x) dx := \int_a^c f(x) dx + \int_c^b f(x) dx.$$

das *uneigentliche Riemann-integral* auf (a, b) .

Beispiele 4.37. (a) Seien $a, b \in \mathbb{R}$ und sei $f : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar. Dann gilt

$$\lim_{\substack{r \rightarrow a \\ a < r < b}} \int_r^b f(x) dx = \int_{[a,b]} f(x) dx = \lim_{\substack{r \rightarrow b \\ a < r < b}} \int_a^r f(x) dx.$$

(b) $f : [0, \infty) \rightarrow \mathbb{R}$, $f(x) = \exp(-x)$ besitzt keinen kompakten Träger und ist nicht Riemann-integrierbar. Eine Stammunfunktion ist durch $F(x) := -\exp(-x)$ gegeben. Also gilt für $r > 0$,

$$\int_0^r \exp(-x) dx = [-\exp(-x)]_0^r = -\exp(-r) + 1 \rightarrow 1 \text{ für } r \rightarrow \infty.$$

Also ist $\exp(-x)$ *uneigentlich Riemann-integrierbar* auf $[0, \infty)$.

(c) Sei $\alpha \in \mathbb{R}$, $r > 0$:

$$\begin{aligned} & \int_r^1 x^\alpha dx \\ &= \begin{cases} \left[\frac{1}{\alpha+1} x^{\alpha+1} \right]_r^1 = \frac{1}{\alpha+1} (1 - r^{\alpha+1}) \rightarrow \frac{1}{\alpha+1} \text{ für } r \rightarrow 0 & \text{falls } \alpha > -1 \\ [\ln(x)]_r^1 = (0 - \ln(r)) \rightarrow \infty \text{ für } r \rightarrow 0 & \text{falls } \alpha = -1 \\ \left[\frac{1}{\alpha+1} x^{\alpha+1} \right]_r^1 = \frac{1}{\alpha+1} (1 - r^{\alpha+1}) \rightarrow \infty \text{ für } r \rightarrow 0 & \text{falls } \alpha < -1 \end{cases} \end{aligned}$$

Sei $\alpha \in \mathbb{R}$, $r > 0$:

$$\begin{aligned} & \int_1^r x^\alpha dx \\ &= \begin{cases} \left[\frac{1}{\alpha+1} x^{\alpha+1} \right]_1^r = \frac{1}{\alpha+1} (r^{\alpha+1} - 1) \rightarrow \infty \text{ für } r \rightarrow \infty & \text{falls } \alpha > -1 \\ [\ln(x)]_1^r = (\ln(r) - 0) \rightarrow \infty \text{ für } r \rightarrow \infty & \text{falls } \alpha = -1 \\ \left[\frac{1}{\alpha+1} x^{\alpha+1} \right]_1^r = \frac{1}{\alpha+1} (r^{\alpha+1} - 1) \rightarrow -\frac{1}{\alpha+1} \text{ für } r \rightarrow \infty & \text{falls } \alpha < -1 \end{cases} \end{aligned}$$

Fazit:

- x^α ist auf $(0, 1)$ genau dann *uneigentlich Riemann-integrierbar*, wenn $\alpha > -1$.
- x^α ist auf $(1, \infty)$ genau dann *uneigentlich Riemann-integrierbar*, wenn $\alpha < -1$.
- x^α ist auf $(0, \infty)$ nicht *uneigentlich Riemann-integrierbar*.

²⁵Korrektur: 1.7.2020

Kontrollfrage 12.1. Sei $f : [0, 1] \rightarrow \mathbb{R}$ Riemann-integrierbar. Welche der nachfolgenden Aussagen sind dann richtig?

(A) die Funktion $[0, 1] \rightarrow \mathbb{R}, x \mapsto (f(x))^2$ ist Riemann-integrierbar.

✓: Siehe Lemma 4.24 (a).

(B) die Funktion $[0, 1] \rightarrow \mathbb{R}, x \mapsto \frac{1}{f(x)}$ ist Riemann-integrierbar, falls $f(x) > 0$ für alle $x \in [0, 1]$.

✗: Gegenbeispiel $f(x) = 1$, falls $x = 0$, und $f(x) = x$, falls $x \neq 0$, zusammen mit Beispiel 4.37 oder Lemma 4.16.

(C) die Funktion $[0, 1] \rightarrow \mathbb{R}, x \mapsto \frac{1}{f(x)}$ ist Riemann-integrierbar, falls f stetig ist und $f(x) \neq 0$ für alle $x \in [0, 1]$.

✓: Siehe Satz 4.22. (Denn $x \mapsto 1/f(x)$ ist dann auch stetig.)

Kontrollfrage 12.2. Welches Kriterium ist hinreichend für die Riemann-Integrierbarkeit einer Funktion $v : [0, 1] \rightarrow \mathbb{R}$?

(A) v ist monoton,

✓: Siehe Satz 4.22 (b).

(B) $v(x)$ ist positiv für alle $x \in [0, 1]$,

✗: Gegenbeispiel: Unbeschränkte Funktion. (Siehe Korollar 4.16.)

(C) v ist beschränkt,

✗: Gegenbeispiel: Kontrollfrage 11.7.

(D) v ist stetig,

✓: Siehe Satz 4.22 (a).

(E) v ist ein Polynom.

✓: Auch Polynome sind stetig.

Kontrollfrage 12.3. Welche der nachfolgenden Aussagen sind richtig?

(A) Jede Riemann-integrierbare Funktion besitzt eine Stammfunktion.

✗: Siehe Beispiel 4.29.

(B) Es existiert eine Funktion auf $(0, 1)$ mit genau einer Stammfunktion.

✗: Siehe Bemerkung 4.26.

(C) Jede beschränkte Funktion $f : [a, b] \rightarrow \mathbb{R}$ mit beschränkter Stammfunktion, ist Riemann-integrierbar.

✗: Gegenbeispiel siehe Beispiele 4.29 und Übung 13.2.

Kontrollfrage 12.4. Sei $f : [0, 1] \rightarrow \mathbb{R}$ eine beliebige Funktion mit zwei Stammfunktionen F und G . Welche der nachfolgenden Aussagen sind dann richtig?

(A) $F - G = 0$

$\times$: Siehe Bemerkung 4.26.

(B) $F - G$ ist konstant

$\checkmark$: Siehe Bemerkung 4.26.

(C) $F - G$ ist streng monoton

$\times$: Konstant ist nicht streng monoton.

(D) $F' = G'$

$\checkmark$: $F' = G' = f$.

Kontrollfrage 12.5. Sei $F : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \sin^2(x) \cos^5(x)$ und $v = F'$. Welche der nachfolgenden Aussagen ist dann richtig?

(A) v ist nicht Riemann-integrierbar.

$\times$: Doch! Siehe Satz 4.22 (a).

(B) $\int_0^{\pi/2} v = \frac{\pi}{2}$.

$\times$

(C) $\int_0^{\pi/2} v = 0$.

$\checkmark$: $F(\frac{\pi}{2}) = F(0) = 0$ und Satz 4.28.

(D) Es existiert $r \in (0, \pi/2)$ mit $\int_0^r v = (\frac{\sqrt{2}}{2})^7$.

$\checkmark$: Für alle $r \in (0, \frac{\pi}{2})$ ist nach Satz 4.28 dann $\int_0^r v = F(r)$. Und $F(\frac{\pi}{4}) = (\frac{1}{2}\sqrt{2})^2(\frac{1}{2}\sqrt{2})^5$.

Kontrollfrage 12.6. Wir betrachten die Funktion

$$f : [0, 1] \rightarrow \mathbb{R}, \quad f(x) := \begin{cases} 1 & \text{falls } x = \frac{1}{2} \\ 0 & \text{sonst.} \end{cases}$$

Welche der nachfolgenden Aussagen ist wahr?

(A) f ist nicht Riemann-integrierbar, da f nicht monoton ist.

$\times$: f ist nicht monoton, aber trotzdem Riemann-integrierbar.

(B) f ist nicht Riemann-integrierbar, da f nicht stetig ist.

$\times$: f ist nicht stetig, aber trotzdem Riemann-integrierbar.

(C) f ist Riemann-integrierbar, weil $f \geq 0$ und weil es Treppenfunktionen $\bar{g}_k \geq f$ gibt mit $\int \bar{g}_k = \frac{1}{k}$ für alle $k \in \mathbb{N}_+$.

✓: $\bar{g}_k = \mathbf{1}_{[\frac{1}{2}-\frac{1}{2k}, \frac{1}{2}+\frac{1}{2k})}$ und Proposition 4.14.

(D) f ist Riemann-integrierbar, weil $f = \mathbf{1}_{[0, \frac{1}{2}]} - \mathbf{1}_{[0, \frac{1}{2})}$ und die beiden Indikatorfunktionen Riemann-integrierbar sind.

✓: $\mathbf{1}_{[0, \frac{1}{2})}$ ist eine Treppenfunktion und somit Riemann-integrierbar. Aus $\mathbf{1}_{[0, \frac{1}{2})} \leq \mathbf{1}_{[0, \frac{1}{2}]} \leq \mathbf{1}_{[0, \frac{1}{2}+\frac{1}{n})}$ für alle $n \in \mathbb{N}_+$ und Proposition 4.14 folgt die Riemann-Integrierbarkeit von $\mathbf{1}_{[0, \frac{1}{2}]}$.

Kontrollfrage 12.7. Wir betrachten das uneigentliche Integral²⁶

$$\int_0^\infty \sin(x) dx.$$

Welche der nachfolgenden Aussagen trifft zu?

(A) Das uneigentliche Integral hat den Wert 0, denn $\sin$ ist 2π -periodisch und es gilt

$$\int_0^{2\pi} \sin(x) dx = [-\cos(x)]_0^{2\pi} = -\cos(2\pi) + \cos(0) = 0.$$

×

(B) Das uneigentliche Integral hat den Wert 2, denn es gilt

$$\int_0^{(2k+1)\pi} \sin(x) dx = [-\cos(x)]_0^{(2k+1)\pi} = -(-1) - (-1) = 2.$$

für alle $k \in \mathbb{N}$.

×

(C) Das uneigentliche Integral existiert nicht, weil der Grenzwert $\lim_{r \rightarrow \infty} \int_0^r \sin(x) dx$ nicht existiert.

✓: Die Rechnungen von (A) und (B) verdeutlichen, dass der Grenzwert $\lim_{r \rightarrow \infty} \int_0^r \sin$ nicht existieren kann. (Denn $2 \neq 0$.)

Kontrollfrage 12.8. Wir betrachten die Schraubkurve

$$f : [0, 100] \rightarrow \mathbb{R}^3, \quad f(t) := \frac{1}{\sqrt{2}}(\sin(t), \cos(t), t).$$

Welche der nachfolgenden Aussagen ist richtig?

(A) f ist nach Bogenlänge parametrisiert

$$\checkmark: \frac{\cos^2(t) + \sin^2(t) + 1^2}{\sqrt{2}^2} = 1 \text{ für alle } t \in [0, 100].$$

²⁶Korrektur 6.7.2020

$$(B) \quad L(f) = 100$$

$$\checkmark: \int_0^{100} \sqrt{1} dt = 100.$$

$$(C) \quad L(f|_{[0,\ell]}) \neq \ell \text{ für ein } 0 < \ell < 100.$$

$$\times: \int_0^l 1 dl = l \text{ für alle } l.$$

Hier endet der prüfungsrelevante Stoff für die Studierenden im Lehramt.

4.4 Das n -dimensionale Volumen

- Jordan-messbare Menge, Volumen einer Menge
- Integral von Funktionen auf Jordan-messbaren Mengen
- Satz von Fubini und Cavallieri

Motivation:

- Sei $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$. Häufig möchte man f nur über eine Teilmenge $\Omega \subseteq U$ integrieren. Wir setzen

$$\int_{\Omega} f(x) dx = \int_{\mathbb{R}^n} \mathbf{1}_{\Omega}(x) f(x) dx.$$

wobei wir auf der rechten Seite $f(x) = 0$ für $x \in \mathbb{R}^n \setminus U$ setzen.

Der Integrand auf der rechten Seite, d.h. die Funktion $x \mapsto \mathbf{1}_{\Omega}(x)f(x)$ ist offensichtlich nur von $f|_{\Omega}$ abhängig.

Damit das Integral auf der rechten Seite definiert ist, muss das Produkt $\mathbf{1}_{\Omega}f$ Riemann-integrierbar sein.

Insbesondere möchten wir die konstante Funktion $f = 1$ betrachten. Dafür muss $\mathbf{1}_{\Omega}$ Riemann-integrierbar sein.

- Wir möchten das Volumen einer Menge $\Omega \subseteq \mathbb{R}^n$ definieren. Ist Ω ein Quader (oder die endliche Vereinigung von Quadern), dann gilt $\text{Vol}_n(\Omega) = \int_{\mathbb{R}^n} \mathbf{1}_{\Omega}(x) dx$.

Definition 4.38. Sei $\Omega \subseteq \mathbb{R}^n$.

(a) $\overline{\text{Vol}}_n(\Omega) := \int_{\mathbb{R}^n} \mathbf{1}_{\Omega}$ heißt äußerer Jordaninhalt von Ω .

(b) $\underline{\text{Vol}}_n(\Omega) := \int_{\mathbb{R}^n} \mathbf{1}_{\Omega}$ heißt innerer Jordaninhalt von Ω .

(c) Ω heißt Jordan-messbar, wenn $\mathbf{1}_{\Omega} \in R(\mathbb{R}^n)$. In diesem Fall heißt

$$\text{Vol}_n(\Omega) := \int_{\mathbb{R}^n} \mathbf{1}_{\Omega}.$$

das n -dimensionale Volumen (oder den Jordaninhalt oder das Maß) von Ω .

(d) Eine Jordan-messbare Menge $\Omega \subseteq \mathbb{R}^n$ mit $\text{Vol}_n(\Omega) = 0$ heißt Jordan-Nullmenge.

Beispiel 4.39. $A := \mathbb{Q}^n \cap [0, 1]^n$ ist nicht Jordan-messbar, da $\overline{\text{Vol}}_n(A) = \text{Vol}_n([0, 1]^n) = 1$ und $\underline{\text{Vol}}_n(A) = 0$.

Lemma 4.40. $\Omega \subset \mathbb{R}^n$ Jordan-messbar $\Rightarrow \Omega$ ist beschränkt.

Beweis. $\mathbf{1}_\Omega$ Riemann-integrierbar $\Rightarrow \mathbf{1}_\Omega$ hat kompakten Träger. $\square$

Bemerkung 4.41 (Geometrische Interpretation). *Wir formalisieren hier die Motivation zum Volumenbegriff, welche wir am Beginn des Kapitels ausgeführt haben. Definiere*

$$\mathcal{I}(\mathbb{R}^n) := \left\{ A = \bigcup_{k=1}^N Q_k : Q_1, \dots, Q_N \in \mathcal{Q}(\mathbb{R}^n) \text{ paarweise disjunkt, } N \in \mathbb{N} \right\}$$

(a) Für $A = \bigcup_{k=1}^N Q_k \in \mathcal{I}(\mathbb{R}^n)$ ist $\mathbf{1}_A$ eine Treppenfunktion. Damit ist A Jordan-messbar.

Es gilt

$$\text{Vol}_n(A) = \sum_{k=1}^N \text{Vol}_n(Q_k).$$

(b) $\mathcal{I}(\mathbb{R}^n)$ ist abgeschlossen unter Vereinigung, Durchschnitts- und Differenzenbildung, d.h. sind A, B in $\mathcal{I}(\mathbb{R}^n)$, dann auch $A \cup B$, $A \cap B$ und $A \setminus B$. (Vgl. Lemma 4.3)

(c) Sei $\Omega \subset \mathbb{R}^n$ beschränkt. Es gilt

$$\begin{aligned} \overline{\text{Vol}}_n(\Omega) &= \inf \left\{ \text{Vol}_n(A) : A \in \mathcal{I}(\mathbb{R}^n) \text{ mit } \Omega \subseteq A \right\}, \\ \underline{\text{Vol}}_n(\Omega) &= \sup \left\{ \text{Vol}_n(A) : A \in \mathcal{I}(\mathbb{R}^n) \text{ mit } A \subseteq \Omega \right\}. \end{aligned}$$

(d) Ω ist genau dann Jordan-messbar, wenn $\overline{\text{Vol}}_n(\Omega) = \underline{\text{Vol}}_n(\Omega)$ gilt.

(e) Sei $\Omega \subseteq \mathbb{R}^n$ beschränkt. Dann gilt

$$\overline{\text{Vol}}_n(\Omega) - \underline{\text{Vol}}_n(\Omega) = \overline{\text{Vol}}_n(\partial\Omega).$$

Lemma 4.42. (a) $\Omega \subseteq \mathbb{R}^n$ ist genau dann Jordan-messbar, wenn für alle $\varepsilon > 0$ Mengen $\overline{A}, \underline{A} \in \mathcal{I}(\mathbb{R}^n)$ mit $\underline{A} \subseteq \Omega \subseteq \overline{A}$ und

$$0 \leq \text{Vol}_n(\overline{A} \setminus \underline{A}) = \text{Vol}_n(\overline{A}) - \text{Vol}_n(\underline{A}) < \varepsilon.$$

(b) $\Omega \subseteq \mathbb{R}^n$ ist genau dann eine Jordan-Nullmenge, wenn es für alle $\varepsilon > 0$ ein $A \in \mathcal{I}(\mathbb{R}^n)$ gibt mit $\Omega \subseteq A$ und $\text{Vol}_n(A) < \varepsilon$.

(c) $\Omega \subseteq \mathbb{R}^n$ ist genau dann Jordan-messbar, wenn $\partial\Omega$ eine Jordan-Nullmenge ist.

(d) Sind $A, B \subseteq \mathbb{R}^n$ Jordan-messbar, dann auch $A \cup B$, $A \cap B$ und $A \setminus B$.

(e) Seien $f, g \in R(\mathbb{R}^n)$. Dann ist $\{(x, y) \in \mathbb{R}^{n+1} : g(x) \leq y < f(x)\}$ Jordan-messbar.²⁷

²⁷Korrektur: 15.7.2020. Achtung hier muss $y < f(x)$ und nicht $y \leq f(x)$ stehen, denn die Menge $\{(x, y) : g(x) \leq y \leq f(x)\}$ kann unbeschränkt und damit nicht Jordan-messbar sein.

Beweis. Aussagen (a) - (c) folgen aus den Beobachtungen von Bemerkung 4.41 und der Definition des Jordaninhaltes. (d) folgt aus den Identitäten

$$\mathbf{1}_{A \cap B} = \mathbf{1}_A \mathbf{1}_B, \quad \mathbf{1}_{A \setminus B} = \max\{\mathbf{1}_A - \mathbf{1}_B, 0\}, \quad \mathbf{1}_{A \cup B} = \mathbf{1}_B + \mathbf{1}_{A \setminus B}$$

und der Abgeschlossenheit des Raumes der Riemann-integrierbaren Funktionen unter Produktbildung.

Aussage (e) ist für Treppenfunktionen f, g klar, da dann die Menge eine endliche Vereinigung Quadern ist. Mit dem Riemannschen Integrierbarkeitskriterium kann die allgemeine Situation auf die Situation mit Treppenfunktionen zurückgeführt werden. □

Lemma 4.43. *Seien $f, g \in R(\mathbb{R}^n)$. Dann gilt $fg \in R(\mathbb{R}^n)$.*

Beweis. **1.** Spezialfall: $f, g \geq 0$. Sei $\varepsilon > 0$. Seien $\bar{f}, \underline{f}, \bar{g}, \underline{g} \in T(\mathbb{R}^n)$ mit

$$0 \leq \underline{f} \leq f \leq \bar{f}, \quad 0 \leq \underline{g} \leq g \leq \bar{g}, \quad \int_{\mathbb{R}^n} \bar{f} - \underline{f} + \int_{\mathbb{R}^n} \bar{g} - \underline{g} \leq \varepsilon.$$

Weiter seien $\bar{f}, \underline{f}, \bar{g}, \underline{g}$ durch $C := \sup_{\mathbb{R}^n} f + \sup_{\mathbb{R}^n} g$ beschränkt. Dann gilt

$$\underline{f} \underline{g} \leq fg \leq \bar{f} \bar{g},$$

sowie

$$\int_{\mathbb{R}^n} (\bar{f} \bar{g}) - (\underline{f} \underline{g}) = \int_{\mathbb{R}^n} (\bar{f} - \underline{f}) \bar{g} + \int_{\mathbb{R}^n} \underline{f} (\bar{g} - \underline{g}) \leq C\varepsilon.$$

Damit ist fg nach dem Riemannschen Integrierbarkeitskriterium Riemann-integrierbar.

2. Sei $f^+ := \max\{f, 0\}$ der Positivteil und $f^- := (-f)^+$ der Negativteil²⁸. Dann sind f^+, f^- Riemann-integrierbar. Selbiges gilt für g^+ und g^- . Weiter gilt

$$fg = (f^+ - f^-)(g^+ - g^-) = f^+g^+ - f^-g^+ - f^+g^- + f^-g^-,$$

Die Produkte auf der rechten Seite sind Produkte nicht-negativer, Riemann-integrierbarer Funktionen und nach Schritt 1 Riemann-integrierbar. □

Definition 4.44. *Sei $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ eine Funktion. Sei $\Omega \subseteq U$ Jordan-messbar. Wir nennen f Riemann-integrierbar auf Ω , falls $\mathbf{1}_\Omega f \in R(\mathbb{R}^n)$. In diesem Fall schreiben wir*

$$\int_{\Omega} f(x) dx := \int_{\mathbb{R}^n} \mathbf{1}_\Omega(x) f(x) dx.$$

Bemerkung 4.45. *Durch Definition 4.11 und Definition 4.44 wird der gleiche Integrierbarkeitsbegriff für Funktionen auf Jordan-messbaren Mengen definiert: Sei $\Omega \subseteq \mathbb{R}^n$ Jordan-messbar. Dann ist jede Funktion $f \in R(\Omega)$ (im Sinne von Definition 4.11) Riemann-integrierbar auf Ω (im Sinne von Definition 4.44). Andererseits: Ist f Riemann-integrierbar auf Ω im Sinne von Definition 4.44, dann gilt $f|_\Omega \in R(\Omega)$. Aus Gründen der Bequemlichkeit schreiben wir auch dann einfach $f \in R(\Omega)$.*

²⁸Korrektur: 1.7.20

Satz 4.46. Sei $K \subseteq \mathbb{R}^n$ kompakt. Dann ist K Jordan-messbar genau dann, wenn $C(K) \subseteq R(K)$.

Beweis. 1. Argument für $\Leftarrow$. Da die konstante Funktion $x \mapsto 1$ stetig auf K ist, gilt $\mathbf{1}_K \in R(\mathbb{R}^n)$ und damit ist K Jordan-messbar.

2. Zwischenbeobachtung: Sei $Q \in \mathcal{Q}(\mathbb{R}^n)$ und sei $f : Q \rightarrow \mathbb{R}$ gleichmäßig stetig. Das Argument aus Beispiel 4.10 zeigt nun, dass f auf Q Riemann-integrierbar ist.

3. Argument für $\Rightarrow$. Wir setzen $f(x) = 0$ für $x \in \mathbb{R}^n \setminus K$ und $C := \max_{x \in K} |f(x)|$. Sei $\varepsilon > 0$. Wähle $\bar{A}, \underline{A} \in \mathcal{I}(\mathbb{R}^n)$ mit

$$\underline{A} \subseteq K \subseteq \bar{A}, \quad \text{Vol}_n(\bar{A}) - \text{Vol}_n(\underline{A}) < \varepsilon.$$

Setze

$$\begin{aligned} \bar{f} : \mathbb{R}^n &\rightarrow \mathbb{R}, & \bar{f}(x) &:= \begin{cases} C & x \in \bar{A} \setminus \underline{A} \\ f(x) & x \in \underline{A} \\ 0 & \text{sonst.} \end{cases} \\ \underline{f} : \mathbb{R}^n &\rightarrow \mathbb{R}, & \underline{f}(x) &:= \begin{cases} -C & x \in \bar{A} \setminus \underline{A} \\ f(x) & x \in \underline{A} \\ 0 & \text{sonst.} \end{cases} \end{aligned}$$

Dann gilt

- $\underline{f} \leq \mathbf{1}_K f \leq \bar{f}$,
- $\underline{f}, \bar{f} \in R(\mathbb{R}^n)$, denn diese Funktionen können als endliche Summen von Funktionen von der Form aus Schritt 2 dargestellt werden.
- Wegen der Monotonie gilt $\int_{\mathbb{R}^n} \underline{f} \leq \int_{\mathbb{R}^n} \mathbf{1}_K f \leq \int_{\mathbb{R}^n} \bar{f}$.
- Nach Konstruktion gilt $\int_{\mathbb{R}^n} \bar{f} - \underline{f} \leq \int_{\mathbb{R}^n} 2C \mathbf{1}_{\bar{A} \setminus \underline{A}} \leq 2C\varepsilon$.

Da $\varepsilon > 0$ beliebig klein gewählt werden kann, müssen das Ober- und Unterintegral von $\mathbf{1}_K f$ übereinstimmen. Also ist $\mathbf{1}_K f$ Riemann-integrierbar und damit ist f auf K Riemann-integrierbar. $\square$

Lemma 4.47. Sei $\Omega \subseteq \mathbb{R}^n$ Jordan-messbar und $f \in R(\Omega)$ (und damit beschränkt). Dann gilt

$$\left| \int_{\Omega} f \, dx \right| \leq \int_{\Omega} |f| \, dx \leq \sup_{\Omega} |f| \text{Vol}_n(\Omega).$$

Beweis. Das folgt aus der Dreiecksungleichung für Integrale und der Monotonität des Integrals. $\square$

Lemma 4.48. Seien $A, B \subseteq \mathbb{R}^n$ disjunkt und Jordan-messbar, sei $N \subseteq \mathbb{R}^n$ eine Jordan-Nullmenge. Sei $f \in R(A \cup B \cup N)$. Dann gilt

$$\begin{aligned} (a) \quad \int_{A \cup B} f(x) \, dx &= \int_A f(x) \, dx + \int_B f(x) \, dx \\ (b) \quad \int_N f(x) \, dx &= 0 \\ (c) \quad \int_A f(x) \, dx &= \int_{A \setminus N} f(x) \, dx \end{aligned}$$

Beweis. (a) Das folgt aus der Identität $\mathbf{1}_{A \cup B} = \mathbf{1}_A + \mathbf{1}_B$ und der Linearität des Integrals.

(b) Das folgt aus $\text{Vol}_n(N) = 0$ und Lemma 4.47.

(c) Das folgt aus $A = (A \setminus N) \cup N$ und Teil (a) und (b). □

29.6.2020

#41

29.6.2020

Definition 4.49. Sei $\Omega \subseteq \mathbb{R}^n$ eine Menge und $c \in \mathbb{R}$. Wir nennen die Menge

$$\Omega_c := \{(x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1} : (x_1, \dots, x_{n-1}, c) \in \Omega\}$$

den c -Schnitt von Ω nach der n -ten Koordinate. Die c -Schnitte nach anderen Koordinaten definieren wir genauso.

Beispiel 4.50. (a) Sei $\Omega := [0, 1] \times [0, 2] \times [-1, 2]$. Dann gilt

$$\Omega_c = \begin{cases} [0, 1] \times [0, 2] & \text{für } c \in [-1, 2], \\ \emptyset & \text{sonst.} \end{cases}$$

Ω_c ist Jordan-messbar für alle $c \in \mathbb{R}$. Es gilt

$$1 \cdot 2 \cdot 3 = \text{Vol}_3(\Omega) = \text{Vol}_2([0, 1] \times [0, 2]) \text{Vol}_1([-1, 2]) = \int_{\mathbb{R}} \text{Vol}(\Omega_c) \, dc.$$

(b) Sei $\Omega := [0, 1] \times [0, 1] \times [-1, 0] \cup [0, 2] \times [0, 2] \times (0, 1]$. Dann gilt

$$\Omega_c = \begin{cases} [0, 1] \times [0, 1] & \text{für } c \in [-1, 0], \\ [0, 2] \times [0, 2] & \text{für } c \in (0, 1], \\ \emptyset & \text{sonst.} \end{cases}$$

Ω_c ist Jordan-messbar für alle $c \in \mathbb{R}$. Es gilt

$$1 + 4 = \text{Vol}_3(\Omega) = \int_{\mathbb{R}} \text{Vol}(\Omega_c) \, dc.$$

(c) Sei $\Omega := (\mathbb{Q} \cap [0, 1]) \times \{1\}$. Dann ist Ω Jordan-messbar, aber $\Omega_1 = \mathbb{Q} \cap [0, 1]$ ist nicht Jordan-messbar. Das Beispiel zeigt, dass c -Schnitte von Jordan-messbaren Mengen nicht Jordan-messbar sein müssen. Man kann zeigen, dass die Menge aller $c \in \mathbb{R}$ für die Ω_c nicht Jordan-messbar ist, eine Jordan-Nullmenge ist.

Satz 4.51 (Satz von Fubini). Sei $\Omega \subseteq \mathbb{R}^n$ Jordan-messbar und $f \in R(\mathbb{R}^n)$. Für $c \in \mathbb{R}$ bezeichnen wir mit Ω_c den c -Schnitt von Ω nach der n ten Koordinate. Für jedes $c \in \mathbb{R}$ sei $\Omega_c \subseteq \mathbb{R}^{n-1}$ Jordan-messbar und $f(\cdot, c) \in R(\Omega_c)$. Dann gilt

$$\int_{\Omega} f(x) dx = \int_{\mathbb{R}} \left(\int_{\Omega_c} f(x', c) dx' \right) dc$$

Beweis. Es gilt mit Proposition 4.18 (d) und wegen $\mathbf{1}_{\Omega}(x', c) = \mathbf{1}_{\Omega_c}(x')$,

$$\begin{aligned} \int_{\Omega} f &= \int_{\mathbb{R}^n} f(x) \mathbf{1}_{\Omega}(x) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}^{n-1}} f(x', c) \mathbf{1}_{\Omega}(x', c) dx' \right) dc \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}^{n-1}} f(x', c) \mathbf{1}_{\Omega_c}(x') dx' \right) dx \\ &= \int_{\mathbb{R}} \left(\int_{\Omega_c} f(x', c) dx' \right) dc. \end{aligned}$$

□

Mit der konstanten Funktion $f = 1$ folgt:

Korollar 4.52 (Prinzip von Cavallieri). In der Situation von Satz 4.51 wähle $c^- \leq c^+$ mit $\Omega \subseteq \mathbb{R}^{n-1} \times [c^-, c^+]$. Dann gilt

$$\text{Vol}_n(\Omega) = \int_{c^-}^{c^+} \text{Vol}_{n-1}(\Omega_c) dc.$$

Beispiel 4.53. Volumen der Einheitskugel $K := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\} \subset \mathbb{R}^2$. Für $c \in \mathbb{R}$ gilt

$$K_c := \{x \in \mathbb{R} : x^2 + c^2 \leq 1\} = \begin{cases} [-\sqrt{1-c^2}, \sqrt{1-c^2}] & \text{falls } |c| \leq 1 \\ \emptyset & \text{sonst} \end{cases}$$

Also

$$\begin{aligned} \text{Vol}_2(K) &= 2 \int_{-1}^1 \sqrt{1-c^2} dc \stackrel{(*)}{=} 2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{1-\sin^2(t)} \cos(t) dt \\ &= 2 \int_{-\pi/2}^{\pi/2} \cos(t) \cos(t) dt \\ &= \int_{-\pi/2}^{\pi/2} \cos^2(t) dt + \int_{-\pi/2}^{\pi/2} \cos(t) \sin'(t) dt \\ &\stackrel{(**)}{=} \int_{-\pi/2}^{\pi/2} \cos^2(t) + \sin^2(t) dt + [\cos(t) \sin(t)]_{-\pi/2}^{\pi/2} \\ &= \pi. \end{aligned}$$

(*) folgt mit der Substitution $c = \sin(t)$ und (**) mit partieller Integration.

4.5 Transformationssatz

Wie verhält sich das Volumen eines Quaders $Q = [a, b]^n \in \mathcal{Q}(\mathbb{R}^n)$ unter einer affinen Transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $\Phi(x) := Ax + z$?

- Translationsinvarianz: Sei $A = I$, also $T(Q) = Q + z$. Es gilt $\text{Vol}_n(Q + z) = \text{Vol}_n([a + z, b + z]^n) = \text{Vol}_n(Q)$.
- Dilatation: Sei $z = 0$ und A eine Diagonalmatrix mit $A_{jj} = \lambda_j$. Dann gilt

$$\begin{aligned} \text{Vol}_n(T(Q)) = \text{Vol}_n([Aa, Ab]^n) &= |\lambda_1|(b_1 - a_1) \cdots |\lambda_n|(b_n - a_n) = |\lambda_1 \cdots \lambda_n| \text{Vol}_n(Q) \\ &= |\det A| \text{Vol}_n(Q). \end{aligned}$$

- Orthogonale Transformation: Sei $z = 0$ und $A \in O(n)$. Man kann zeigen, dass

$$\text{Vol}_n(T(Q)) = \text{Vol}_n(Q).$$

- Allgemeiner Fall: Mit Hilfe der Singulärwertzerlegung kann jede Matrix $A \in \mathbb{R}^{n \times n}$ als Produkt der Form $A = RDR'$ geschrieben werden mit einer Diagonalmatrix D und $R, R' \in O(n)$. Also gilt

$$\text{Vol}_n(T(Q)) = |\det D| \text{Vol}_n(Q) = |\det A| \text{Vol}_n(Q) = |\det J_T| \text{Vol}_n(Q).$$

Sei $\Phi : U \rightarrow V$ ein Diffeomorphismus zwischen offenen Mengen $U, V \subseteq \mathbb{R}^n$. Sei $x_0 \in U$ und $Q_h := h[-\frac{1}{2}, \frac{1}{2}]^n$ ein kleiner Quader mit $x_0 + Q_h \subseteq U$. Aus der Taylorentwicklung von Φ um x_0 folgt, dass $\Phi(x) = \Phi(x_0) + J_\Phi(x_0)(x - x_0) + |x - x_0|^\rho(x - x_0)$, und damit gilt approximativ (für kleines h)

$$\Phi(x_0 + Q_h) \approx T(x_0 + Q_h) = \Phi(x_0) + J_\Phi(x_0)Q_h, \quad T(x) := \Phi(x_0) + J_\Phi(x_0)(x - x_0)$$

sowie

$$\text{Vol}_n(\Phi(x_0 + Q_h)) \approx \text{Vol}_n(T(x_0 + Q_h)) = \text{Vol}_n(J_\Phi(x_0)Q_h) = |\det J_\Phi(x_0)| \text{Vol}_n(Q_h).$$

Satz 4.54 (Transformationssatz). *Seien $U, V \subseteq \mathbb{R}^n$ offene Mengen und $\Phi : U \rightarrow V$ ein Diffeomorphismus. Sei $f : V \rightarrow \mathbb{R}$ Riemann-integrierbar auf V . Dann ist die Funktion $U \rightarrow \mathbb{R}$, $x \mapsto (f \circ \Phi)(x) |\det J_\Phi(x)|$ Riemann-integrierbar auf U und es gilt*

$$\int_U f(\Phi(x)) |\det J_\Phi(x)| dx = \int_V f(y) dy.$$

Wir beweisen diesen Satz nicht, denn Sie werden einen Beweis dieses Satzes für das modernere Lebesgue-Integral in der Vorlesung “Maß- und Integrationstheorie” kennenlernen.

Beispiel 4.55 (Polarkoordinaten). *Seien*

$$\Phi : (0, \infty) \times (0, 2\pi) \rightarrow \mathbb{R}^2 \setminus ([0, \infty) \times \{0\}), \quad \Phi(r, \alpha) := (r \cos(\alpha), r \sin(\alpha))$$

Dann gilt

$$J_\Phi(r, \alpha) = \begin{pmatrix} \cos(\alpha) & -r \sin(\alpha) \\ \sin(\alpha) & r \cos(\alpha) \end{pmatrix}, \quad \det J_\Phi(r, \alpha) = r.$$

Aus dem Transformationssatz folgt nun

$$\begin{aligned}\int_{\mathbb{R}^2} f(x) dx &= \int_{\mathbb{R}^2 \setminus ([0, \infty) \times \{0\})} f(x) dx && ([0, \infty) \times \{0\} \text{ ist Jordan-Nullmenge}) \\ &= \int_{(0, \infty) \times (0, 2\pi)} f(\Phi(r, \alpha)) r d(r, \alpha) && (\text{Transformationssatz}) \\ &= \int_0^\infty \left(\int_0^{2\pi} f(\Phi(r, \alpha)) d\alpha \right) r dr,\end{aligned}$$

falls die Integrale existieren. Insbesondere,

$$\begin{aligned}\int_{\mathbb{R}^2} \exp(-\|x\|^2) dx &= 2\pi \int_0^\infty \exp(-r^2) r dr = 2\pi \left[-\frac{1}{2} \exp(-r^2) \right]_0^\infty \\ &= \pi.\end{aligned}$$

Achtung: Das Integral $\int_0^\infty dr$ ist ein uneigentliches Integral, d.h. man muss zunächst $\int_c^C \exp(-r^2) r dr$ betrachten und dann zu den Grenzwerten $c \rightarrow 0$ und $C \rightarrow \infty$ übergehen.

#43

29.6.2020

6.7.2020

5 Lokale Extrema unter Nebenbedingungen

Sei $\Omega \subseteq \mathbb{R}^n$, $f : \Omega \rightarrow \mathbb{R}$ und $F : \Omega \rightarrow \mathbb{R}$. Wir betrachten das Minimierungsproblem

$$(5.1) \quad \min F(x), \quad \text{wobei } x \in \Omega \text{ und } f(x) = 0.$$

Die Bedingung “ $f(x) = 0$ ” heißt Nebenbedingung (genauer: Gleichungsnebenbedingung).

Ein $x_* \in \Omega$ heißt lokale Lösung von (5.1), falls $f(x_*) = 0$ und falls es eine offene Kugel B um x_* gibt, so dass

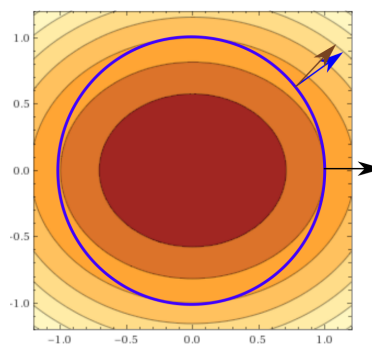
$$F(x_*) \leq F(x) \quad \text{für alle } x \in \Omega \cap B \text{ mit } f(x) = 0.$$

Beispiel 5.1. Wir möchten den Ausdruck

$$x_1^2 + \frac{3}{2}x_2^2 \text{ minimieren unter der Nebenbedingung } \sqrt{x_1^2 + x_2^2} = 1.$$

Wir setzen

$$\begin{aligned}F : \mathbb{R}^2 &\rightarrow \mathbb{R}, & F(x) &:= x_1^2 + \frac{3}{2}x_2^2 \\ f : \mathbb{R}^2 &\rightarrow \mathbb{R}, & f(x) &:= \|x\|_2^2 - 1\end{aligned}$$



Satz 5.2 (Notwendige Bedingung, Lagrange-Multiplikator). Sei $\Omega \subseteq \mathbb{R}^n$ offen, seien $f, F \in C^1(\Omega)$, sei x_* eine lokale Lösung zu (5.1) und es gelte $\nabla f(x_*) \neq 0$. Dann gibt es ein $\lambda \in \mathbb{R}$ mit

$$\nabla F(x_*) = \lambda \nabla f(x_*).$$

Die Zahl λ heißt Lagrange Multiplikator.

Beweis. Wegen $\nabla f(x_*) \neq 0$ könne wir O.B.d.A. annehmen, dass $\frac{\partial}{\partial x_n} f(x_*) \neq 0$ (andernfalls nennen wir die Koordinaten um).

Wir schreiben

$$x_* = (x'_*, x_{*n}), \quad x'_* = (x_{*1}, \dots, x_{*n-1}).$$

Wir lösen mit dem Satz über Implizite Funktionen (Satz 3.14) die Gleichung $f(x', x_n) = 0$ nach x_n auf: Es existiert eine offene Umgebungen U' von x'_* und $g \in C^1(U')$, so dass

$$x_* = (x'_*, g(x'_*)), \quad f(x', g(x')) = 0 \text{ und } (x', g(x')) \in \Omega \quad \text{für alle } x' \in U'.$$

Betrachte $h : U' \rightarrow \mathbb{R}$, $h(x') := f(x', g(x'))$. Diese Funktion ist konstant 0. Mit der Kettenregel folgt nun für $i = 1, \dots, n-1$:

$$\begin{aligned} 0 &= \frac{\partial h}{\partial x_i}(x'_*) = \frac{\partial f}{\partial x_i}(x'_*, g(x'_*)) + \frac{\partial g}{\partial x_i}(x'_*) \frac{\partial f}{\partial x_n}(x'_*, g(x'_*)) \\ (5.2) \quad &= \frac{\partial f}{\partial x_i}(x_*) + \frac{\partial g}{\partial x_i}(x'_*) \frac{\partial f}{\partial x_n}(x_*). \end{aligned}$$

Betrachte nun

$$\tilde{F} : U' \rightarrow \mathbb{R}, \quad \tilde{F}(x') := F(x', g(x')).$$

Dann ist x'_* ein lokales Minimum von $\tilde{F}$. Folglich gilt

$$\begin{aligned} 0 &= \frac{\partial \tilde{F}}{\partial x_i}(x'_*) \quad (\text{Kettenregel}) \\ &= \frac{\partial F}{\partial x_i}(x'_*, g(x'_*)) + \frac{\partial g}{\partial x_i}(x'_*) \frac{\partial F}{\partial x_n}(x'_*, g(x'_*)) \\ (5.3) \quad &= \frac{\partial F}{\partial x_i}(x_*) + \frac{\partial g}{\partial x_i}(x'_*) \frac{\partial F}{\partial x_n}(x_*), \end{aligned}$$

Multiplikation von (5.2) und (5.3) mit $\frac{\partial F}{\partial x_n}(x_*)$ und $\frac{\partial f}{\partial x_n}(x_*)$ und Gleichsetzen führt auf

$$\frac{\partial F}{\partial x_i}(x_*) \frac{\partial f}{\partial x_n}(x_*) = \frac{\partial f}{\partial x_i}(x_*) \frac{\partial F}{\partial x_n}(x_*)$$

Mit

$$\lambda := \frac{\frac{\partial F}{\partial x_n}(x_*)}{\frac{\partial f}{\partial x_n}(x_*)}$$

folgt die Behauptung □

Beispiel 5.3. 1. Sei $F : \mathbb{R}^n \rightarrow \mathbb{R}$ die quadratische Abbildung $F(x) := x \cdot Ax$ mit symmetrischer Matrix $A \in \mathbb{R}^{n \times n}$. Wir betrachten das Minimierungsproblem

$$\min F(x) \quad \text{wobei } x \in \mathbb{R}^n \text{ mit } \|x\|_2 = 1.$$

Wir setzen daher $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f(x) := \|x\|_2^2 - 1$. Dann gilt

$$\nabla F(x) = 2Ax, \quad \nabla f(x) = 2x.$$

Da F stetig ist und die Menge $S := \{x \in \mathbb{R}^n : f(x) = 0\}$ eine kompakte Menge ist, existiert ein Minimum $x_* \in S$ mit $F(x_*) \leq F(x)$ für alle $x \in S$.

Für alle $x \in S$ gilt $\nabla f(x) = 2x \neq 0$.

Nach Satz (5.2) existiert also $\lambda \in \mathbb{R}$ mit $\nabla F(x_*) = \lambda \nabla f(x_*)$ und damit

$$Ax_* = \lambda x_*.$$

Wir folgern, dass (λ, x_*) ein Eigenpaar ist!

2. Allerdings ist nicht jedes Eigenpaar eine Lösung des Minimierungsproblems. Das zeigt, dass die Existenz eines Lagrange-Multiplikators nur eine notwendige Bedingung ist! Besitz A z.B. Eigenwerte $0 < \lambda_1 < \dots < \lambda_n$ mit normierten Eigenvektoren $v_1, \dots, v_n \in S$, dann gilt

$$F(v_i) = v_i \cdot Av_i = \lambda_i v_i \cdot v_i = \lambda_i,$$

und damit ist genau v_1 die Lösung des Minimierungsproblems. Man nennt dieses Verfahren zur Bestimmung des kleinsten Eigenwertes das Rayleigh-Ritz-Prinzip.

6.7.2020

#44

6.7.2020

6 Kurvenintegral

Motivation:

- Ein konstantes Kraftfeld $F \in \mathbb{R}^3$ leistet entlang einer Strecke von $x_0 \in \mathbb{R}^3$ nach $x_1 \in \mathbb{R}^3$ die Arbeit

$$\|F\|_2 \|x_1 - x_0\|_2 \cos \alpha = F \cdot (x_1 - x_0), \quad \alpha := \angle(F, x_1 - x_0).$$

- Betrachte einen Streckenzug durch die Punkte $x_0, x_1, \dots, x_N \in \mathbb{R}^3$ und ein Kraftfeld, dass auf den Streckenabschnitten $[[x_{i-1}, x_i]]$ konstant den Wert $F_i \in \mathbb{R}^3$ annimmt. Dann gilt für die Arbeit

$$\sum_{i=1}^N F_i \cdot (x_i - x_{i-1}).$$

- Wir beschreiben den Streckenzug als Raumkurve $\gamma : [a, b] \rightarrow \mathbb{R}^3$ mit $\gamma(t_i) = x_i$ für eine Zerlegung $a = t_0 < t_1 < \dots < t_N = b$. Mit $F(x_i) := F_i$ gilt

$$\sum_{i=1}^N \left(F(\gamma(t_i)) \cdot \frac{\gamma(t_i) - \gamma(t_{i-1})}{t_i - t_{i-1}} \right) (t_i - t_{i-1})$$

- Ist nun $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ stetig und $\gamma : [a, b] \rightarrow \mathbb{R}^3$ eine stetig differenzierbare Kurve, so ist die längs der Kurve f verrichtete Arbeit:

$$\int_a^b F(\gamma(t)) \cdot \gamma'(t) dt.$$

Definition 6.1 (Kurvenintegral). Sei $\Omega \subseteq \mathbb{R}^n$, sei $F : \Omega \rightarrow \mathbb{R}^n$ stetig und $\gamma : [a, b] \rightarrow \mathbb{R}^n$ stetig differenzierbar. Dann heißt

$$\int_{\gamma} F(x) \cdot dx := \int_a^b F(\gamma(t)) \cdot \gamma'(t) dt$$

das Kurvenintegral von F von $\gamma(a)$ nach $\gamma(b)$ entlang γ .

Ist $\gamma : [a, b] \rightarrow \mathbb{R}^n$ stückweise stetig differenzierbar mit $\gamma|_{[t_{i-1}, t_i]} \in C^1([t_{i-1}, t_i])$ für eine Zerlegung $a = t_0 < \dots < t_N = b$, dann setzen wir

$$\int_{\gamma} F(x) \cdot dx = \sum_{i=1}^N \int_{\gamma|_{[t_{i-1}, t_i]}} F \cdot d\gamma.$$

Bemerkung 6.2. • (Invarianz unter Umparametrisierung). Sei $\Omega \subseteq \mathbb{R}^n$, sei $F : \Omega \rightarrow \mathbb{R}^n$ stetig und $\gamma : [a, b] \rightarrow \mathbb{R}^n$ stetig differenzierbar. Sei $\varphi : [\alpha, \beta] \rightarrow [a, b]$ ein C^1 -Diffeomorphismus. Dann gilt mit $\tilde{\gamma} := \gamma \circ \varphi$ wegen der Substitutionsregel:

$$\begin{aligned} \int_{\alpha}^{\beta} F(\tilde{\gamma}(s)) \cdot \tilde{\gamma}'(s) ds &= \int_{\alpha}^{\beta} F(\gamma(\varphi(s))) \cdot \gamma'(\varphi(s)) \varphi'(s) ds \\ &= \begin{cases} \int_a^b F(\gamma(t)) \cdot \gamma'(t) dt & \text{falls } \varphi' > 0, \\ -\int_a^b F(\gamma(t)) \cdot \gamma'(t) dt & \text{falls } \varphi' < 0, \end{cases} \end{aligned}$$

d.h. es gilt im Fall einer orientierungserhaltenden Umparametrisierung ($\varphi' > 0$):

$$\int_{\gamma} F \cdot d\gamma = \int_{\tilde{\gamma}} F \cdot d\tilde{\gamma}.$$

- Zwei Kurven $\gamma : [a, b] \rightarrow \mathbb{R}^n$ und $\tilde{\gamma} : [\alpha, \beta] \rightarrow \mathbb{R}^n$ heißen äquivalent bis auf orientierungserhaltende Umparametrisierungen, falls ein Diffeomorphismus $\varphi : [\alpha, \beta] \rightarrow [a, b]$ existiert mit $\tilde{\gamma} = \gamma \circ \varphi$ und $\varphi' > 0$. Das Kurvenintegral ist für äquivalente Kurven identisch.
- Weiter kann man zeigen, dass das Kurvenintegral einer stückweise stetig differenzierbaren Kurve unabhängig von der Zerlegung ist.
- In der Literatur findet man häufig folgende Schreibweise:

$$\int_{\Gamma} F(x) \cdot dx,$$

wobei $\Gamma \subseteq \mathbb{R}^n$ eine “geometrische” Kurve bezeichnet. Gemeint ist hierbei das Kurvenintegral $\int_{\gamma} F(x) \cdot dx$, wobei γ eine parametrisierte Kurve mit Spur Γ bezeichnet. Aus dem Kontext

muss aber klar sein, wie Γ (modulo Äquivalenz) zu parametrisieren ist. Beispielsweise gibt es für $\Gamma = \{x \in \mathbb{R}^2 : \|x\|_2 = 1\}$ (Einheitssphäre) unendlich viele parametrisierte Kurven mit Spur Γ :

$$\gamma_k : [0, 1] \rightarrow \mathbb{R}^2, \quad \gamma_k(t) := (\cos(2\pi kt), \sin(2\pi kt)), \quad k \in \mathbb{N}_+.$$

Die Kurve γ_k durchläuft k -mal die Einheitssphäre. Für verschiedene k sind die Kurven nicht äquivalent und es gilt

$$\int_{\gamma_k} F(x) \cdot dx = k \int_{\gamma_1} F(x) \cdot dx.$$

Lemma 6.3. Sei $\Omega \subseteq \mathbb{R}^n$ offen, $f \in C^1(\Omega)$, $\gamma : [a, b] \rightarrow \mathbb{R}^n$ eine stückweise stetig differenzierbare Kurve mit Spur in Ω . Dann gilt

$$\int_{\gamma} \nabla f(x) \cdot dx = f(\gamma(b)) - f(\gamma(a)).$$

Beweis. **1.** Sei γ stetig differenzierbar. Dann gilt nach dem Hauptsatz der Differential- und Integralrechnung und der Kettenregel:

$$f(\gamma(b)) - f(\gamma(a)) = \int_a^b (f \circ \gamma)'(t) dt = \int_a^b \nabla f(\gamma(t)) \cdot \gamma'(t) dt = \int_{\gamma} \nabla f(x) \cdot dx.$$

2. Sei γ stückweise stetig differenzierbar. Dann existiert eine Zerlegung $a = t_0 < \dots < t_N = b$, so dass $\gamma|_{[t_{i-1}, t_i]} \in C^1([t_{i-1}, t_i])$. Nach Schritt 1 folgt

$$f(\gamma(t_i)) - f(\gamma(t_{i-1})) = \int_{\gamma|_{[t_{i-1}, t_i]}} \nabla f(x) \cdot dx.$$

Ein Teleskopsummenargument ergibt

$$f(\gamma(b)) - f(\gamma(a)) = \sum_{i=1}^N \int_{\gamma|_{[t_{i-1}, t_i]}} \nabla f(x) \cdot dx = \int_{\gamma} \nabla f(x) \cdot dx.$$

□

#45

6.7.2020

6.7.2020

Definition 6.4 (Konservatives Vektorfeld). Sei $\Omega \subseteq \mathbb{R}^n$ offen und wegzusammenhängend. Ein stetiges Vektorfeld $F : \Omega \rightarrow \mathbb{R}^n$ heißt konservativ auf Ω , falls für alle Punkte $x_a, x_b \in \Omega$ das Kurvenintegral von F von x_a nach x_b "nicht vom Weg abhängig" ist, d.h.: Sind $\gamma : [a, b] \rightarrow \Omega$ und $\tilde{\gamma} : [c, d] \rightarrow \Omega$ stückweise stetig differenzierbaren Kurven von x_a nach x_b , dann gilt

$$\int_{\gamma} F(x) \cdot dx = \int_{\tilde{\gamma}} F(x) \cdot dx.$$

Satz 6.5. Sei $\Omega \subseteq \mathbb{R}^n$ offen und wegzusammenhängend und $F : \Omega \rightarrow \mathbb{R}^n$ ein stetiges Vektorfeld. Dann gilt:

$$F \text{ ist Gradientenfeld} \quad \Leftrightarrow \quad F \text{ ist konservativ.}$$

Beweis. **1.** Die Richtung $\Rightarrow$ folgt direkt aus Lemma 6.3.

2. Behauptung: Seien $x_a, x_b \in \Omega$. Dann existiert eine stückweise stetig differenzierbare Kurve $\gamma : [a, b] \rightarrow \Omega$ mit $\gamma(a) = x_a$ und $\gamma(b) = x_b$. Da Ω wegzusammenhängend ist existiert eine stetige Kurve. Da Ω offen und die Spur der stetigen Kurve kompakt ist, können wir die stetige Kurve durch einen Streckenzug in Ω approximieren. Wir überlassen die Details dem interessierten Leser.

3. Argument für $\Leftarrow$: Wir fixieren einen Punkt $x_a \in \Omega$ und definieren $f : \Omega \rightarrow \mathbb{R}$ durch

$$f(x) := \int_{\gamma} F \cdot dx,$$

wobei γ eine stückweise stetig differenzierbare Kurve von x_a nach x ist. Eine solche Kurve existiert nach Schritt 2. Der Wert $f(x)$ ist nicht von der gewählten Kurve abhängig, da F konservativ ist.

Wir zeigen nun, dass f partiell differenzierbar ist und $\nabla f(x) = F(x)$ gilt.

Sei $x \in \Omega$ und $1 \leq i \leq n$. Wähle $r > 0$ mit $B(x; 2r) \subseteq \Omega$. Wähle eine stückweise stetig differenzierbare Kurve $\gamma : [a, b] \rightarrow \Omega$ von x_a nach x und setze

$$\tilde{\gamma} : [a, b + r] \rightarrow \Omega, \quad \tilde{\gamma}(t) := \begin{cases} \gamma(t) & \text{falls } t \in [a, b] \\ \gamma(b) + e_i(t - b) & \text{sonst.} \end{cases}$$

Dann ist $\tilde{\gamma}$ eine stückweise stetig-differenzierbare Kurve und insbesondere auf $[b, b + r]$ differenzierbar.

Nach Konstruktion gilt

$$x + he_i = \tilde{\gamma}(b + h) \quad \text{für alle } 0 \leq h \leq r.$$

Also

$$\begin{aligned} f(x + he_i) - f(x) &= \int_{\tilde{\gamma}} F \cdot dx - \int_{\gamma} F \cdot dx \\ &= \int_b^{b+h} F(\tilde{\gamma}(t)) \cdot \tilde{\gamma}'(t) dt, \end{aligned}$$

wobei die Definition des Kurvenintegrals und Lemma 4.23 verwendet wurden. Auf $(b, b + h)$ gilt $\tilde{\gamma}'(t) = e_i$. Mit dem Mittelwertsatz für Integrale folgt nun

$$\frac{f(x + he_i) - f(x)}{h} = \frac{1}{h} \int_b^{b+h} F_i(\tilde{\gamma}(t)) dt = F_i(\tilde{\gamma}(\tau_h))$$

für ein $\tau_h \in [b, b + h]$. Für $h \rightarrow 0$ gilt $\tilde{\gamma}(\tau_h) \rightarrow x$ und da F stetig ist, folgt

$$\lim_{\substack{h \rightarrow 0 \\ h > 0}} \frac{f(x + he_i) - f(x)}{h} = F_i(x),$$

Mit dem gleichen Argument zeigt man, dass .Also gilt $\frac{\partial}{\partial x_i} f(x) = F_i(x)$.

□

Lemma 6.6. Sei $\Omega \subseteq \mathbb{R}^n$, $f \in C^2(\Omega)$ und $F := \nabla f$. Dann gilt

$$\partial_i F_j(x) = \partial_j F_i(x)$$

für alle $x \in \Omega$ und $1 \leq i, j \leq n$.

Beweis. Das folgt aus dem Lemma von Schwartz. □

Satz 6.7. Sei $\Omega \subseteq \mathbb{R}^n$ offen und sternförmig, d.h. es existiert $x_* \in \Omega$, so dass für alle $x \in \Omega$ die Verbindungsstrecke $[[x, x_*]]$ in Ω enthalten ist. Sei $F : \Omega \rightarrow \mathbb{R}^n$ stetig differenzierbar mit

$$(6.1) \quad \partial_i F_j(x) = \partial_j F_i(x)$$

für alle $x \in \Omega$ und $1 \leq i, j \leq n$, dann ist F ein Gradientenfeld auf Ω .

Bemerkung 6.8. Man nennt (6.1) Integrabilitätsbedingung.

Beweis. O.B.d.A. sei $x_* = 0$ (andernfalls verschiebe Ω und F). Die Funktion

$$g : [0, 1] \times \Omega \rightarrow \mathbb{R}, \quad g(t, x) := F(tx) \cdot x$$

ist wohldefiniert (da $[[0, x]] \subseteq \Omega$) und stetig differenzierbar. Setze

$$f(x) := \int_0^1 g(t, x) dt.$$

(das ist gerade das Kurvenintegral von F entlang der Kurve $\gamma_x : [0, 1] \rightarrow \mathbb{R}^n$, $\gamma_x(t) := tx$.)

Es gilt

$$\begin{aligned} \frac{\partial}{\partial x_j} g(t, x) &= \sum_{i=1}^n \frac{\partial}{\partial x_j} (F_i(tx) x_i) \\ &= F_j(tx) + \sum_{i=1}^n t \frac{\partial F_i}{\partial x_j}(tx) x_i \\ &\stackrel{(6.1)}{=} F_j(tx) + \sum_{i=1}^n t \frac{\partial F_j}{\partial x_i}(tx) x_i \\ &= F_j(tx) + t \nabla F_j(tx) \cdot x \\ &= h_{j,x}(t) + t h'_{j,x}(t) \quad \text{mit} \quad h_{j,x}(t) := F_j(tx). \end{aligned}$$

Wir dürfen partielles Ableiten und Integrieren hier vertauschen (für die Begründung siehe Lemma 6.10 unten) und erhalten so

$$\begin{aligned} \frac{\partial}{\partial x_j} f(x) &= \int_0^1 \frac{\partial}{\partial x_j} g(t, x) dt \\ &= \int_0^1 h_{j,x}(t) + t h'_{j,x}(t) dt = \int_0^1 h_{j,x}(t) - h_{j,x}(t) dt + [t h_{j,x}(t)]_0^1 \\ &= h_{j,x}(1) = F_j(x). \end{aligned}$$

□

Beispiel 6.9. Wir betrachten

$$\Omega := \mathbb{R}^2 \setminus \{0\}, \quad F : \Omega \rightarrow \mathbb{R}^2, F(x) = \frac{x^\perp}{\|x\|_2^2}, \quad x^\perp := (-x_2, x_1).$$

Dann gilt

- Ω ist nicht sternförmig.
- $\partial_1 F_2 = \partial_2 F_1$ auf Ω . Also existiert für jede sternförmige Teilmenge $B \subseteq \Omega$ ein Potential $f \in C^2(B)$ mit

$$\nabla f(x) = F(x) \quad \text{für alle } x \in B.$$

- Andererseits gilt für die Kurve $\gamma : [0, 2\pi] \rightarrow \Omega$, $\gamma(t) := (\cos(t), \sin(t))$, deren Spur der Einheitskreis S^1 ist:

$$\int_\gamma F \cdot dx = \int_0^{2\pi} F(\gamma(t)) \cdot \gamma'(t) dt = \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} \cdot \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} dt = 2\pi.$$

Da $\gamma(0) = \gamma(2\pi)$ (d.h. die Kurve ist geschlossen), kann F nicht konservativ sein und es kann kein Potential $f \in C^2(\Omega)$ existieren mit

$$\nabla f(x) = F(x) \quad \text{für alle } x \in \Omega.$$

Lemma 6.10 (Vertauschen von Integral und partieller Ableitung). Sei $\Omega \subseteq \mathbb{R}^n$ offen, $g : [a, b] \times \Omega \rightarrow \mathbb{R}$, $(t, x) \mapsto g(t, x)$ stetig. Sei g partiell nach x_j differenzierbar und die partielle Ableitung

$$\frac{\partial}{\partial x_j} g : [a, b] \times \Omega \rightarrow \mathbb{R}$$

sei stetig. Dann ist

$$G : \Omega \rightarrow \mathbb{R}, \quad G(x) := \int_a^b g(t, x) dt$$

partiell nach x_j differenzierbar und es gilt

$$\frac{\partial}{\partial x_j} G(x) = \int_a^b \frac{\partial}{\partial x_j} g(t, x) dt.$$

Beweis. Sei $x \in \Omega$ und $U \subseteq \Omega$ eine kompakte Umgebung von x . Für $x, x + he_j \in U$ mit $h \neq 0$ gilt

$$\frac{1}{h}(G(x + he_j) - G(x)) = \int_a^b \frac{g(t, x + he_j) - g(t, x)}{h} dt$$

Nach dem Mittelwertsatz existiert eine Zwischenstelle $\tau_{t,h} \in (-h, h)$ mit

$$\begin{aligned} \frac{g(t, x + he_j) - g(t, x)}{h} &= \frac{\partial}{\partial x_j} g(t, x + \tau_{t,h} e_j) \\ &= \frac{\partial}{\partial x_j} g(t, x) + r_h(t, x) \end{aligned}$$

mit Rest

$$r_h(t, x) := \frac{\partial}{\partial x_j} g(t, x + \tau_{t,h} e_j) - \frac{\partial}{\partial x_j} g(t, x).$$

Auf der kompakten Menge $[a, b] \times U$ ist $\frac{\partial}{\partial x_j} g$ gleichmäßig stetig. Für alle $\varepsilon > 0$ existiert also ein $\delta > 0$, so dass für alle $0 < |h| < \delta$ (und damit $\|\tau_{t,h} e_j\|_\infty < \delta$):

$$|r_h(t, x)| = \left| \frac{\partial}{\partial x_j} g(t, x + \tau_{t,h} e_j) - \frac{\partial}{\partial x_j} g(t, x) \right| < \varepsilon.$$

Es folgt für $0 < |h| < \delta$:

$$\begin{aligned} \left| \frac{1}{h} (G(x + h e_j) - G(x)) - \int_a^b \frac{\partial}{\partial x_j} g(t, x) dt \right| &\leq \int_a^b \underbrace{\left| \frac{g(t, x + h e_j) - g(t, x)}{h} - \frac{\partial}{\partial x_j} g(t, x) \right|}_{< \varepsilon} dt \\ &\leq (b - a) \varepsilon. \end{aligned}$$

Da $\varepsilon > 0$ beliebig klein gewählt werden kann, folgt

$$\lim_{h \rightarrow 0} \frac{1}{h} (G(x + h e_j) - G(x)) = \int_a^b \frac{\partial}{\partial x_j} g(t, x) dt.$$

□

#46

6.7.2020
13.7.2020

7 Topologie und Kompaktheit in metrischen Räumen

7.1 Topologie

Definition 7.1. Sei (M, d) ein metrischer Raum. Wir nennen das System der offenen Mengen

$$\mathcal{T} := \{O \subseteq M : O \text{ ist offen} \}$$

die durch d induzierte Topologie auf M .

Satz 7.2. Sei (M, d) ein metrischer Raum und $\mathcal{T}$ die durch d induzierte Topologie. Dann gilt:

- (a) $\emptyset, M \in \mathcal{T}$,
- (b) Seien $O_1, O_2 \in \mathcal{T}$. Dann gilt $O_1 \cap O_2 \in \mathcal{T}$,
- (c) Sei $(O_i)_{i \in I}$ eine (möglicherweise überabzählbare) Familie von Mengen in $\mathcal{T}$. Dann gilt $\bigcup_{i \in I} O_i \in \mathcal{T}$.

Beweis. Siehe AN10 (Analysis - Grundlegende Konzepte). □

Bemerkung 7.3 (Topologischer Raum). • Wir werden in diesem Abschnitt das Mengensystem $\mathcal{T}$ untersuchen und sehen, dass man mit $\mathcal{T}$ die Konvergenz von Folgen und die Stetigkeit von Funktionen untersuchen kann ohne die Metrik d zu benutzen.

- Sei M eine Menge. Ein Mengensystem $\mathcal{T} \subseteq \mathcal{P}(M)$ heißt Topologie auf M , falls (a) - (c) erfüllt sind. Man nennt dann das Paar $(M, \mathcal{T})$ einen topologischen Raum. In einem topologischen Raum heißen Mengen $O \subseteq M$ offen, falls $O \in \mathcal{T}$. Wir nennen $A \subseteq M$ abgeschlossen, falls $M \setminus A \in \mathcal{T}$. Eine Menge $D \subseteq M$ heißt Umgebung eines Punktes $x \in M$, falls $O \in \mathcal{T}$ existiert mit $x \in O$ und $O \subseteq D$.
- Nicht jede Topologie wird durch eine Metrik induziert. Dennoch kann man in topologischen Räumen den Begriff der Stetigkeit und Konvergenz definieren.
- Eine wichtiger topologischer Raum (der durch keine Metrik induziert ist) ist beispielsweise der Raum der Testfunktionen $C_c^\infty(\mathbb{R}^n)$ aus der Distributionentheorie nach L. Schwarz. Auf diesem topologischen Raum ist die lineare Abbildung

$$\delta : C_c^\infty(\mathbb{R}^n) \rightarrow \mathbb{R}, \quad \varphi \mapsto \varphi(0)$$

stetig. Stetige Funktionale auf $C_c^\infty(\mathbb{R}^n)$ heißen Distributionen. Die “Delta-Distribution” δ spielt beispielsweise in der Theorie partieller Differentialgleichungen, in der Signalanalyse (Elektrotechnik) und in Modellen der Physik (Feldgleichungen) eine wichtige Rolle.

Korollar 7.4. Sei (M, d) ein metrischer Raum. Dann gilt

- (a) $\emptyset, M$ sind abgeschlossen,
- (b) Seien $A_1, \dots, A_N$ endlich viele abgeschlossene Mengen, dann ist auch $\bigcup_{i=1}^N A_i$ abgeschlossen.
- (c) Sei $(A_i)_{i \in I}$ eine (möglicherweise überabzählbare) Familie abgeschlossener Mengen, dann ist auch $\bigcap_{i \in I} A_i$ abgeschlossen.

Beweis. Das folgt aus den Morgan’schen Regeln für Mengen (siehe AN10, Satz 1.7). □

Proposition 7.5 (Inneres und Abschluss). Sei (M, d) ein metrischer Raum und $D \subseteq M$.

- (a) (Inneres von D). Es gilt $\text{Int}(D) = \bigcup \{O : O \text{ offen und } O \subseteq D\}$
- (b) (Abschluss von D). Es gilt $\overline{D} = \bigcap \{A : A \text{ abgeschlossen und } D \subseteq A\}$.

Beweis. Vgl. AN10 Satz 6.4 □

Proposition 7.6 (Grenzwert und Häufungspunkt von Folgen). Sei (M, d) ein metrischer Raum. Sei $(a_n)_{n \in \mathbb{N}} \subseteq M$ eine Folge. Dann gilt: $a \in M$

- (a) ist ein Häufungspunkt der Folge, falls jede Umgebung von a unendlich viele Folgenglieder enthält.
- (b) ist Grenzwert der Folge, falls jede Umgebung von a alle bis auf endliche viele Folgenglieder enthält.

Beweis. Siehe AN10. □

Proposition 7.7 (Stetigkeit). Seien (M_1, d_1) und (M_2, d_2) metrische Räume, $f : M_1 \rightarrow M_2$. Dann sind äquivalent:

(a) f ist stetig

(b) Für alle offenen $O \subseteq M_2$ ist $f^{-1}(O) \subseteq M_1$ offen.

Beweis. Übung

□

13.7.2020

#47

13.7.2020

7.2 Kompaktheit

Eine offene Überdeckung eines metrischen Raumes (M, d) ist eine Familie offener Mengen $(O_i)_{i \in I} \subseteq \mathcal{T}$ mit $\bigcup_{i \in I} O_i \supseteq M$.

Satz 7.8. Sei (M, d) ein metrischer Raum. Dann sind äquivalent:

(a) (Folgenkompaktheit). Jede Folge in M besitzt eine konvergente Teilfolge

(b) (Überdeckungskompaktheit). Jede offene Überdeckung $(O_i)_{i \in I}$ von M besitzt eine “endliche Teilüberdeckung”, d.h. eine endliche Teilfamilie $J = \{i_1, \dots, i_N\} \subseteq I$ mit $\bigcup_{j \in J} O_j \supseteq M$.

(c) M ist vollständig und zu jedem $\varepsilon > 0$ existieren endlich viele offene ε -Kugeln, die M überdecken, d.h.

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \exists x_1, \dots, x_N : M \subseteq \bigcup_{i=1}^N B(x_i, \varepsilon).$$

Beweis. ²⁹

1. Argument für $(b) \Rightarrow (a)$ durch Kontraposition. Sei (x_n) eine Folge in M . Annahme: (x_n) besitzt keine konvergente Teilfolge. Dann existiert zu jedem $y \in M$ eine offene Menge O_y , so dass O_y nur höchstens endlich viele Folgenglieder von (x_n) enthält. Nach Konstruktion ist $(O_y)_{y \in M}$ eine offene Überdeckung von M . Da M Überdeckungskompakt ist existiert eine endliche Teilüberdeckung $(O_y)_{y \in J}$, d.h. $J \subseteq M$ endliche Menge mit $\bigcup_{y \in J} O_y \supseteq M$. Damit kann die Folge nur endlich viele Elemente besitzen und folglich muss (x_n) eine konvergente Teilfolge besitzen - ein Widerspruch!

2. Argument für $(a) \Rightarrow (c)$. M ist wegen (a) offensichtlich vollständig. Wir zeigen, dass für vollständige metrische Räume “nicht (c)” die Aussage “nicht (a)” impliziert. Es gelte “nicht (c)”. Wähle $\varepsilon > 0$, so dass sich M nicht durch endlich viele ε -Kugeln überdecken lässt. Dann können wir induktiv $x_0 \in M$ und

$$x_i \in M \setminus \left(\bigcup_{k=0, \dots, i-1} B(x_k, \varepsilon) \right) \quad \text{für } i \in \mathbb{N}_+$$

wählen. Nach Konstruktion gilt $d(x_i, x_j) \geq \varepsilon$, d.h. der paarweise Abstand beliebiger Folgenglieder ist $\geq \varepsilon$. Damit kann keine Teilfolge von $(x_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge sein. Damit besitzt (x_n) keine konvergente Teilfolge.

²⁹Korrektur 13.7.2020

3. Argument für (c) $\Rightarrow$ (b). Widerspruchsargument: Es gelte (c) und “nicht (b)”, d.h. wir können eine offene Überdeckung $(O_i)_{i \in I}$ von M wählen, mit der Eigenschaft

$$(7.1) \quad J \subseteq I \text{ endlich} \quad \Rightarrow \quad \bigcup_{j \in J} O_j \neq M.$$

2.1 Behauptung: Für $n \in \mathbb{N}$ existieren Kugeln $B_n = B(x_n; 2^{-n})$ mit $B_n \cap B_{n-1} \neq \emptyset$, welche nicht von endlich vielen O_i überdeckt werden. Argument: Wir konstruieren diese Induktiv. Für $n = 1$ wähle (gemäß (c)) endlich viele Kugeln mit Radius $1/2$, welche M überdecken. Sei $B_1 = B(x_1; 1/2)$ eine davon, welche nicht von endlich vielen O_i überdeckt werden kann (eine solche muss es geben, sonst gilt (7.1) nicht). Für den Induktionsschritt $n - 1 \mapsto n$ gehen wir von B_{n-1} aus und wählen endlich viele Kugeln mit Radius 2^{-n} , welche B_{n-1} überdecken und mit B_{n-1} einen nichtleeren Durchschnitt haben. Sei $B_n = B(x_n; 2^{-n})$ eine davon, welche nicht von endlich vielen O_i überdeckt werden kann (eine solche muss es geben nach Konstruktion von B_{n-1}).

2.2 Die Folge $(x_n)_{n \in \mathbb{N}}$ ist eine Cauchyfolge, denn es gilt nach Konstruktion $d(x_{n-1}, x_n) < 2^{-(n-1)}$. Da M vollständig ist, konvergiert $(x_n) \subseteq M$ und es existiert ein $i \in I$ mit $x := \lim_{n \rightarrow \infty} x_n \in O_i$. Da O_i offen ist, existiert $\varepsilon > 0$ mit $B(x; \varepsilon) \subseteq O_i$ und damit

$$\exists N \in \mathbb{N}, \forall n \geq N : B_n \subseteq B(x; \varepsilon) \subseteq O_i.$$

Also werden alle bis auf endliche viele der Kugeln B_n von endlich vielen der Mengen O_i überdeckt — im Widerspruch zur Konstruktion. □

Definition 7.9 (Relativ kompakt). (a) Eine Teilmenge K eines metrischen Raumes (M, d) heißt kompakt, falls (K, d) (mit der induzierten Metrik) Überdeckungskompakt (oder Folgenkompakt) ist.

(b) Eine Teilmenge K eines metrischen Raumes (M, d) heißt relativ kompakt, wenn der Abschluss von K in (M, d) kompakt ist.,

Lemma 7.10 (Heine-Borel). Eine Teilmenge $K \subseteq \mathbb{R}^d$ ist genau dann kompakt, wenn K beschränkt und abgeschlossen ist.

Beweis. Siehe Lemma 1.55. □

Beispiele 7.11. • $M = \mathbb{R}^d$ wird durch die Mengen $O_x := B(x; 2\sqrt{d})$ mit $x \in \mathbb{Z}^d$ überdeckt. $(O_x)_{x \in \mathbb{Z}^d}$ besitzt keine endliche Teilüberdeckung, denn für jede endliche Menge $J \subseteq \mathbb{Z}^d$ gilt

$$\bigcup_{x \in J} O_x \subseteq B(0; R), \quad R := 2\sqrt{d} + \max_{x \in J} \|x\|_2.$$

- Sei $M = \{x \in \text{Abb}(\mathbb{N}; \mathbb{R}) \text{ mit } \|x\|_\infty \leq 2\}$ die abgeschlossene Kugel mit Radius 2 im Raum der beschränkten Folgen. Sei $d(a, b) := \|a - b\|_\infty$. Da $(\text{Abb}(\mathbb{N}; \mathbb{R}), \|\cdot\|_\infty)$ ein Banachraum ist, ist die Menge (M, d) vollständig. Es bezeichne $e_i \in M$ die Folge mit $e_i = (e_{ij})_{j \in \mathbb{N}}$, $e_{ij} := \delta_{ij}$. Dann sind für alle $0 < \varepsilon < 1/4$ die Mengen

$$(B(e_i; \varepsilon))_{i \in \mathbb{N}} \quad \text{offen und disjunkt}$$

und für den Abstand zweier solcher Mengen gilt

$$\inf\{\|a - b\|_\infty : a \in B(e_i; \varepsilon), b \in B(e_j; \varepsilon)\} > \frac{1}{4} \quad \text{für } i \neq j.$$

Damit kann (M, d) nicht (c) aus Satz 7.8 erfüllen. Damit ist M zwar eine abgeschlossene und beschränkte Teilmenge eines Banachraumes, aber nicht kompakt!

Zum Abschluss möchten wir uns die Frage stellen, wie kompakte Teilmengen im Raum der stetigen Funktionen aussehen.

Satz 7.12 (Arzela-Ascoli). *Sei $K \subseteq \mathbb{R}^n$ kompakt. Für eine Teilmenge B des Banachraumes $(C(K), \|\cdot\|_\infty)$ sind die folgenden Aussagen dann äquivalent:*

(a) B ist relativ kompakt.

(b) B ist beschränkt und gleichgradig stetig für alle $x \in K$, d.h.

$$\forall x \in K \quad \forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall f \in B \quad \forall y \in K : \left(\|x - y\| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon \right)$$

Beweis. Wir geben nur den Beweis für die Richtung (b) $\Rightarrow$ (a) an: Sei $\varepsilon > 0$. Es genügt zu zeigen, dass $\overline{B}$ durch endlich viele Kugeln mit Radius $4\varepsilon > 0$ überdeckt werden kann. Man sieht leicht, dass auch $\overline{B}$ gleichgradig stetig ist, d.h.

$$\forall x \in K \quad \exists \delta_x > 0 \quad \forall f \in \overline{B} \quad \forall y \in K : \left(\|x - y\| < \delta_x \Rightarrow |f(x) - f(y)| < \varepsilon \right).$$

Da K kompakt ist und $K \subseteq \bigcup_{x \in K} B(x; \delta_x)$ gilt

$$\exists n \in \mathbb{N}, x_1, \dots, x_n \in K : K \subseteq \bigcup_{i=1}^n B(x_i; \delta_i), \text{ wobei } \delta_i := \delta_{x_i}.$$

Da $\overline{B}$ weiterhin beschränkt ist, existiert ein kompaktes Intervall $C \subseteq \mathbb{R}$ mit

$$\forall f \in \overline{B} : \text{Bild}(f) \subseteq C.$$

Da C kompakt ist, gibt es $y_1, \dots, y_m \in \mathbb{R}$, sodass

$$\begin{aligned} C &\subseteq \bigcup_{j=1}^m B(y_j; \varepsilon) \\ \Rightarrow \quad \forall i \in \{1, \dots, n\} \quad \forall f \in \overline{B} \quad \exists j \in \{1, \dots, m\} : f(x_i) &\in B(y_j, \varepsilon) \\ \Rightarrow \quad \overline{B} &= \bigcup_{\varphi \in \Phi} L_\varphi, \text{ wobei } \Phi = \{\varphi : \{1, \dots, n\} \rightarrow \{1, \dots, m\}\} \\ \text{und } L_\varphi &= \{f \in \overline{B} \mid f(x_i) \in B(y_{\varphi(i)}; \varepsilon)\}. \end{aligned}$$

Da Φ endlich ist genügt es zu zeigen, dass alle L_φ durch endlich viele Kugeln mit Radius 4ε überdeckt werden können. Sei $\varphi \in \Phi$, o.B.d.A. sei $f_\varphi \in L_\varphi \neq \emptyset$ fest gewählt. Dann gilt $\forall f \in L_\varphi$:

$$\begin{aligned} \|f_\varphi - f\|_\infty &\leq \max_i \max_{x \in B(x_i, \delta_i)} |f_\varphi(x) - f(x)|, \text{ wobei} \\ |f_\varphi(x) - f(x)| &\leq \underbrace{|f_\varphi(x) - f_\varphi(x_i)|}_{< \varepsilon} + \underbrace{|f_\varphi(x_i) - y_{\varphi(i)}|}_{< \varepsilon} + \underbrace{|y_{\varphi(i)} - f(x_i)|}_{< \varepsilon} + \underbrace{|f(x_i) - f(x)|}_{< \varepsilon} < 4\varepsilon \\ \Rightarrow \quad L_\varphi &\subseteq B(f_\varphi, 4\varepsilon). \end{aligned}$$

□

Bemerkung 7.13. Falls K nicht nur einen Punkt enthält, ist $C(K)$ ein unendlich dimensionaler Banachraum. Die abgeschlossene Einheitskugel $K(0;1) \subseteq C(K)$ ist in diesem Fall nicht kompakt.

Beispiel 7.14. Sei $I \subseteq \mathbb{R}$ ein kompaktes Intervall. Wir betrachten die Menge

$$B := \{f \in C^1(I) : \|f\|_\infty + \|f'\|_\infty < 1\}.$$

Für alle $f \in B$ gilt

$$|f(x) - f(y)| \leq \|f'\|_\infty |x - y| < |x - y|,$$

also ist B beschränkt und gleichgradig stetig für alle $x \in I$. Nach dem Satz von Arzela-Ascoli ist $\overline{B}$ kompakt. Für jede Folge in $(f_n) \subseteq B$ finden wir eine gleichmäßig konvergente Teilfolge mit Grenzwert $f \in C(I)$ mit $\|f\|_\infty \leq 1$.

A Schwerpunkte der Vorlesung

Kapitel	Strukturen	Verständnis	Rechenmethoden
1.1 - 1.5	• Grundbegriffe zum Vektorraum (Dimension, Linear Abbildung, Darstellungsmatrix, Basis); Beispiele • Normierter Raum, Banachraum • Operatornorm • Homöomorphismus, Isomorphismus, Isometrie	• Beispiele (insbesondere Funktionenräume, $\mathbb{K}^d$, p-Norm) • Stetigkeit linearer Abbildungen (Satz 1.25) • Äquivalenz von Normen • endlichdimensionaler Fall: Satz 1.65, Satz 1.66. • Satz vom Minimum und Maximum	• Höldersche und Youngsche Ungleichungen anwenden • Operatornorm, Matrixnorm ausrechnen
1.6	• Grundbegriffe Hilbertraum (Skalarprodukt, Winkel) • (Un)gleichungen: Cauchy-Schwarz, Pythagoras, Polarisationsformel, Parallelogrammgleichung • Orthogonalität, ONB, Orthogonales Komplement • Euklidischer Vektorraum	• Beispiele, insbesondere der Form $(x, y) := x \cdot Ay$ mit $A \in \mathbb{R}^{d \times d}$. • Lemma 1.77 • Invarianz unter Euklidischen Transformationen	• Hilbertraum-Eigenschaften nachprüfen • ONB, orthogonale Komplemente, Winkel bestimmen
1.7	• Punktweise und Gleichmäßige Konvergenz • Banachräume $(B(D), \ \cdot\ _\infty)$, $(C(K), \ \cdot\ _\infty)$	• Offene und abgeschlossene Mengen in $(B(D), \ \cdot\ _\infty)$, $(C(K), \ \cdot\ _\infty)$ • Satz 1.97	• Konvergenz (punktweise / glm.) von Funktionenfolgen und -reihen untersuchen

Kapitel	Strukturen	Verständnis	Rechenmethoden
2.1, (4.3)	• Grundbegriffe Kurven im $\mathbb{R}^d$ (Spur, Kurvenpunkt, Doppelpunkt, Schnittpunkt, Schnittwinkel, reguläre Kurve, rektifizierbare Kurve) • Beispiele für Kurven • Länge einer Kurve • Parametertransformation, Bogenlängenparametrisierung	• Eigenschaften der Länge einer Kurve • Mittelwertsatz und Zwischenwertsatz gelten nicht für Kurven	• Tangentialvektoren, Schnittpunkte und Schnittwinkel bestimmen • Berechnung der Länge (via Definition oder Satz 4.33)
2.2, 2.3	• Grundbegriffe: differenzierbar, partiell differenzierbar, totale Ableitung, partielle Ableitung, Richtungsableitung, Jacobi-Matrix, Hessische Matrix • Rechenregeln (Summen, Produkt und Kettenregeln)	• totale Ableitung versus Jacobi-Matrix • Implikationen zwischen: differenzierbar, partiell differenzierbar, stetig partiell differenzierbar, stetig; insbesondere Bemerkung 2.60, Satz 2.52 • Schrankensatz • Lemma von Schwarz	• (totale/partielle) Differenzierbarkeit überprüfen • (höhere, partielle, totale) Ableitungen berechnen

Kapitel	Strukturen	Verständnis	Rechenmethoden
2.4, 2.5, 2.6, 5 ⁰	• Skalarfeld, Gradient, Niveaumenge, Tangentialvektor, -raum und -ebene. • Taylorentwicklung (Taylorpolynom und Restterm) • Lokale Extrema (lokale/isolierte Extrema), stationärer Punkte	• MWS für Skalarfelder • Satz 2.80 und 2.84 • Geometrische Bedeutung der Taylorpolynome der Ordnung 1 und 2 • Kriterien für lokale Extrema (Satz 2.93, Satz 2.96) • <i>Lagrange Multiplikator (Satz 5.2)</i>⁰	• Gradient, Richtung des steilsten Anstiegs, Tangentialeben, -raum, -vektor bestimmen. • Taylorpolynom (insbesondere quadratische Entwicklung) bestimmen • lokale / globale Extrema und isolierte Punkte bestimmen • <i>Extremwertaufgaben unter Nebenbedingunge</i>⁰
2.7, 6 ⁰	• Vektorfeld, Gradientenfeld, Potential • Divergenz, Rotation, Laplace-Operator • harmonisches Skalarfeld • <i>konservatives Vektorfeld</i>⁰	• Beispiele für rotations- / divergenzfreie Vektorfelder, harmonische Skalarfelder • Gradientenfeld $\Rightarrow$ rotationsfrei • <i>Zusammenhang: Gradientenfeld, rotationsfrei, Integrabilitätsbedingung (Satz 6.5, Satz 6.7)</i>⁰	• Divergenz & Rotation, Laplace berechnen

⁰Nicht relevant für Lehramtsstudierende

Kapitel	Strukturen	Verständnis	Rechenmethoden
3	• strikte Kontraktion, Fixpunkt • Diffeomorphismus / Koordinatentransformation • Implizite Funktionen	• Banachscher Fixpunktsatz • totale Ableitung von Umkehrabbildungen (Lemma 3.7) • Lokaler und globaler Umkehrsatz (Satz 3.11, Satz 3.12) • Satz über implizite Funktionen (Satz 3.14)	• Abbildungen auf strikte Kontraktionseigenschaft untersuchen, Fixpunkt berechnen • Nichtlinearer Gleichungssysteme: Existenz einer Lösung beweisen, Gleichung lokal auflösen, Ableitung der Lösungsfunktion bestimmen • Umkehrabbildung: Differenzierbarkeit, Bestimmung der Ableitung
4.1, 4.2	• Treppenfunktion, Integral einer Treppenfunktion • Riemann-integrierbare Funktionen, Riemannintegral, Ober- und Unterintegral	• Eigenschaften des Integrals einer Treppenfunktion / des Riemannintegrals (Proposition 4.7, 4.18) • Zusammenhang: stetig + kompakter Träger, Riemann-integrierbar, beschränkt (Satz 4.13, Korollar 4.16) • Riemannsches Integrierbarkeitskriterium • Beispiele	• Riemann-Integrierbarkeit untersuchen • Integral von Treppenfunktionen berechnen

⁰Nicht relevant für Lehramtsstudierende

Kapitel	Strukturen	Verständnis	Rechenmethoden
4.3	• 1D-Riemannintegral • Stammfunktion, unbestimmtes Integral • Uneigentliches Riemannintegral	• Kriterien für Riemann-Integrierbarkeit in 1d: Lemma 4.20, Satz 4.22, Lemma 4.23, 4.24) • Hauptsatz der Differential- und Integralrechnung: Satz 4.28 • Mittelwertsatz, partielle Integration, Substitutionsregel • Zusammenhang: Riemann-integrierbar, Stammfunktion	• 1d-Integrale berechnen • Stammfunktion bestimmen • uneigentliche Riemannintegrale berechnen • Ober- und Unterintegral berechnen
4.4 ⁰	• <i>Jordan-messbare Mengen, n-dimensionale Volumen⁰</i> • <i>Riemann-Integral von Funktionen auf Jordan-messbaren Mengen</i>	• <i>Satz von Fubini, Prinzip von Cavalieri⁰</i> • <i>TransformationssatzSatz von Fubini, Prinzip von Cavalieri⁰</i>	• <i>Jordanmessbarkeit untersuchen</i> • <i>n-dimensionales Volumen / mehrdimensionale Integral mittels Cavalieri und Transformationssatz bestimmen</i> • <i>Fläche zwischen Funktionsgraphen bestimmen</i>

⁰Nicht relevant für Lehramtsstudierende