Zusatzmaterial zur Analysis II – Serie 1

Vektorräume

- Sei $\mathbb{K}=\mathbb{R}$ oder $\mathbb{K}=\mathbb{C}$. Eine Menge X zusammen mit einer Verknüpfung $+ : X\times X \to X$, genannt die Addition auf X, und einer (äußeren) Operation $\cdot : \mathbb{K}\times X \to X$, genannt skalare Multiplikation, heißt ein $\mathbb{K}$**-Vektorraum** oder ein **linearer Raum über** $\mathbb{K}$, falls gilt:

  (a) $(X,+)$ ist eine abelsche Gruppe und

  (b) für alle $\lambda,\mu\in\mathbb{K}$ und $x,y\in X$ gilt

  $$1_{\mathbb{K}}\cdot x = x,\quad (\lambda+\mu)\cdot x = \lambda\cdot x+\mu\cdot y,\quad \lambda\cdot(x+y)=\lambda\cdot x+\lambda\cdot y,\quad (\lambda\mu)\cdot x=\lambda\cdot(\mu\cdot x).$$

- Ein bereits bekanntes Beispiel aus dem ersten Semester ist zum Beispiel für $\mathbb{K}=\mathbb{R}$ der n-dimensionale reelle Vektorraum $X=\mathbb{R}^n$ mit der komponentenweisen Addition und Multiplikation. Für zwei Vektoren $x=(x_1,\dots,x_n)$ und $y=(y_1,\dots,y_n)$ in X und $\lambda\in\mathbb{R}$ sei $x+y:=(x_1+y_1,\dots,x_n+y_n)$ und $\lambda\cdot x:=(\lambda x_1,\dots,\lambda x_n)$. Dann ist X ein $\mathbb{R}$-Vektorraum. Ebenso ist $C([a,b],\mathbb{R}):=\big\{f:[a,b]\to\mathbb{R}\;\big|\; f \text{ stetig}\big\}$ ein Vektorraum über $\mathbb{R}$.

- Auf einem reellen Vektorraum X heißt $\langle\cdot,\cdot\rangle : X\times X\to\mathbb{R}$ **Skalarprodukt**, falls

  (i) $\langle x,y\rangle = \langle y,x\rangle$ (Symmetrie)

  (ii) $\langle \lambda x+\mu y,z\rangle = \lambda\langle x,z\rangle + \mu\langle y,z\rangle$ (Bilinearität)

  (iii) $\langle x,x\rangle\ge 0$ und $\langle x,x\rangle = 0 \iff x=0$ (Positiv-Definitheit)

  für alle $x,y,z\in X$ und für alle $\lambda,\mu\in\mathbb{R}$ gilt. Ein Tupel $(X,\langle\cdot,\cdot\rangle)$ nennen wir einen **Euklidischen Vektorraum**.

Normierte Räume

- Sei X ein $\mathbb{K}$-Vektorraum. Eine Abbildung $\|\cdot\| : X\to[0,\infty[$ heißt eine **Norm** auf X, wenn für alle $x,y\in X$ und $\lambda\in\mathbb{K}$ folgende drei Eigenschaften erfüllt sind:

  (i) $\|x\|=0 \quad\Leftrightarrow\quad x=0$ (Definitheit),

  (ii) $\|\lambda x\| = |\lambda|\|x\|$ (Homogenität),

  (iii) $\|x+y\|\le\|x\|+\|y\|$ (Dreiecksungleichung).

- Ein Vektorraum X, versehen mit einer Norm $\|\cdot\|$, heißt ein **normierter Raum** $(X,\|.\|)$.

- Ein Euklidischer Vektorraum $(X,\langle\cdot,\cdot\rangle)$ wird mit der vom Skalarprodukt induzierten Norm $\|x\|_{\langle.,.\rangle}:=\sqrt{\langle x,x\rangle}$ zu einem normierten Raum $\big(X,\|.\|_{\langle.,.\rangle}\big)$.

- Häufig verwendete Normen auf dem $\mathbb{R}$-Vektorraum $\mathbb{R}^n$ sind zum Beispiel

  $$\|x\|_2 := \sqrt{\sum_{i=1}^n |x_i|^2}\ \Big(=\sqrt{\langle x,x\rangle}\dots \text{2-Norm mit dem Euklidischen Skalarprodukt } \langle x,y\rangle:=\sum_{i=1}^n x_iy_i\Big),$$

  $$\|x\|_1 := \sum_{i=1}^n |x_i| \text{ (1-Norm)},\qquad \|x\|_\infty := \max_{i=1,\dots,n}|x_i| \text{ (Maximumsnorm)},$$

  wobei die 1-Norm und die 2-Norm Spezialfälle der für reellwertiges $p\in[1,\infty[$ gegebenen p-Norm

  $$\|x\|_p := \sqrt[p]{\sum_{i=1}^n |x_i|^p}$$

sind. Die Dreiecksungleichung bezeichnet man in diesem Fall als **Minkowski-Ungleichung**

$$\|x+y\|_p \le \|x\|_p + \|y\|_p \ .$$

Beweisen kann man sie für $p \in]1,\infty[$ (vgl. Forster I, §16) mittels der **Hölder-Ungleichung**

$$\sum_{k=1}^n |x_k y_k| \ \le\ \|x\|_p \|y\|_q \ =\ \left(\sum_{k=1}^n |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^n |y_k|^q\right)^{\frac{1}{q}} \qquad \Big(\text{wobei } \tfrac{1}{p}+\tfrac{1}{q}=1\Big),$$

die für $p=2$ mit der **Cauchy-Schwarz-Ungleichung** $|\langle x,y\rangle| \le \|x\|_2\|y\|_2$ zusammenfällt.

- Auf dem Vektorraum $C([a,b],\mathbb{R}) := \big\{ f : [a,b] \to \mathbb{R} \ \big|\ f \text{ stetig } \big\}$ kennen wir die Normen

  (a) $\|f\|_{\langle\cdot,\cdot\rangle_{C([a,b])}} := \sqrt{\int_a^b (f(x))^2\, dx} \ \ \Big(= \sqrt{\langle x,x\rangle} \text{ mit } \langle f,g\rangle_{C([a,b])} := \int_a^b (fg)(x)dx \Big)$

  (b) $\|f\|_\infty \ :=\ \max\limits_{x\in[a,b]} |f(x)|$ (existiert, da $|f|$ stetig, $[a,b]$ abgeschlossen und beschränkt)

Metrische Räume

- Sei $M \neq \emptyset$ eine Menge. Eine Abbildung $d : M \times M \to \mathbb{R}$ heißt **Metrik** auf M, falls

  (i) $\forall x,y \in M \ :\ d(x,y) = 0 \iff x = y\,,$ (Definitheit)

  (ii) $\forall x,y \in M \ :\ d(x,y) = d(y,x)\,,$ (Symmetrie)

  (iii) $\forall x,y,z \in M \ :\ d(x,z) \le d(x,y) + d(y,z)$ (Dreiecksungleichung).

- Ist d eine Metrik auf M, so heißt das Paar (M,d) **metrischer Raum**. Für je zwei Elemente $x,y \in M$ heißt die Zahl $d(x,y)$ der Abstand oder die Distanz von x und y.

- Jeder normierte Vektorraum $(X,\|\cdot\|)$ wird mit der norminduzierten Metrik $d_{\|\cdot\|}(x,y) := \|x-y\|$ zu einem metrischen Raum $\big(X, d_{\|\cdot\|}\big)$.

- In einem metrischen Raum (M,d) nennen wir eine Menge $U \subset M$ genau dann **offen**, wenn es zu jedem $x \in U$ ein $\varepsilon > 0$ existiert, so dass die offene Kugel $B_\varepsilon(x)$ um x von Radius ε bezüglich der Metrik d vollständig in U enthalten ist; mit anderen Worten, wenn

  $$\forall x \in U \exists \varepsilon > 0 \ : B_\varepsilon(x) \subset U, \qquad \text{wobei } B_\varepsilon(x) := \{y \in M |\ d(x,y) < \varepsilon\}\,.$$

- Unter der **Einheitskugel** eines normierten Raumes $(X,\|\cdot\|)$ versteht man die bezüglich der durch die Norm induzierten Metrik offene Kugel um den Ursprung von Radius 1, d.h. die Menge

  $$\{x \in X \mid d_{\|\cdot\|}(0,x) < 1\} \ =\ \{x \in X \mid \|x\| < 1\}.$$

- Die Menge T_d aller offenen Teilmengen eines metrischen Raumes (M,d) nennt man die (von der Metrik d induzierte) **Topologie** von M. Die Menge T_d besitzt die Eigenschaften:

  - Beliebige Vereinigungen offener Mengen sowie endliche Durchschnitte offener Mengen sind wieder offen.
  - $\emptyset \in T_d$ und $M \in T_d$.

  Allgemein heißt ein System T von Teilmengen von M mit diesen beiden Eigenschaften eine **Topologie** auf der Menge M, das Paar (M,T) nennt man einen **topologischen Raum** und die Elemente von T **offen**.

  Achtung:

  Nicht jede Topologie wird durch eine Metrik induziert, es gibt z.B. Topologien, die nicht **Hausdorffsch** sind, d.h. für die es nicht zu je zwei (verschiedenen) Punkten $x,y \in M$ offene Mengen (sog. Umgebungen) U,V mit $x \in U$, $y \in V$ und $U \cap V = \emptyset$ gibt.

- Zwei Metriken d_1, d_2 auf $M \neq \emptyset$ heißen **äquivalent**, wenn sie dieselbe Topologie induzieren.

Zusatzaufgabe 1.1:

(a) Zeigen Sie, dass die Nichtnegativität einer Norm bzw. einer Metrik nicht explizit gefordert werden muss, sondern sich aus den übrigen Eigenschaften einer Norm bzw. Metrik ergibt.

(b) Geben Sie alle Normen auf $\mathbb{R}$ an und beweisen Sie, dass Sie wirklich alle gefunden haben.

(c) Wie lassen sich die Begriffe der Konvergenz von Folgen in bzw. Stetigkeit von Funktionen zwischen metrischen Räumen aus den aus der Analysis I bekannten Begriffen übertragen ?

(d) Zeigen Sie: Für vier beliebige Punkte $w, x, y, z \in X$ eines metrischen Raumes (X, d) ist stets die sogenannte **Vierecksungleichung** $|d(w, y) - d(x, z)| \leq d(w, x) + d(y, z)$ erfüllt.

Lösung zu Zusatzaufgabe 1.1:

(a)
- $\forall x \in X : 0 = 0\|0\| = \|0\| = \|x - x\| \leq \|x\| + \| - x\| = 2\|x\| \implies \|x\| \geq 0$
- $\forall x, y \in M : 0 = d(x, x) \leq d(x, y) + d(y, x) = 2d(x, y) \implies d(x, y) \geq 0$

(b) Jede Norm auf $\mathbb{R}$ besitzt die Gestalt $\|x\| = c|x|$ mit einer Konstanten $c > 0$ (genauer $c = \|1\|$), was unmittelbar aus der Gleichung $\|x\| = |x|\|1\|$ folgt.

(c)
- $x_n \xrightarrow{d} x \;:\Longleftrightarrow\; \forall \varepsilon > 0 \;\exists N_\varepsilon \in \mathbb{N} \;\forall n \in \mathbb{N} \;:\; \Big(n \geq N_\varepsilon \implies d(x_n, x) < \varepsilon\Big)$
- Seien (X, d_X) und (Y, d_Y) metrische Räume. Dann ist $f : (X, d_X) \to (Y, d_Y)$ stetig in $a \in X \;:\Longleftrightarrow\; \forall \varepsilon > 0 \;\exists \delta_{\varepsilon,a} > 0 \;\forall x \in X \;:\; \Big(d_X(x, a) < \delta \implies d_Y(f(x), f(a)) < \varepsilon\Big)$

(d) Für beliebige Punkte $w, x, y, z \in X$ ergeben sich nach zweifacher Anwendung der Dreiecksungleichung die beiden Ungleichungsketten

$$\begin{array}{rcccl} d(w,y) & \leq & d(w,x) + d(x,y) & \leq & d(w,x) + d(x,z) + d(z,y) \;, \\ d(x,z) & \leq & d(x,w) + d(w,z) & \leq & d(x,w) + d(w,y) + d(y,z) \;. \end{array}$$

Unter Verwendung der Symmetrie der Metrik und mittels Subtraktion von $d(x, z)$ beziehungsweise $d(w, y)$ gelangen wir zu

$$\begin{array}{rcl} d(w,y) - d(x,z) & \leq & d(w,x) + d(y,z) \;, \\ d(x,z) - d(w,y) & \leq & d(w,x) + d(y,z) \;. \end{array}$$

Dies ist jedoch äquivalent zur Behauptung.

Zusatzaufgabe 1.2:

(a) Wie sieht die offene Einheitskugel im $\mathbb{R}^2$ um den Nullpunkt bezüglich der Metrik $d_*(x, y) := \|x - y\|_*$ für $* \in \{1, 2, \infty\}$ aus?

(b) Geben Sie alle Topologien von $M = \{a, b\}$ an.

(c) Zeigen Sie, dass $\{\emptyset, \{u\}, X\}$ eine Topologie auf $X := \{u, v, w\}$ definiert, die jedoch nicht Hausdorffsch ist.

(d) Zeigen Sie, dass jeder metrische Raum (M, d) Hausdorffsch ist.

Lösung zu Zusatzaufgabe 1.2:

(a) (i) Raute $\big\{(x_1, x_2) \in \mathbb{R}^2 \,\big|\, |x_1| + |x_2| < 1\big\}$;

(ii) Einheitskreis $\big\{(x_1, x_2) \in \mathbb{R}^2 \,\big|\, x_1^2 + x_2^2 < 1\big\}$.

(iii) Quadrat $]-1, 1[\;\times\;]-1, 1[\;:=\; \big\{(x_1, x_2) \in \mathbb{R}^2 \,\big|\, -1 < x_1 < 1, -1 < x_2 < 1\big\}$;

(b) $T_1 = \{\emptyset, M\}$, $T_2 = \{\emptyset, \{a\}, M\}$, $T_3 = \{\emptyset, \{b\}, M\}$, $T_4 = \{\emptyset, \{a\}, \{b\}, M\} = \mathcal{P}(M)$ sind genau alle Topologien auf $M = \{a, b\}$

(c) Nachzuprüfen ist, dass neben $\emptyset$ und X auch jeder endliche Schnitt und jede beliebige Vereinigung von Elementen aus $\{\emptyset, \{u\}, X\}$ wieder in $\{\emptyset, \{u\}, X\}$ liegen. Dies ist offensichtlich der Fall, da wegen $\emptyset \subset \{u\} \subset X$ jeder Schnitt und jede Vereinigung genau einer dieser Mengen ergibt.
Jedoch besitzt beispielsweise der Punkt w nur die offene Umgebung $X = \{u, v, w\}$ und damit keine Umgebung, die nicht auch u oder w enthalten würde. Somit ist das Hausdorffsche Trennungsaxiom nicht erfüllt.

(d) Seien $x \neq y$ aus M. Aufgrund der Definitheit gilt dann $d(x, y) = \delta > 0$. Betrachte nun für $\varepsilon := \frac{\delta}{3}$ die offenen ε-Kugeln

$$B_\varepsilon(x) \;=\; \{z \in M \;:\; d(x, z) < \varepsilon\}\ , \qquad B_\varepsilon(y) \;=\; \{z \in M \;:\; d(x, z) < \varepsilon\}\ .$$

Dann gilt $B_\varepsilon(x) \cap B_\varepsilon(y) = \emptyset$. Wäre nämlich $z \in B_\varepsilon(x) \cap B_\varepsilon(y)$, so folgte nach Dreiecksungleichung $\delta = d(x, y) \leq d(x, z) + d(z, y) < \frac{\delta}{3} + \frac{\delta}{3} = \frac{2\delta}{3}$, Widerspruch. Somit existieren zu je zwei verschiedenen Punkten $x, y \in M$ disjunkte offene $U, V \subset M$ mit $x \in U$ und $y \in V$.

Zusatzaufgabe 1.3:

(a) Welche Folgen konvergieren in (M, d_{disk}) mit $d_{\text{disk}}(x, y) := \begin{cases} 1 & \text{bei } x \neq y, \\ 0 & \text{bei } x = y \end{cases}$?

(b) Sei $-\infty < a < b < \infty$ und $[a, b] \subset \mathbb{R}$ ein Intervall. Weiter sei $X = C([a, b], \mathbb{R})$ und $d : X \times X \to \mathbb{R}$ für $f, g \in X$ definiert durch $d(f, g) := \sup\limits_{x \in [a,b]} \{|f(x) - g(x)|\}$.

 (i) Zeigen Sie, dass (X, d) ein metrischer Raum ist.

 (ii) Konvergiert die Folge durch $f_n(x) := \frac{\sin(x)}{n}$ definierte Folge $(f_n)_{n \in \mathbb{N}}$ in (X, d)?
 Was sagt dies über die punktweise und gleichmäßige Konvergenz der Folge $(f_n)_{n \in \mathbb{N}}$?

(c) Ist $d(f, g) := |f(0) - g(0)|$ eine Metrik auf $C([0, 1], \mathbb{R}) = \{f : [0, 1] \to \mathbb{R} \mid f \text{ stetig }\}$?

Lösung zu Zusatzaufgabe 1.3:

(a) Wegen $B_{\frac{1}{2}}(x) = \{x\}$ konvergieren in (M, d_{disk}) nur Folgen, welche ab einem endlichen Index K konstant sind.

(b) (i) Es gilt $X \neq \emptyset$, da die Nullfunktion stetig ist. Desweiteren erfüllt d die Eigenschaften:

- $d(f, g) = 0 \iff \sup\limits_{x \in [a,b]} \{|f(x) - g(x)|\} = 0 \iff \forall x \in [a, b] \;:\; |f(x) - g(x)| = 0$
 $\overset{|.|\text{ Norm}}{\iff} \forall x \in [a, b] \;:\; f(x) - g(x) = 0 \iff \forall x \in [a, b] \;:\; f(x) = g(x) \iff f = g$;
- $d(f, g) = \sup\limits_{x \in [a,b]} |f(x) - g(x)| = \sup\limits_{x \in [a,b]} |(-1)(g(x) - f(x))|$
 $\overset{|.|\text{ Norm}}{=} \sup\limits_{x \in [a,b]} |g(x) - f(x)| = d(g, f)$;
- $d(f, h) = \sup\limits_{x \in [a,b]} |f(x) - h(x)| \leq \sup\limits_{x \in [a,b]} \{|f(x) - g(x)| + |g(x) - h(x)|\}$
 $\leq \sup\limits_{x \in [a,b]} |f(x) - g(x)| + \sup\limits_{x \in [a,b]} |g(x) - h(x)| = d(f, g) + d(g, h)$

 (ii) Sei $f \equiv 0$. Dann konvergiert die Folge $(f_n)_{n \in \mathbb{N}}$ in (X, d) gegen f wegen

$$d(f_n, f) = \sup_{x \in [a,b]} |f_n(x) - f(x)| = \sup_{x \in [a,b]} \left|\frac{\sin(x)}{n} - 0\right| = \frac{1}{n} \sup_{x \in [a,b]} |\sin(x)| \;\leq\; \frac{1}{n} \overset{n \to \infty}{\longrightarrow} 0.$$

 Die Konvergenz in (X, d) ist jedoch genau die gleichmäßige Konvergenz auf dem Intervall $[a, b]$, welche zudem die punktweise Konvergenz impliziert.

(c) Nein, denn beispielsweise für $f(x) = x^2$ und $g(x) = 0$ gilt einerseits $f(0) = g(0)$, also demzufolge $d(f,g) = 0$, aber offenbar ist $f \neq g$, womit die Definitheit verletzt wäre.

Zusatzaufgabe 1.4:

(a) Zeigen Sie, dass durch $d(x,y) := \arctan(|x-y|)$ eine Metrik auf $\mathbb{R}$ definiert wird.

(b) Ist die Metrik d aus (a) äquivalent zur Euklidischen Metrik $d_{Euklid}(x,y) := |x-y|$?

(c) Sei $f : \mathbb{R}^n \to \mathbb{R}^n$ injektiv. Zeigen Sie, dass durch $d_f(x,y) := \|f(x) - f(y)\|_2$ eine Metrik d_f auf $\mathbb{R}^n$ definiert wird.

(d) Betrachte auf dem $\mathbb{R}^n$ die zu $f(x) := (\mathrm{e}^{x_1}, \ldots, \mathrm{e}^{x_n})$ gehörige Metrik d_f aus (c). Ist die Folge $x^{(k)} := \left(\ln\left(\frac{1}{k}\right), \ldots, \ln\left(\frac{1}{k}\right)\right)$ eine Cauchy-Folge in $(\mathbb{R}^n, d_f)$? Konvergiert $x^{(k)}$ in $(\mathbb{R}^n, d_f)$?

Lösung zu Zusatzaufgabe 1.4:

(a) Wegen $\arctan(0) = 0$, $\arctan'(x) = \frac{1}{1+x^2} > 0$ für alle $x \in \mathbb{R}$ wissen wir somit, dass $\arctan(x)$ auf $]0,\infty[$ streng monoton wächst (also insbesondere injektiv) ist und dass die Ableitung für $x \geq 0$ wegen $\arctan''(x) = -\frac{2x}{(1+x^2)^2} \leq 0$ monoton fallend ist. Somit erfüllt d alle Eigenschaften einer Metrik, da

- $\forall x,y \in \mathbb{R} \;:\; d(x,y) = 0 \iff \arctan(|x-y|) = 0 \iff |x-y| = 0 \iff x = y\,.$
- $\forall x,y \in \mathbb{R} \;:\; d(x,y) = \arctan(|x-y|) = \arctan(|y-x|) = d(y,x)$
- $\forall x,y,z \in \mathbb{R} \;:$

$$\begin{aligned} d(x,y) = \arctan(|x-y|) &\leq \arctan(|x-z| + |z-y|) \\ &\leq \arctan(|x-z|) + \arctan(|z-y|) = d(x,z) + d(z,y)\,. \end{aligned}$$

Letzteres folgt aus $\arctan''(x) \leq 0$ und $\arctan(0) = 0$ und dem Argument, dass aufgrund der Monotonie des Integrals für $a,b > 0$ die folgende Abschätzung gilt:

$$\arctan(a+b) - \arctan(a) = \int_0^b \frac{1}{(a+x)^2+1}dx \leq \int_0^b \frac{1}{x^2+1}dx = \arctan(b)\,.$$

(b) Zwei Metriken sind äquivalent, falls sie dieselbe Topologie erzeugen, d.h. hier, dass eine Menge genau dann bezüglich d offen ist, wenn sie offen bezüglich d_{Euklid} ist. Da in $(\mathbb{R}, d_{Euklid})$ die offenen ε-Kugeln um ein x_0 genau die offenen Intervalle $]x_0 - \varepsilon, x_0 + \varepsilon[$ sind, folgt die Behauptung sofort aus der (rechtsseitigen) Stetigkeit von $\arctan(x)$ (bzw. $\tan(x)$) bei Null.

(c) Die drei Eigenschaften einer Metrik sind erfüllt:

- Die Definitheit ist aufgrund der Injektivität von f erfüllt:

$$\forall x,y \in \mathbb{R}^n \;:\; d_f(x,y) = 0 \iff \|f(x) - f(y)\|_2 = 0 \iff f(x) = f(y) \iff x = y.$$

- Die Symmetrie ist erfüllt:

$$\forall x,y \in \mathbb{R}^n \;:\; d_f(x,y) = \|f(x) - f(y)\|_2 = \|f(y) - f(x)\|_2 = d_f(y,x)\,.$$

- Die Dreiecksungleichung ist erfüllt:

$$\begin{aligned} \forall x,y,z \in \mathbb{R}^n \;:\; d_f(x,z) = \|f(x) - f(z)\|_2 &\leq \|f(x) - f(y)\|_2 + \|f(y) - f(z)\|_2 \\ &= d_f(x,y) + d_f(y,z)\,. \end{aligned}$$

(d) Die angegebene Folge $x^{(k)}$ ist eine CAUCHY-Folge, denn zu $\varepsilon > 0$ wähle ein $K \in \mathbb{N}$ mit $K > \frac{2\sqrt{n}}{\varepsilon}$ (beachte, dass n hier als Vektorraumdimension fest ist), dann gilt

$$d_f\left(x^{(k)}, x^{(l)}\right) = \left\| f\left(x^{(k)}\right) - f\left(x^{(l)}\right)\right\|_2 = \left\| \left(\tfrac{1}{k}, \ldots, \tfrac{1}{k}\right)^T - \left(\tfrac{1}{l}, \ldots, \tfrac{1}{l}\right)^T \right\|_2 \leq \left(\tfrac{1}{k} + \tfrac{1}{l}\right)\sqrt{n} < \varepsilon$$

für $k, l \geq K$. Jedoch konvergiert $x^{(k)}$ bezüglich d_f nicht, denn wäre x der Grenzwert, dann müsste $f(x^{(k)})$ bezüglich der Euklidischen Norm gegen $f(x)$ konvergieren. Nun konvergiert aber die Folge $y^{(k)} = f(x^{(k)})$ bezüglich $\|\cdot\|_{\text{Euklid}}$ gegen den Nullpunkt, es gibt jedoch kein $x \in \mathbb{R}^n$ mit $f(x) = 0$ und somit konvergiert $x^{(k)}$ nicht im metrischen Raum $(\mathbb{R}^n, d_f)$.

Zusatzaufgabe 1.5:

(a) Sei (X, d) ein metrischer Raum und $f : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ eine stetige, streng monoton wachsende Funktion mit $f(0) = 0$, die zusätzlich $f(x + y) \leq f(x) + f(y)$ für alle $x, y \geq 0$ erfüllt.

  (i) Überprüfen Sie, ob dann auch $\tilde{d}(x, y) := f(d(x, y))$ eine Metrik auf X ist.

  (ii) Zeigen Sie, dass die Topologien von (X, d) und $(X, \tilde{d})$ identisch sind.

(b) Zeigen Sie, dass $(\mathbb{N}, d_{\mathbb{N}})$ mit $d_{\mathbb{N}}(m, n) := \left|\frac{1}{n} - \frac{1}{m}\right|$ ein metrischer Raum ist.

Lösung zu Zusatzaufgabe 1.5:

(a) (i) Es gilt offensichtlich $\tilde{d}(x, y) \geq 0$ wegen $f \geq 0$. Da außerdem $f(t) = 0$ wegen des streng monotonen Wachstums genau bei $t = 0$ gilt, ist $\tilde{d}(x, y) = 0$ äquivalent zu $d(x, y) = 0$ und daher zu $x = y$. Die Symmetrie $\tilde{d}(x, y) = \tilde{d}(y, x)$ folgt sofort aus der Symmetrie von $d(x, y)$.

Zu zeigen bleibt also nur noch die Dreiecksungleichung. Diese ergibt sich aus der Dreiecksungleichung für d, indem wir die Sublinearität und das monotone Wachstum von f anwenden:

$$\tilde{d}(x,y)+\tilde{d}(y,z) \;=\; f(d(x,y))+f(d(y,z)) \;\geq\; f(d(x,y)+d(y,z)) \;\geq\; f(d(x,z)) \;=\; \tilde{d}(x,z)$$

(ii) Sei $U \subset (X, d)$ offen und $x_0 \in U$ beliebig. Dann existiert ein $\varepsilon > 0$, so dass die offene Kugel $B_\varepsilon(x_0)$ mit Radius $\varepsilon > 0$ um x_0 (bezüglich der Metrik d) vollständig in U enthalten ist. Wählen wir nun $0 < \tilde{\varepsilon} < f(\varepsilon)$, so ist die offene Kugel $\tilde{B}_{\tilde{\varepsilon}}(x_0)$ um x_0 in X bezüglich der Metrik $\tilde{d}$ in $B_\varepsilon(x_0)$ enthalten. In der Tat, für $x \in X$ mit $\tilde{d}(x, x_0) < \tilde{\varepsilon}$ gilt gemäß Definition zunächst $f(d(x, x_0)) < \tilde{\varepsilon} < f(\varepsilon)$ und damit wegen der strengen Monotonie von f insbesondere $d(x, x_0) < \varepsilon$, also $\tilde{B}_{\tilde{\varepsilon}}(x_0) \subset B_\varepsilon(x_0) \subset U$.

Umgekehrt, sei nun $V \subset (X, \tilde{d})$ offen und $x_0 \in V$ beliebig. Da V offen ist, existiert ein $\tilde{\varepsilon} > 0$, so dass V die offene Kugel $\tilde{B}_{\tilde{\varepsilon}}(x_0)$ um x_0 vom Radius $\tilde{\varepsilon}$ bezüglich der Metrik $\tilde{d}$ enthält. Da f nach Voraussetzung stetig ist, finden wir ein $\tilde{\delta} > 0$, so dass für alle reellen Zahlen $0 \leq t < \tilde{\delta}$ die Ungleichung $f(t) < \tilde{\varepsilon}$ erfüllt ist. Folglich gilt im Falle $0 < \varepsilon < \tilde{\delta}$

$$\tilde{d}(x, x_0) = f(d(x, x_0)) < \tilde{\varepsilon}.$$

für alle $x \in X$ mit $d(x, x_0) < \varepsilon$. Damit erhalten wir also $B_\varepsilon(x_0) \subset \tilde{B}_{\tilde{\varepsilon}}(x_0)$ für jedes $0 < \varepsilon < \tilde{\delta}$.

(b) Damit $(\mathbb{N}, d_{\mathbb{N}})$ ein metrischer Raum wird, muss $d_{\mathbb{N}} : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ die drei Eigenschaften einer Metrik erfüllen:

- $\forall m, n \in \mathbb{N} : d_{\mathbb{N}}(m, n) = \left|\frac{1}{n} - \frac{1}{m}\right| \geq 0$.
  $\forall m, n \in \mathbb{N} : \left(d_{\mathbb{N}}(m, n) = 0 \iff \left|\frac{1}{n} - \frac{1}{m}\right| = 0 \iff \frac{1}{n} = \frac{1}{m} \iff m = n\right)$
- $\forall m, n \in \mathbb{N} : d_{\mathbb{N}}(m, n) = \left|\frac{1}{n} - \frac{1}{m}\right| = \left|\frac{1}{m} - \frac{1}{n}\right| = d_{\mathbb{N}}(n, m)$.
- $\forall l, m, n \in \mathbb{N} : d_{\mathbb{N}}(l, n) = \left|\frac{1}{n} - \frac{1}{l}\right| \leq \left|\frac{1}{n} - \frac{1}{m}\right| + \left|\frac{1}{m} - \frac{1}{l}\right| = d_{\mathbb{N}}(m, n) + d_{\mathbb{N}}(l, m)$.

Alle Übungsaufgaben sind verfügbar unter: `https://studip.uni-rostock.de/studip/`

Zusatzmaterial zur Analysis II – Serie 2

Äquivalenz von Normen

- Zwei Normen auf einem Vektorraum X heißen **äquivalent**, wenn die beiden induzierten Metriken dieselbe Topologie besitzen. Dies ist genau dann der Fall, wenn es Konstanten $C, c > 0$ gibt, so dass für alle $x \in X$ die Ungleichung $c\|x\| \leq \|x\|' \leq C\|x\|$ erfüllt ist.
- Auf dem $\mathbb{R}^n$, also insbesondere in jedem endlich-dimensionalen normierten Raum, sind alle Normen äquivalent.
- Äquivalenz von Normen bildet eine Äquivalenzrelation auf der Menge aller Normen eines Vektorraumes X, d.h. die Äquivalenz von Normen ist reflexiv, symmetrisch und transitiv.

Konvergenz in normierten Räumen

- Eine Folge x_k von Punkten in einem normierten Vektorraum $(X, \|\cdot\|)$ nennt man konvergent gegen den Punkt $x \in X$, falls $\lim_{k\to\infty} \|x_k - x\| = 0$ in $\mathbb{R}$ gilt oder äquivalenterweise falls $\forall \varepsilon > 0\ \exists K \in \mathbb{N}\ \forall k \in \mathbb{N}\ :\ (k \geq K \Longrightarrow \|x_k - x\| \leq \varepsilon)$.
- Für äquivalente Normen konvergiert eine Folge genau dann bezüglich einer der beiden Normen, wenn sie es auch bezüglich der anderen tut.
- **Satz 1 [Forster II, §2]**: Eine Folge $(x_k)_{k\in\mathbb{N}}$ aus dem $\mathbb{R}^n$ konvergiert genau dann gegen ein $x \in \mathbb{R}^n$, wenn alle Komponenten der Folge in $\mathbb{R}$ gegen die jeweilige Komponente von x konvergieren.
- **Achtung:** Die Konvergenz einer Folge in einem endlich-dimensionalen normierten Raum hängt nicht von der Norm ab. Konvergiert sie bezüglich einer Norm, so auch bezüglich aller anderen Normen. In unendlich-dimensionalen normierten Räumen kann die Konvergenz einer Folge sehr wohl von der gewählten Norm abhängen.

Umgebung, Berührpunkt, Abschluss, Inneres, Rand: Sei (X, d) ein metrischer Raum.

- Eine Teilmenge $V \subset X$ heißt **Umgebung** des Punktes $x \in X$, falls ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subset V$ – oder äquivalenterweise – falls eine offene Menge $U \subset V$ mit $x \in U$ existiert.
- **Definition [abgeschlossene Teilmenge]**: Eine Teilmenge $A \subset X$ heißt **abgeschlossen**, wenn ihr Komplement $U = X \setminus A$ offen ist.
- **Definition [Berührpunkt]**: Ein Punkt x heißt **Berührpunkt** einer Teilmenge $M \subset X$, wenn es eine Folge $\{x_k\}_{k\in\mathbb{N}} \subset M$ mit $\lim_{k\to\infty} x_k = x$ gibt, oder äquivalenterweise, wenn in jeder Umgebung von x mindestens ein Punkt aus M liegt.
- **Definition [Abschluss]**: Die Menge aller Berührpunkte von M bezeichnet man mit $\overline{M}$ und nennt sie den **Abschluss** von M. Insbesondere gilt $M \subset \overline{M}$, da für $x \in M$ die konstante Folge $x, x, \ldots$ offensichtlich gegen x konvergiert.
- **Satz 2 [Forster II, §2]**: Eine Teilmenge $A \subset X$ ist genau dann **abgeschlossen**, wenn für jede Folge $(x_k)_{k\in\mathbb{N}} \subset A$, die in X konvergiert, auch der Grenzwert $\lim_{k\to\infty} x_k$ in A liegt, – mit anderen Worten – wenn sie jeden ihrer Berührpunkte enthält, d.h. wenn $A = \overline{A}$ gilt.

- In der Tat ist $\overline{M}$ die minimale M enthaltende in (X,d) abgeschlossene Menge. Analog nennt man die maximale in M enthaltene offene Menge das **Innere** von M und bezeichnet sie mit $\overset{\circ}{M}$. Die Menge $\overset{\circ}{M}$ besteht genau aus den **inneren Punkten** von M, d.h. aus denjenigen Punkten $x \in M$, für die es eine Kugel $B_r(x)$ $(r > 0)$ um x mit $B_r(x) \subset M$ gibt. Zuletzt definiert man noch den **Rand** ∂M einer Teilmenge $M \subset X$ durch $\partial M := \overline{M} \setminus \overset{\circ}{M}$. Der Rand von M besteht genau aus den Punkten x, für die in jeder Umgebung von x sowohl Punkte aus M als auch aus $X \setminus M$ liegen, d.h.,

$$\partial M \;=\; \{x \in X \mid \forall \varepsilon > 0 \; \exists y \in B_\varepsilon(x) \cap M, \exists z \in B_\varepsilon(x) \setminus M\} \,.$$

- **Definition**: Für $\emptyset \neq A \subset X$ bezeichnet $\operatorname{diam}(A) := \sup\limits_{x,y \in A} d(x,y)$ den **Durchmesser**.
- Es gilt: A ist beschränkt $\iff$ $\operatorname{diam}(A) < \infty$.
- **Satz 4 [Schachtelungsprinzip, Forster II, §2]**: Ist (X,d) vollständig und $A_1 \supset A_2 \supset \ldots$ absteigende Folge nichtleerer abgeschlossener Teilmengen von X mit $\lim\limits_{k\to\infty} \operatorname{diam}(A_k) = 0$, dann existiert genau ein $x \in X$, so dass $\forall k \in \mathbb{N} \,:\, x \in A_k$.

Vollständigkeit, Banach-Raum, Hilbert-Raum

- **Definition [Cauchy-Folge]:** Eine Folge $(x_n)_{n\in\mathbb{N}}$ in einem metrischen Raum (X,d) heißt eine **Cauchy-Folge**, wenn zu jedem $\varepsilon > 0$ ein Index $N = N(\varepsilon) \in \mathbb{N}$ existiert, so dass für alle $l,k \geq N$ die Ungleichung $d(x_k, x_l) < \varepsilon$ gilt.
- **Definition [Vollständigkeit]:** Ein metrischer Raum (X,d) heißt **vollständig**, wenn in (X,d) jede Cauchy-Folge $(x_n)_{n\in\mathbb{N}}$ konvergiert.
- **Definition [Banach-Raum]**: Einen vollständigen normierten Raum $(X, \|\cdot\|)$ nennt man einen **Banach-Raum**.
  Beispiele: $(\mathbb{R}^n, \|\cdot\|)$ mit beliebiger Norm (Satz 3, §2 Forster II) und $\big(C([a,b],\mathbb{R}), \|.\|_\infty\big)$.
- **Definition [Hilbert-Raum]**: Ein vollständiger Euklidischer Raum[1] heißt **Hilbert-Raum**.

Stetigkeit: Seien (X, d_X), (Y, d_Y), (Z, d_Z) metrische Räume.

- **Definition [Stetigkeit]**: Eine Abbildung $f : X \to Y$ heißt **stetig** im Punkt $a \in X$

  $:\iff$ Für jede Folge $(x_k)_{k\in\mathbb{N}}$ aus X mit $\lim\limits_{k\to\infty} x_k = a$ folgt $\lim\limits_{k\to\infty} f(x_k) = f(a)$.

  $:\iff$ [**ε-δ-Kriterium**] Zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$, so dass für alle $x \in X$ die Implikation $d_X(x,a) \leq \delta \implies d_Y(f(x), f(a)) \leq \varepsilon$ gilt (vgl. Satz 8, §2, Forster II).

  $:\iff$ Zu jeder Umgebung $V \subset Y$ von $f(a)$ existiert eine Umgebung $U \subset X$ von a mit $f(U) \subset V$ (vgl. Satz 11, §2, Forster II).

  $\iff$ Urbilder offener Mengen in Y sind offen in X (vgl. Satz 11, §2, Forster II)

  $\iff$ Urbilder abgeschlossener Mengen in Y sind abgeschlossen in X (vgl. Corollar).

  Ist f in allen $a \in X$ stetig, so bezeichnen wir f als **stetig** auf X oder einfach stetig.

- **Satz 5 [Forster II, §2]**: Sind $g : X \to Y$ in $a \in X$ und $h : Y \to Z$ in $g(a) \in Y$ stetig, dann ist auch die Komposition $h \circ g : X \to Z$ stetig.
- **Satz 6 [Forster II, §2]**: Die Abbildung $f = (f_1, f_2) : X \to Y \times Z$ ist genau dann in $a \in X$ stetig, wenn die Komponentenfunktionen $f_1 : X \to Y$ und $f_2 : X \to Z$ stetig sind.

[1] Der $\mathbb{R}^n$ wird mit dem üblichen Skalarprodukt zu einem Hilbert-Raum (vgl. Normäquivalenz auf dem $\mathbb{R}^n$).

- Beispiele für stetige Abbildungen sind die konstanten Abbildungen zwischen metrischen Räumen, die Norm sowie die Addition und skalare Multiplikation eines normierten Raumes oder die Projektion $X \times Y \ni (x,y) \mapsto x \in X$. Darüber hinaus sind für stetige Funktionen $f, g : X \to \mathbb{R}$ die Abbildungen $f+g$ und $f \cdot g$ stetig. Gilt zusätzlich $g(x) \neq 0$ für alle $x \in X$, so ist auch $\frac{f}{g}$ stetig auf X (vgl. Satz 7 und Corollar, Forster II, §2).

- **Definition [Homöomorphismus]**: Eine stetige Bijektion $f : X \to Y$, deren Umkehrabbildung $f^{-1} : Y \to X$ auch stetig ist, heißt ein **Homöomorphismus** zwischen (X, d_X) und (Y, d_Y) . Die Räume (X, d_X) und (Y, d_Y) werden dann **homöomorph** genannt. Für solch eine Abbildung ist $x_k \overset{d_X}{\to} x$ äquivalent zu $f(x_k) \overset{d_Y}{\to} f(x)$.

- **Definition [Lipschitz-stetig, gleichmäßig stetig]**: Eine Abbildung $f : X \to Y$ heißt
  - **Lipschitz-stetig** $:\iff \exists L \geq 0 \forall x, x' \in X \ : \ d_Y(f(x), f(x')) \leq L\, d_X(x, x')$.
  - **glm. stetig** $:\iff \forall \varepsilon > 0\, \exists \delta > 0\, \forall x, x' \in X : (d_X(x,x') \leq \delta \implies d_Y(f(x), f(x')) \leq \varepsilon)$

- Es gilt: Lipschitz-Stetigkeit $\implies$ gleichmäßige Stetigkeit $\implies$ Stetigkeit.

- Analog können wir den Grenzwert einer Abbildung $f : D \to Y$ ($D \subset X$) in einem Berührpunkt a von $D \setminus \{a\}$ und – im Falle der Existenz – die stetige Abänderung/Fortsetzung von f in a definieren.

Gleichmäßige Konvergenz: Seien (X, d_X) und (Y, d_Y) metrische Räume.

- **Definition [gleichmäßige Konvergenz]**: Eine Folge $f_k : X \to Y$ von Abbildungen zwischen metrischen Räumen (X, d_X) , (Y, d_Y) heißt **gleichmäßig konvergent** gegen $f : X \to Y$, falls $\lim\limits_{k\to\infty} \big(\sup\limits_{x \in X} d_Y(f_k(x), f(x)) \big) = 0$ erfüllt ist.

- **Satz 9 [Forster II, §2]**: Die Grenzabbildung $f(x) = \lim\limits_{k\to\infty} f_k(x)$ einer gleichmäßig konvergenten Folge $f_k : X \to Y$ stetiger Abbildungen ist selbst wieder stetig.

Stetige lineare Abbildungen: Seien $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ normierte Vektorräume.

- **Satz 10 [Forster II, §2]**: Eine lineare Abbildung $A : X \to Y$ ist stetig $\iff$

$$\exists L < \infty \ \forall x \in X \ : \ \|Ax\|_Y \ \leq \ L\|x\|_X \ . \tag{2.1}$$

- Insbesondere gelten für eine lineare Abbildung $A : X \to Y$ folgende Äquivalenzen:

$$A \text{ stetig} \iff A \text{ stetig in } 0_X \iff \exists L < \infty \ \forall x \in X : \|Ax\|_Y \leq L\|x\|_X \iff A \text{ glm. stetig}$$

- Die Menge $L(X,Y)$ aller stetigen linearen Abbildungen bildet einen Vektorraum. Durch

$$\|A\|_{L(X,Y)} \ := \ \sup_{\substack{x \in X \\ x \neq 0_X}} \frac{\|Ax\|_Y}{\|x\|_X} \ = \ \sup_{\substack{x \in X \\ \|x\|_X = 1}} \|Ax\|_Y \tag{2.2}$$

  wird auf $L(X,Y)$ eine Norm definiert, welche **Operatornorm** von $A : X \to Y$ genannt wird, welche genau die kleinstmögliche Lipschitz-Konstante in Gleichung (2.1) darstellt.

- Speziell ist $L(X,X)$ der Vektorraum aller stetigen linearen Abbildungen $A : X \to X$, welche **Endomorphismen** genannt werden. Ist X ein Banach-Raum, dann ist $L(X,X)$ mit der Operatornorm ebenfalls vollständig.

Zusatzaufgabe 2.1: Zeigen Sie:

(a) Jede abgeschlossene Teilmenge $M \subset X$ eines vollständigen metrischen Raumes (X, d) ist mit der eingeschränkten Metrik $d|_{M \times M}$ wieder vollständig. Insbesondere ist jeder abgeschlossene lineare Unterraum eines Banach-Raumes erneut selbst (versehen mit der ursprünglichen Norm) ein Banach-Raum.

(b) Ist umgekehrt (X, d) ein metrischer Raum und $M \subset X$ derart, dass (M, d_M) ($d_M := d\big|_M$) zu einem vollständigen metrischen Raum wird, dann ist M abgeschlossen in (X, d).

(c) Der Rand einer offenen Kugel in einem metrischen Raum (X, d) ist abgeschlossen.

(d) Jede offene Teilmenge $U \neq \emptyset, X$ eines Banach-Raumes $(X, \|\cdot\|)$ ist nicht vollständig.

Lösung zu Zusatzaufgabe 2.1:

(a) Ist $x_k \in M$ eine Cauchy-Folge, dann konvergiert x_k aufgrund der Vollständigkeit von (X, d) gegen einen Punkt $x \in X$. Aber aufgrund der Abgeschlossenheit von M liegt x schon in M (vgl. Satz 2, Forster II, §2), d.h. x_k konvergiert in M. Somit ist M vollständig.

(b) Sei $x \in X$ ein Berührpunkt von M. Dann existiert eine in X konvergente Folge $\{x_k\}_k$ aus M. Insbesondere ist dann $\{x_k\}_k$ eine Cauchy-Folge in M. Da (M, d_M) vollständig ist, konvergiert die Folge $\{x_k\}_k$ bereits in M (also gegen ein $m \in M$). Aufgrund der Eindeutigkeit des Grenzwertes muss nun aber $x = m$, also $x \in M$ gelten. Also enthält M alle seine Berührpunkte in X und ist somit eine abgeschlossene Teilmenge von X.

(c) Nach Beispiel (2.2) im Forster II, §2, ist die Abbildung $f : X \to \mathbb{R}$, definiert durch $f_a(x) := d(x, a)$ für ein festes $a \in X$, stetig. Da sich Rand einer offenen Kugel als Urbild $\partial B_r(a) = \{x \in X \mid d(x, a) = r\} = f_a^{-1}(\{r\})$ der einelementigen (und daher in $(\mathbb{R}, d_{|.|})$ abgeschlossenen) Menge $\{r\} \subset \mathbb{R}$ schreiben läss t, folgt nun auch die Abgeschlossenheit von $\partial B_r(a)$ (vgl. Corollar zu Satz 11).

(d) Der Rand von U ist nicht leer, also gibt es eine Folge $x_n \in U$ mit $x_n \to x$ in X für ein $x \in \partial U \setminus U$, und diese konvergiert nicht in U.

Zusatzaufgabe 2.2:

(a) Sei $r > 0$ eine reelle Zahl und $M \subset B_r(0) \subset \mathbb{R}^2$ eine abgeschlossene konvexe Menge, so dass $0 \in \overset{\circ}{M}$ und $x \in M \iff -x \in M$ gilt. Zeigen Sie:

   (i) Für jedes $x \neq 0$ existiert $\max\{\lambda \mid \lambda x \in M\} > 0$.

   (ii) Durch $\|x\|_M := \begin{cases} \frac{1}{\max\{\lambda \mid \lambda x \in M\}} & \text{für } x \neq 0, \\ 0 & \text{für } x = 0, \end{cases}$ wird eine Norm auf $\mathbb{R}^2$ definiert.

   (iii) Jede beliebige Norm auf $\mathbb{R}^2$ wird durch ein geeignetes M auf diese Weise induziert.

(b) Zeigen Sie: Die offene Einheitskugel $B_1(0)$ in einem normierten Vektorraum $(X, \|.\|)$ ist homöomorph zu X.

Lösung zu Zusatzaufgabe 2.2:

(a) (i) Zunächst gilt $\{\lambda \mid \lambda x \in M\} \neq \emptyset$, da $0 \in \overset{\circ}{M}$ gilt. Betrachte nun die stetige Funktion $g : \lambda \mapsto \lambda x$, dann enthält $g^{-1}\big(\overset{\circ}{M}\big)$ als offene Umgebung von $0 \in \mathbb{R}$ ein offenes Intervall um $0 \in \mathbb{R}$. Somit existiert ein $\lambda > 0$ mit $g(\lambda) = \lambda x \in \overset{\circ}{M}$, also gilt $\sup\{\lambda \mid \lambda x \in M\} > 0$. Nun ist noch zu zeigen, dass das Supremum angenommen wird.

Dazu beobachte man, dass M wegen $M \subset B_r(0)$ für ein festes $r > 0$ beschränkt ist. Also ist $\{\lambda \mid \lambda x \in M\}$ beschränkt, denn aus $\|\lambda x\| \leq r$ folgt $|\lambda| \leq \frac{r}{\|x\|}$. Darüber hinaus ist $g^{-1}(M) = \{\lambda \mid \lambda x \in M\}$ abgeschlossen als Urbild einer abgeschlossenen Menge unter einer stetigen Funktion. Daher ist das Supremum bereits ein Maximum.

(ii) Die Definitheit folgt aus (i). Die Homogenität folgt für positive μ aus

$$\max\{\lambda \mid \lambda\mu x \in M\} = \max\left\{\frac{\lambda\mu}{\mu} \mid \lambda\mu x \in M\right\} = \frac{1}{\mu}\max\{\tilde{\lambda} \mid \tilde{\lambda} x \in M\}$$

und überträgt sich auf negative μ aufgrund von $x \in M \iff -x \in M$. Zur Dreiecksungleichung: Nach Definition gilt $\frac{x}{\|x\|} \in M$ und $\frac{y}{\|y\|} \in M$, also $x + y \in (\|x\|M + \|y\|M)$. Aufgrund der Konvexität von M gilt

$$\frac{\|x\|}{\|x\| + \|y\|}M + \frac{\|y\|}{\|x\| + \|y\|}M \subset M$$

und daher $\|x\|M + \|y\|M \subset (\|x\| + \|y\|)\, M$. Somit ergibt sich $x+y \in (\|x\| + \|y\|)\, M$, also

$$\frac{1}{\|x\| + \|y\|}\,(x + y) \in M\ .$$

Dies bedeutet aber genau $\frac{1}{\|x+y\|} = \max\{\lambda \mid \lambda(x+y) \in M\} \geq \frac{1}{\|x\|+\|y\|}$ und darum folgt $\|x + y\| \leq \|x\| + \|y\|$, wie behauptet.

(iii) Sei $\|.\|_*$ eine beliebige Norm. Wählen wir nun $M := \overline{B_1(0)} = \{x \in \mathbb{R}^2 \mid \|x\|_* \leq 1\}$, dann ist – wie oben gesehen – die in (ii) zu M definierte Norm durch die Inklusion $\frac{x}{\|x\|} \in M$ charakterisiert (d.h., eindeutig bestimmt). Die ursprüngliche Norm $\|.\|_*$ erfüllt aufgrund der Homogenität jedoch ebenso $\frac{x}{\|x\|_*} \in M$. Also stimmen $\|.\|_*$ und $\|.\|$ überein.

(b) Die offene Kugel $B_1(0)$ in einem normierten Vektorraum $(X, \|.\|)$ ist homöomorph zu X, denn die Abbildung $f : B_1(0) \to X$, $x \mapsto \frac{x}{1-\|x\|}$ ist

- stetig, da die Identität und die Norm stetig sind und $\|x\| < 1$ für $x \in B_1(0)$ gilt,
- bijektiv, da $f(x) = y \iff x = \frac{y}{1+\|y\|}$ für $x \in B_1(0)$ und $y \in X$ gilt,
- und sogar ein Homöomorphismus, da auch die Umkehrabbildung $f^{-1}(y) := \frac{y}{1+\|y\|}$ stetig ist (aufgrund der Stetigkeit der Identität und der Norm).

Zusatzaufgabe 2.3:

(a) Die Menge $A := S^1 \cap H_+$ mit der Sphäre $S^1 := \{(x, y) \in \mathbb{R}^2 \mid \|(x, y)\|_2 = 1\}$ und dem Halbraum $H_+ := \{(x, y) \in \mathbb{R}^2 \mid y \geq 0\}$ ist abgeschlossen in $\left(\mathbb{R}^2, d_{\|\cdot\|_2}\right)$.

(b) Die Abbildung $f : \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\to \{(x, y) \in S^1 \mid x > 0\}$, $t \mapsto (\cos(t), \sin(t))$, ist ein Homöomorphismus.

(c) Die Abbildung $f : [0, 2\pi[\to S^1$, $t \mapsto (\cos(t), \sin(t))$, ist kein Homöomorphismus.

Lösung zu Zusatzaufgabe 2.3:

(a) Da sowohl die Norm als auch die Projektion stetige Funktionen sind, ist sowohl S^1 als Urbild der einelementigen (und somit abgeschlossenen) Menge $\{1\}$ unter einer stetigen Funktion als auch H_+ als Urbild des in $\mathbb{R}$ abgeschlossenen Intervalls $[0, \infty[$ selbst wieder abgeschlossen (vgl. Corollar zu Satz 11). Somit ist A als endlicher Schnitt abgeschlossener Mengen selbst abgeschlossen in $(\mathbb{R}^2, \|\cdot\|_2)$.

(b) Es ist zu zeigen, dass f stetig und bijektiv ist und dass die Umkehrabbildung ebenfalls stetig ist. Die Stetigkeit von $f : t \mapsto (\cos(t), \sin(t))$ auf $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ folgt nach Satz 6, Forster II, §2 (wenn man f als Abbildung von $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ nach $\mathbb{R} \times \mathbb{R}$ betrachtet), da die Komponentenfunktionen stetig sind. Wegen $(\cos(t))^2 + (\sin(t))^2 = 1$ und $\cos(t) > 0$ auf $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ gilt wie behauptet $f\left(\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\right) = \{(x,y) \in S^1 \mid x > 0\}$, also $f : \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\to \{(x,y) \in S^1 \mid x > 0\}$ surjektiv. Aufgrund der Monotonie und Stetigkeit von $\sin(t)$ auf $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ folgt ebenso die Injektivität, also insgesamt $f : \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\to \{(x,y) \in S^1 \mid x > 0\}$ bijektiv. Da wir die Umkehrabbildung $g = f^{-1} : \{(x,y) \in S^1 \mid x > 0\} \to \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ direkt angeben können mit $g(x,y) = \arctan\left(\frac{y}{x}\right)$ und diese als Verkettung stetiger Funktionen selbst wieder stetig ist, haben wir nun gezeigt, dass f zu einem Homöomorphismus wird.

(c) Analog zu (b) können wir zwar zeigen, dass f stetig und bijektiv ist. Die Umkehrabbildung kann nach Satz 11, Forster II, §2, aber nicht stetig sein, da das Urbild jeder (genügend kleinen) offenen Umgebung um den Punkt $(1,0) \in S^1$ die Gestalt $[0, \varepsilon[\;\cup\;]2\pi - \varepsilon, 2\pi[$ besitzt und damit nicht offen sein kann (0 ist kein innnerer Punkt).

Zusatzaufgabe 2.4:

(a) Sei (X,d) ein metrischer Raum und seien $f, g : X \to \mathbb{R}$ stetig. Zeigen Sie: Die Funktionen

$$\text{(i)}\quad x \mapsto \max(f(x), g(x)) \qquad \text{und} \qquad \text{(ii)}\quad x \mapsto \begin{cases} \frac{f(x)^2}{\sqrt{f(x)^2 + g(x)^2}} & , (f(x), g(x)) \neq (0,0) \\ 0 & , (f(x), g(x)) = (0,0) \end{cases}$$

sind stetig von (X,d) nach $(\mathbb{R}, d_{|.|})$.

(b) Ist die Funktion $f : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$ definiert durch $f(x,y) := \frac{\sin(x^2)}{x^2+y^2}$ beschränkt? Kann man f bezüglich der Euklidischen Norm/Metrik stetig in den Punkt $(0,0)$ fortsetzen?

(c) Für $\left(C^1([-1,1]), \|\,.\,\|_\infty\right)$ und $\left(C([-1,1]), \|\,.\,\|_\infty\right)$ betrachten wir die durch $A(f) := f'$, $f \in C^1([-1,1])$, definierte lineare Abbildung $A : C^1([-1,1]) \to C([-1,1])$. Untersuchen Sie, ob A stetig ist.

Lösung zu Zusatzaufgabe 2.4:

(a) (i) Da $f, g : X \to \mathbb{R}$ nach Voraussetzung beide stetig sind, folgt aus der Konvergenz $y \to x$ insbesondere auch $f(y) \to f(x)$ als auch $g(y) \to g(x)$. Weiter gilt

$$\max(f(x), g(x)) = \frac{1}{2}|f(x) - g(x)| + \frac{1}{2}(f(x) + g(x)),$$

so dass wir unter Berücksichtigung von $\big||a| - |b|\big| \leq |a - b|$ die Abschätzung

$$\begin{aligned} \Big|\max(f,g)(x) - \max(f,g)(y)\Big| &= \frac{1}{2}\Big||f(x) - g(x)| + f(x) + g(x) - |f(y) - g(y)| - f(y) - g(y)\Big| \\ &\leq \frac{1}{2}\left(\Big||f(x) - g(x)| - |f(y) - g(y)|\Big| + |f(x) - f(y)| + |g(x) - g(y)|\right) \\ &\leq \frac{1}{2}\left(\Big|(f(x) - g(x)) - (f(y) - g(y))\Big| + |f(x) - f(y)| + |g(x) - g(y)|\right) \\ &\leq |f(x) - f(y)| + |g(x) - g(y)| \end{aligned}$$

finden. Demzufolge impliziert $y \to x$ auch $\max(f,g)(y) \to \max(f,g)(x)$, also die Stetigkeit von $\max(f,g)$ in x. Da $x \in X$ beliebig war, folgt somit die Behauptung.

(ii) Zunächst folgt nach Lemma 1.39 die Stetigkeit von $x \mapsto (f(x), g(x))$. Weiterhin ist die Funktion

$$(u,v) \mapsto \begin{cases} \frac{u^2}{\sqrt{u^2+v^2}} & ,\ (u,v) \neq (0,0) \\ 0 & ,\ (u,v) = (0,0) \end{cases}$$

von $\mathbb{R}^2$ nach $\mathbb{R}$ ist stetig, da sie einerseits in jedem Punkt $(u,v) \neq (0,0)$ als Quotient der stetigen Funktionen $(u,v) \mapsto u^2$ und $(u,v) \mapsto \|(u,v)\|_2$ stetig ist und andererseits wegen $\frac{u^2}{\sqrt{u^2+v^2}} \leq \frac{u^2}{\sqrt{u^2}} \leq |u|$ im Punkt $(0,0)$ stetig ist, weil diese Ungleichung insbesondere

$$\frac{u^2}{\sqrt{u^2+v^2}} \to 0$$

für $(u,v) \mapsto 0 \in \mathbb{R}^2$ impliziert. Als Komposition von stetigen Abbildungen ist somit insbesondere

$$x \mapsto (f(x), g(x)) \mapsto \frac{f(x)^2}{\sqrt{f^2(x) + g^2(x)}}$$

auf ganz X stetig.

(b) Es gilt $|\sin(x^2)| \leq x^2 \leq x^2 + y^2$, also ist $|f(x,y)| \leq 1$ für alle $(x,y) \neq (0,0)$, und somit ist f beschränkt. Aber f kann nicht stetig in $(0,0)$ fortgesetzt werden, denn einerseits gilt nach den Regeln von L'Hospital

$$\lim_{x\to 0} f(x,0) \;=\; \lim_{x\to 0} \frac{\sin(x^2)}{x^2} \;=\; \lim_{x\to 0} \frac{2x\cos(x^2)}{2x} \;=\; 1,$$

andererseits aber $\lim_{y\to 0} f(0,y) = 0$.

(c) Offenbar ist die Abbildung A bezüglich der angegebenen Räume nicht stetig, da beispielsweise für die Funktionenfolgen $f_n(x) := x^n$ oder $g_n(x) := \sin(nx)$ zwar jeweils

$$\| f_n \|_\infty \;:=\; \max_{x\in[-1,1]} |f_n(x)| = 1\,, \quad \text{bzw.} \quad \| g_n \|_\infty \;:=\; \max_{x\in[-1,1]} |g_n(x)| = 1\,,$$

aber auch

$$\| Af_n \|_\infty \;=\; \| f_n' \|_\infty \;=\; \max_{x\in[-1,1]} |nx^{n-1}| \;=\; n \;\overset{n\to\infty}{\longrightarrow}\; \infty$$
$$\| Ag_n \|_\infty \;=\; \| g_n' \|_\infty \;=\; \max_{x\in[-1,1]} |n\cos(nx)| \;=\; n \;\overset{n\to\infty}{\longrightarrow}\; \infty$$

gelten. Somit kann es keine Konstante $C < \infty$ geben, für die gilt, dass

$$\forall f \in C^1([-1,1]) \;:\; \| Af \|_\infty \;\leq\; C \cdot \| f \|_\infty\;.$$

Zusatzaufgabe 2.5:

Jeder metrische Raum kann isometrisch in einem vollständigen metrischen Raum eingebettet werden. Genauer gesagt existiert zu jedem metrischen Raum (X,d) ein vollständiger metrischer Raum $(\tilde{X},\tilde{d})$ und eine Abbildung $\iota : X \to \tilde{X}$, so dass $\iota(X)$ dicht in $\tilde{X}$ liegt, d.h. $\overline{\iota(X)} = \tilde{X}$, und $\tilde{d}(\iota(x),\iota(y)) = d(x,y)$ gilt.

Lösung zu Zusatzaufgabe 2.5:

Wir betrachten die Menge aller Cauchy-Folgen in X und nennen zwei Cauchy-Folgen x_k und y_k äquivalent, wenn $\lim_{k\to\infty} d(x_k,y_k) = 0$ gilt. Dies definiert tatsächlich eine Äquivalenzrelation $\sim$ auf der Menge aller Cauchy-Folgen wegen

- $x_k \sim x_k$ aufgrund von $\lim_{k\to\infty} d(x_k,x_k) = 0$

- $(x_k \sim y_k) \Rightarrow (y_k \sim x_k)$ aufgrund von $\lim\limits_{k\to\infty} d(x_k, y_k) = \lim\limits_{k\to\infty} d(y_k, x_k)$
- $((x_k \sim y_k) \wedge (y_k \sim z_k)) \Rightarrow (x_k \sim z_k)$ aufgrund der Ungleichung $0 \leq d(x_k, z_k) \leq d(x_k, y_k) + d(y_k, z_k) \to 0$ bei $k \to \infty$ und somit $\lim\limits_{k\to\infty} d(x_k, z_k) = 0$.

Nun sei $\tilde{X}$ die Menge aller Äquivalenzklassen $[x_k]$ von Cauchy-Folgen x_k bezüglich der soeben definierten Äquivalenzrelation $\sim$. Auf $\tilde{X}$ kann man dann eine Metrik durch $\tilde{d}([x_k], [y_k]) := \lim\limits_{k\to\infty} d(x_k, y_k)$ definieren.

Tatsächlich existiert dieser Limes in $\mathbb{R}$ für beliebige Cauchy-Folgen x_k und y_k in X, denn wegen $|d(x_k, y_k) - d(x_l, y_l)| \leq d(x_k, x_l) + d(y_k, y_l)$ ist die Folge $d(x_k, y_k)$ eine Cauchy-Folge in $\mathbb{R}$. Darüberhinaus hängt der Grenzwert nicht von der Wahl der Repräsentanten ab, denn gilt $x_k \sim \tilde{x}_k$ und $y_k \sim \tilde{y}_k$, dann folgt wegen

$$|d(\tilde{x}_k, \tilde{y}_k) - d(x_k, y_k)| \leq d(x_k, \tilde{x}_k) + d(y_k, \tilde{y}_k) \to 0$$

für $k \to \infty$ die Gleichheit $\lim\limits_{k\to\infty} d(\tilde{x}_k, \tilde{y}_k) = \lim\limits_{k\to\infty} d(x_k, y_k)$, also ist $\tilde{d}$ wohldefiniert. Desweiteren ist $\tilde{d}$ eine Metrik auf $\tilde{X}$, denn aus $\tilde{d}([x_k], [y_k]) = 0$ folgt $x_k \sim y_k$ nach Definition von $\sim$ und somit $[x_k] = [y_k]$, und die Symmetrie sowie die Dreiecksungleichung für d übertragen sich direkt auf $\tilde{d}$.

Um zu zeigen, dass $(\tilde{X}, \tilde{d})$ vollständig ist, betrachte man eine Cauchy-Folge $[x_k]_n$ in $\tilde{X}$, d.h. eine Folge von Äquivalenzklassen von Cauchy-Folgen in X, für die zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert mit $\tilde{d}([x_k]_m, [x_k]_n) \leq \varepsilon/3$ für alle $m, n \geq N$. Wählt man sich zu jedem n einen Repräsentanten x_{kn} von $[x_k]_n$, so ist für festes n die Folge $k \mapsto x_{kn}$ ja eine Cauchy-Folge, also gibt es ein $K(n) \in \mathbb{N}$ mit $d(x_{kn}, x_{ln}) \leq \varepsilon/3$ für $k, l \geq K(n)$.

Um sich einen Grenzwert der Cauchy-Folge $[x_k]_n$ in $\tilde{X}$ zu verschaffen, betrachte man einfach die Folge $y_n := x_{K(n)n}$. Diese Folge y_n ist eine Cauchy-Folge in X, denn seien $m, n \geq N$, dann gibt es nach Definition von $\tilde{d}$ ein $l \geq K(n), K(m)$ mit $d([x_{lm}, x_{ln}) \leq \varepsilon/3$, und mit diesem gilt

$$d(y_m, y_n) = d(x_{K(m)m}, x_{K(n)n}) \leq d(x_{K(m)m}, x_{lm}) + d(x_{lm}, x_{ln}) + d(x_{ln}, x_{K(n)n}) \leq \varepsilon \quad .$$

Und tatsächlich konvergiert die Cauchy-Folge $[x_k]_n$ in $\tilde{X}$ auch gegen die Äquivalenzklasse $[y_k]$. Denn da (wie gerade nachgewiesen) y_k eine Cauchy-Folge in X ist, gibt es ein M mit $d(y_n, y_k) \leq \frac{2}{3}\varepsilon$ für $n, k \geq M$. Sei nun $n \geq M$, dann gilt für jedes $k \geq M, K(n)$ die Ungleichung

$$d(x_{kn}, y_k) \leq d(x_{kn}, y_n) + d(y_n, y_k) = d(x_{kn}, x_{K(n)n}) + d(y_n, y_k) \leq \varepsilon \quad ,$$

weil $d(x_{kn}, x_{K(n)n}) \leq \varepsilon/3$ aufgrund der Wahl von $k \geq K(n)$ gilt. Also gilt $\tilde{d}([x_k]_n, [y_k]) = \lim\limits_{k\to\infty} d(x_{kn}, y_k) \leq \varepsilon$ für jedes $n \geq M$, d.h. die Cauchy-Folge $[x_k]_n$ konvergiert in $\tilde{X}$ gegen $[y_k]$. Somit ist $(\tilde{X}, \tilde{d})$ vollständig.

Zu guter letzt kann man noch (X, d) in den so gewonnenen metrischen Raum $(\tilde{X}, \tilde{d})$ isometrisch einbetten, indem man einen Punkt x mit der Äquivalenzklasse der konstanten Folge $\iota(x) := x, x, \ldots$ identifiziert. Die so definierte Abbildung $\iota : X \to \tilde{X}$ erfüllt offensichtlich $\tilde{d}(\iota(x), \iota(y)) = d(x, y)$, d.h. ι ist eine Isometrie und insbesondere injektiv. Somit hat X aufgefasst als Teilmenge $\iota(X)$ von $\tilde{X}$ dieselben metrischen Eigenschaften wie vorher. Darüberhinaus liegt X dicht in $\tilde{X}$, d.h. es gibt zu jedem $[x_k] \in \tilde{X}$ und $\varepsilon > 0$ ein $\iota(x)$ mit $\tilde{d}([x_k], \iota(x)) \leq \varepsilon$, denn sei K so groß, dass $d(x_k, x_l) \leq \varepsilon$ für alle $k, l \geq K$ gilt, dann erfüllt $x := x_K$ diese Bedingung.

Zusatzmaterial zur Analysis II – Serie 3

Kompaktheit

- **Definition [offene Überdeckung]**: Sei $(X, \mathcal{T})$ ein topologischer Raum. Eine Familie $(U_i)_{i\in I}$ (mit einer beliebig großen Indexmenge I) von offenen Mengen $U_i \in T$ heißt **offene Überdeckung** von $A \subset X$, falls gilt
$$A \subset \bigcup_{i\in I} U_i \ .$$
- **Definition [Kompaktheit]**: Sei $(X, \mathcal{T})$ ein topologischer Raum. Eine Teilmenge $K \subset X$ heißt **(überdeckungs)kompakt**, falls zu jeder offenen Überdeckung $(U_i)_{i\in I}$ von K eine endliche Teilüberdeckung existiert, d.h., falls mit endlich vielen Indizes $i_1, \ldots, i_m \in I$ bereits
$$X \subset \bigcup_{j=1}^{m} U_{i_j}$$
erfüllt ist.
- In metrischen Räumen (X, d) ist die Kompaktheit einer Teilmenge $K \subset X$ äquivalent dazu, dass jede Folge in K eine konvergente Teilfolge besitzt[1] (vgl. Satz 8, Forster, Analysis 2 und Aufgabe 3.1 (b)). Letztere Eigenschaft wird auch **Folgen-Kompaktheit** genannt.
- Insbesondere folgt, dass eine kompakte Teilmenge eines metrischen Raumes (X, d) beschränkt und abgeschlossen sein muss (Satz 3). Die Umkehrung gilt jedoch i.A. nicht.[2]
- **Satz 4**: Sei (X, d) ein metrischer Raum und $K \subset X$ eine kompakte Teilmenge. Dann ist jede in (X, d) abgeschlossene Teilmenge $A \subset K$ wieder kompakt.
- **Satz 5 [Heine-Borel]**: In endlich-dimensionalen normierten Räumen $(X, \|\cdot\|)$ ist eine Teilmenge $K \subset X$ genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.
- Stetige Abbildungen zwischen metrischen Räumen erhalten Kompaktheit in dem Sinne, als dass Bilder $f(K) \subset Y$ von kompakten Mengen $K \subset X$ unter einer stetigen Abbildung $f : X \to Y$ wieder kompakt sind (Satz 6). Eine direkte Folge davon ist, dass im Falle $Y = \mathbb{R}$ die stetige Abbildung $f : X \to \mathbb{R}$ auf jedem Kompaktum $K \subset X$ ein Maximum und ein Minimum annimmt (Satz 7).
- Sind $(X, d_X), (Y, d_Y)$ metrische Räume und X kompakt, dann ist jede stetige Abbildung $f : X \to Y$ gleichmäßig stetig (Satz 9). Ist f zusätzlich bijektiv, dann ist auch ihre Umkehrabbildung $f^{-1} : Y \to X$ stetig, d.h. jede stetige Bijektion von einem kompakten metrischen Raum in einen metrischen Raum bildet einen Homöomorphismus (vgl. Zusatzaufgabe 3.2).

Zusatzaufgabe 3.1:

(a) Beweisen Sie mittels der Definition, dass die Menge $A := \left\{\frac{1}{n} \middle| \, n \in \mathbb{N}\right\} \cup \{0\}$ als Teilmenge des Euklidischen $\mathbb{R}$ (d.h., im metrischen Raum $(\mathbb{R}, d_{|.|})$) kompakt ist.

(b) Zeigen Sie, dass jede endliche Vereinigung von kompakten Teilmengen eines metrischen Raumes (X, d) wieder kompakt ist.

[1] Diese Eigenschaft ist auch als Satz von Bolzano-Weierstraß bekannt.
[2] Siehe auch Beispiel (3.3), Forster, Analysis 2.

Lösung zu Zusatzaufgabe 3.1:

(a) Es ist zu zeigen, dass zu einer beliebigen offenen Überdeckung $\{U_i\}_{i\in I}$ von A eine endliche Teilüberdeckung existiert. Sei dazu $\{U_i\}_{i\in I}$ eine beliebige offene Überdeckung von A. Dann existiert zu jedem $n \in \mathbb{N}$ ein U_{i_n}, so dass $1/n \in U_{i_n}$. Weiterhin gibt es auch ein U_{i_0}, so dass $\{0\} \in U_{i_0}$. Da U_{i_0} offene Teilmenge in $\mathbb{R}$ bildet, existiert ein $\varepsilon > 0$, so dass $B_\varepsilon(0) \subset U_{i_0}$. Da weiterhin $\{1/n\}_{n\in\mathbb{N}}$ ein Nullfolge in $\mathbb{R}$ bildet, existiert zu diesem $\varepsilon > 0$ ein $N(\varepsilon) \in \mathbb{N}$, so dass $n \geq N(\varepsilon)$ unmittelbar $1/n \in B_\varepsilon(0)$ impliziert. Demzufolge enthält U_0 bereits alle bis auf endlich viele der Folgenglieder $\frac{1}{n}$, so dass bereits

$$A \subset \bigcup_{k=0}^{N-1} U_{i_k}$$

gilt und die Mengen $U_{i_0}, \ldots, U_{i_{N-1}}$ damit eine endliche Teilüberdeckung von A bilden. Da $\{U_i\}_{i\in I}$ beliebig war, folgt daraus die Kompaktheit der Menge A.

(b) Sei $\{U_i\}_{i\in I}$ eine beliebige offene Überdeckung der endlichen Vereinigung

$$A = \bigcup_{l=1}^{n} K_l$$

von kompakten Teilmengen K_l eines metrischen Raumes X. Dann ist dies insbesondere auch für jedes einzelne K_l eine offene Überdeckung. Aufgrund der Kompaktheit der Mengen K_l finden wir jeweils endlich viele Indizes $i_{l1}, \ldots, i_{lk_l}$, für die

$$\bigcup_{j=1}^{k_l} U_{i_{lj}} \supset K_l, \quad l = 1, \ldots, n,$$

gilt. Demnach bildet $\{U_{i_{lj}}\}_{(l,j)\in\{1,\ldots,n\}\times\{1,\ldots,k_l\}}$ eine endliche offene Überdeckung für jedes K_l, und somit gilt auch

$$\bigcup_{l=1}^{n} \left(\bigcup_{j=1}^{k_l} U_{i_{lj}} \right) \supset A,$$

weil für jedes $a \in A$ ein $l \in \{1, \ldots, n\}$ existiert, so dass $a \in K_l$. Auf diese Weise können wir zu jeder beliebigen offenen Überdeckung von A eine endliche Teilüberdeckung von A konstruieren, was gerade der Definition von Kompaktheit entspricht.

Zusatzaufgabe 3.2:

Zeigen Sie: Sind $(X, d_X), (Y, d_Y)$ metrische Räume und X kompakt, dann ist jede bijektive stetige Abbildung $f : X \to Y$ ein Homöomorphismus.

Lösung zu Zusatzaufgabe 3.2:

Es ist zu zeigen, dass $f^{-1} : Y \to X$ stetig ist. Da die Urbilder von f^{-1} genau die Bilder von f sind, müssen wir nur zeigen, dass das Bild einer offenen Menge unter f offen ist, oder äquivalenterweise, dass das Bild einer abgeschlossenen Menge $A \subset X$ unter f abgeschlossen ist. Nun ist A nach Satz 4 kompakt, also auch $f(A)$ als Bild eines Kompaktums unter einer stetigen Abbildung nach Satz 6 kompakt in Y, und somit ist insbesondere $f(A)$ in Y wiederum nach Satz 4 abgeschlossen, was zu zeigen war.

Zusatzaufgabe 3.3:

(a) Zeigen Sie, dass die Menge $A := \left\{3 - \frac{5}{n} \;\middle|\; n \in \mathbb{N}\right\}$ als Teilmenge des Euklidischen $\mathbb{R}$ nicht kompakt sein kann, indem Sie eine offene Überdeckung von A angeben, aus der keine endliche Teilüberdeckung ausgewählt werden kann.

(b) Sei A Teilmenge eines mit der diskreten Metrik versehenen metrischen Raumes (M, d_{diskret}). Zeigen Sie: A ist genau dann kompakt, wenn A nur endlich viele Elemente enthält.

(c) (**Satz von Dini**) Sei (X, d) ein kompakter metrischer Raum, $f_n : X \to \mathbb{R}$ stetig und $f_{n+1} \leq f_n$ für $n \in \mathbb{N}$. Die Folge (f_n) konvergiere punktweise gegen 0.
Zeigen Sie, dass dann (f_n) gleichmäßig gegen die Nullfunktion konvergiert.

Lösung zu Zusatzaufgabe 3.3:

(a) Wegen $n-1 < n < n+1 \iff \frac{5}{n+1} < \frac{5}{n} < \frac{5}{n-1} \iff 3 - \frac{5}{n-1} < 3 - \frac{5}{n} < 3 - \frac{5}{n+1}$ haben wir durch $\{A_n \ : \ n \in \mathbb{N}\}$ mit $A_1 :=]-3, \frac{1}{2}[$ und $A_n := \left]3 - \frac{5}{n-1}, 3 - \frac{1}{n+1}\right[$ für $n \geq 2$ eine offene Überdeckung von A bestehend aus abzählbar unendlich vielen offenen Mengen gegeben, von der aufgrund von $A \cap A_n = \left\{3 - \frac{5}{n}\right\}$ keine Menge weggelassen werden kann. Demzufolge existiert zu der offenen (unendlichen) Überdeckung $\{A_n \ : \ n \in \mathbb{N}\}$ von A keine endliche Teilüberdeckung, d.h., insbesondere kann A nicht kompakt sein.

(b) Nach Definition ist A genau dann kompakt, wenn zu jeder offenen Überdeckung $(U_i)_{i \in I}$ eine endliche Teilüberdeckung existiert.

„$\Leftarrow$“ Habe nun A endlich viele Elemente, d.h. $A = \{a_k \in M \ : \ 1 \leq k \leq n\}$ für ein $n \in \mathbb{N}$. Dann existiert in jeder offenen Überdeckung $(U_i)_{i \in I}$ von A zu jedem der a_k ein $U_{i_{a_k}}$ mit $a_k \in U_{i_{a_k}}$, so dass offenbar $\left\{U_{i_{a_k}} \ : \ 1 \leq k \leq n\right\}$ eine endliche Teilüberdeckung von A ist, d.h. $A \subset \bigcup_{k=1}^{n} U_{i_{a_k}}$. Somit ist A kompakt.

„$\Rightarrow$“ Sei nun A kompakt. Dann existiert zu der offenen Überdeckung $\left\{B_{\frac{1}{2}}(a) \ : \ a \in A\right\}$ mit

$$B_{\frac{1}{2}}(a) := \left\{x \in M \ : \ d_{\text{diskret}}(a, x) < \tfrac{1}{2}\right\}$$

aller offenen Kugeln mit Radius $\frac{1}{2}$ und Mittelpunkt $a \in A$ eine endliche Teilüberdeckung, etwa $\left\{B_{\frac{1}{2}}(a_k) \ : \ 1 \leq k \leq n\right\}$ für ein $n \in \mathbb{N}$. Wegen $d_{\text{diskret}}(a, b) = 1 > \frac{1}{2}$ für $a \neq b$ enthält jede dieser offenen Umgebungen $B_{\frac{1}{2}}(a_k)$ jedoch nur jeweils den Mittelpunkt a_k. Demzufolge kann die Teilmenge

$$A \subset \bigcup_{k=1}^{n} B_{\frac{1}{2}}(a_k) = \{a_k \ : \ 1 \leq k \leq n\}$$

nur endlich viele Elemente enthalten.

(c) Sei $\varepsilon > 0$ beliebig, jedoch fest gewählt. Dann gibt es aufgrund der punktweisen Konvergenz von (f_n) gegen 0 zu jedem $\xi \in X$ ein $m = m(\varepsilon, \xi) \in \mathbb{N}$, für dass $|f_m(\xi)| < \varepsilon$. Aufgrund der Stetigkeit der Funktionen f_n ist auch $|f_m|$ stetig, so dass ein $\delta = \delta(\varepsilon, \xi) > 0$ existiert, für dass $|f_m(x)| < \varepsilon$ für alle $x \in B_\delta(\xi) \cap X$ gilt. Aufgrund der monotonen Konvergenz ($0 \leq f_{n+1} < f_n$) gilt sogar $|f_n(x)| < \varepsilon$ für alle $x \in B_\delta(\xi) \cap X$ und alle $n \geq m(\varepsilon, \xi)$.
Das System der $B_{\delta(\varepsilon,\xi)}(\xi)$ $(\xi \in X)$ bildet eine offene Überdeckung von X. Da X nach Voraussetzung kompakt ist, können wir endlich viele Punkte $\xi_1, \ldots, \xi_k$ finden, so dass die zugehörigen Umgebungen $U^{(k)} := B_{\delta(\varepsilon,\xi_k)}(\xi_k)$ bereits X überdecken.
Mit $n_0 := \max\{m(\varepsilon, \xi_1), \ldots, m(\varepsilon, \xi_k)\}$ gilt somit für jedes beliebige $x \in X$ (welches sich in mindestens einem der $U^{(k)}$ befindet) $|f_n(x)| < \varepsilon$ für alle $n \geq n_0$.
Da n_0 konstruktionsgemäß nicht von x abhängig ist (sondern nur von ε), bedeutet dies genau die gleichmäßige Konvergenz der Folge (f_n) auf X gegen 0.

Zusatzaufgabe 3.4:

(a) Zeigen Sie oder widerlegen Sie: Die Menge $M := \{(x, y, z) \mid x^2 + y^2 + yz + z^2 = 1\}$ ist eine kompakte Teilmenge im metrischen Raum $(\mathbb{R}^3, d_{\|\cdot\|})$.

(b) Ist die Funktion $f : M \to \mathbb{R}$, $(x, y, z) \mapsto x^2 + y^2 + z^2$ mit M aus (a) beschränkt ? Falls ja, nimmt sie ihr Maximum an ?

Lösung zu Zusatzaufgabe 3.4:

(a) Mit der stetigen Funktion $g : \mathbb{R}^3 \to \mathbb{R}$, definiert durch $g(x,y,z) = x^2 + y^2 + yz + z^2$, ist $M = g^{-1}(\{1\})$ als Urbild einer einelementigen (und daher trivialerweise abgeschlossenen) Menge unter einer stetigen Funktion abgeschlossen in $(\mathbb{R}^3, d_{\|.\|})$. Wegen

$$2g(x,y,z) \;=\; 2x^2{+}y^2{+}(y{+}z)^2{+}z^2 \;\geq\; x^2{+}y^2{+}z^2 \;\geq\; \max\{x^2,y^2,z^2\} \;=\; (\max\{|x|,|y|,|z|\})^2$$

gilt offenbar $2g(x,y,z) \geq 4$ für jeden Punkt (x,y,z) außerhalb der Kugel $B_2((0,0,0))$ bezüglich der Maximumnorm $\|.\|_\infty$, womit dieser nicht in $M = g^{-1}(\{1\})$ liegen kann. Daher folgt $M \subset [-2,2]^3 = \overline{B_2((0,0,0))}$, also die Beschränktheit von M in $(\mathbb{R}^3, d_{\|.\|})$. Da $(\mathbb{R}^3, d_{\|.\|})$ identisch mit dem endlich-dimensionalen normierten Vektorraum $(\mathbb{R}^3, \|.\|)$ ist, folgt nach dem Satz von Heine-Borel somit aus der Beschränktheit und der Abgeschlossenheit von M bereits die Kompaktheit von M in $(\mathbb{R}^3, d_{\|.\|})$.

(b) Da f die Einschränkung der auf dem $\mathbb{R}^3$ definierten stetigen Funktion $(x,y,z) \mapsto \|(x,y,z)\|_2^2$ auf die nach Aufgabenteil (a) kompakte Teilmenge M ist und da nach Satz 6 das Bild $f(M)$ in $(\mathbb{R}, d_{|.|})$ kompakt ist, folgt die Beschränktheit aus Satz 3. Dies ist auch gerade in Satz 7 zusammengefasst, welcher garantiert, dass f auf M sein Maximum annimmt.

Zusatzaufgabe 3.5:

(a) Gegeben sei die Menge $M := \{(x,y,z) \in \mathbb{R}^3 \mid x^2 + y^4 + z^6 = 1 \,,\;\; xy + yz + xz = 0\}$. Zeigen Sie, dass die Funktion $f : M \to \mathbb{R}$, definiert durch $f(x,y,z) := x^2 + y^3 + z^4$ beschränkt ist und dass sie ein Maximum und ein Minimum annimmt.

(b) Entscheiden Sie, ob die Menge $\mathbb{Q}$ in $(\mathbb{R}, d_{|.|})$ kompakt ist.

(c) Auf einem metrischen Raum (X,d) betrachten wir die Teilmenge $A \subset X$. Zu einem beliebigen $x \in X$ wird dann durch

$$\operatorname{dist}(x,A) := \inf_{y \in A} d(x,y)$$

eine stetige Funktion $x \mapsto \operatorname{dist}(x,A)$ definiert (vgl. Beispiel (3.5), Forster Analysis 2, §3). Zeigen oder widerlegen Sie: Sind A_1, A_2 abgeschlossene Mengen eines metrischen Raumes (M,d) mit $A_1 \cap A_2 = \emptyset$. Dann gilt $\operatorname{dist}(A_1, A_2) > 0$.

Lösung zu Zusatzaufgabe 3.5:

(a) Die Funktion $g : \mathbb{R}^3 \to \mathbb{R}^2$, $g(x,y,z) = \begin{pmatrix} x^2+y^4+z^6 \\ xy+yz+xz \end{pmatrix}$ ist stetig und $M = g^{-1}\left(\left\{\begin{pmatrix}1\\0\end{pmatrix}\right\}\right)$ nach Satz 11 (§2 Forster Analysis 2) folglich abgeschlossen. Wegen $x^2{+}y^3{+}z^6 \geq \|(x,y,z)\|_\infty$ folgt $M \subset \{(x,y,z) \in \mathbb{R}^3 \mid \|(x,y,z)\|_\infty \leq 1\}$, also auch die Beschränktheit von M. Der Satz von Heine-Borel (Satz 5, §3 Forster Analysis 2) liefert nun die Kompaktheit von M. Als stetige Funktion auf einem Kompaktum ist f beschränkt und nimmt nach Satz 7 (§3 Forster Analysis 2) sowohl Maximum als auch Minimum an.

(b) Aufgrund der Unbeschränktheit kann $\mathbb{Q}$ nicht kompakt sein (vgl. Satz 3, §3 Forster 2).

(c) Betrachten wir die stetige Funktion $f : (\mathbb{R}^2, d_{\|.\|}) \to (\mathbb{R}, d_{|.|})$, $(x,y) \mapsto xy$. Dann gilt (nach den Eigenschaften des Urbildes) $f^{-1}(\{0\}) \cap f^{-1}(\{1\}) = \emptyset$ und als Urbilder einelementiger (und somit trivialerweise abgeschlossener) Mengen unter einer stetigen Funktion sind beide abgeschlossen in $(\mathbb{R}^2, d_{\|.\|})$ (nach Satz 11, §2 Analysis 2, Forster). Jedoch gilt offenbar $(\frac{1}{n}, n) \in f^{-1}(\{1\})$ und $(0,n) \in f^{-1}(\{0\})$, jedoch $d_{\|.\|}\big((\frac{1}{n}, n), (0,n)\big) = \frac{1}{n} \xrightarrow{n\to\infty} 0$ und somit für $A_1 = f^{-1}(\{0\})$ und $A_2 = f^{-1}(\{0\})$ offenbar $\operatorname{dist}(A_1, A_2) = 0$.

Alle Übungsaufgaben sind verfügbar unter: `https://studip.uni-rostock.de/studip/`

Zusatzmaterial zur Analysis II – Serie 4

Kurven, Parametertransformation, Schnittwinkel, Bogenlänge

- **Definition [Kurve]**: Eine stetige Abbildung $c : I \to X$ eines Intervalls $I \subset \mathbb{R}$ in einen metrischen Raum X nennt man eine **parametrisierte Kurve**. Das Bild $C := c(I) \subset X$ einer parametrisierten Kurve c heißt die **Spur** von c oder einfach auch die **Kurve**, die durch c parametrisiert wird.

  Im Weiteren betrachten wir parametrisierte Kurven in Banach-Räumen $(X, \|\cdot\|)$, wobei $X = \mathbb{R}^n$ den einfachsten und für $n = 2$ oder $n = 3$ den anschaulichsten Fall darstellt.[1]

- **Definition [(stetig) differenzierbare Kurve]**: Eine parametrisierte Kurve $c : I \to X$ in einem Banach-Raum X heißt **differenzierbar in** $t^* \in I$, wenn der Grenzwert

$$\dot{c}(t^*) = \lim_{t \to t^*} \frac{c(t) - c(t^*)}{t - t^*}$$

  existiert, und $\dot{c}(t^*)$ nennt man dann die **Ableitung** (oder auch **Tangentialvektor**)[2] von c in t^*. Im Fall $\dot{c}(t^*) \neq 0 \in X$ existiert der **Tangenten-Einheitsvektor** $\frac{\dot{c}(t^*)}{\|\dot{c}(t^*)\|}$. Desweiteren heißt c **differenzierbar**, falls c in jedem Punkt $a \in I$ differenzierbar ist. Ist zusätzlich die durch Differentiation induzierte Abbildung $\dot{c} : t \mapsto \dot{c}(t)$ auch noch stetig, so nennt man c **stetig differenzierbar**.

- Falls $X = \mathbb{R}^n$ ist eine Abbildung $c = (c_1, \ldots, c_n) : I \to \mathbb{R}^n$ nach Satz 6 (Forster 2, §2) genau dann stetig, wenn alle Komponenten $c_i : I \to \mathbb{R}$, $i = 1, \ldots n$, von c stetig sind. Analog ist eine parametrisierte Kurve $c : I \to \mathbb{R}^n$ genau dann differenzierbar, wenn jede Komponentenfunktion $c_i : I \to \mathbb{R}^n,\ i = 1, \ldots, n,$ differenzierbar ist (Satz 1, Forster 2, §2).

- Sei $C \subset X$ eine Kurve im Banach-Raum X und $c : I \to X$ eine Parametrisierung von C. Ist $\varphi : J \to I$ ein Homöomorphismus zwischen Intervallen I, J, dann heißt φ eine C^1**-Parametertransformation**. Durch $c \circ \varphi : J \to X$ ist eine neue Parametrisierung von C gegeben. Dabei heißt φ **orientierungstreu** genau dann, wenn φ (streng) monoton wachsend. Andernfalls heißt φ **orientierungsumkehrend**.

- Ist $c : I \to X$ eine differenzierbare Parametrisierung einer Kurve $C \subset X$ im Banach-Raum X, und $\varphi : J \to I$ eine C^1-Parametertransformation, so ist $c \circ \phi$ (als direkte Folgerung der Kettenregel) in allen $t \in J$ differenzierbar mit Ableitung

$$\frac{d}{dt}(c \circ \varphi)(t) = \dot{c}(\varphi(t)) \cdot \varphi'(t)\ .$$

  Da $\varphi'(t)$ alle möglichen Werte aus $\mathbb{R} \setminus \{0\}$ annehmen kann, ist nur die Menge T_xC aller Tangentialvektoren in einem Punkt $x \in C$ durch C eindeutig bestimmt.

- Ist C eine Kurve in einem Banach-Raum und ist $x \in C$ ein Punkt, an welchen die Ableitung jeder stetig differenzierbaren Parametrisierung von C verschwindet (also $T_xC = \{0\}$ gilt), so heißt x ein **singulärer Punkt** von C. Andernfalls bezeichnet man $x \in C$ als einen **regulären Punkt** von C.

[1] physikalische Interpretation: Eine Kurve beschreibt die zeitliche Bewegung eines Punktes im Raum $\mathbb{R}^n$

[2] physikalische Interpretation: Der Geschwindigkeitsvektor ist genau der Grenzwert des Quotienten aus Ortsdifferenz und Zeitdifferenz. Die Zahl $\|\dot{c}(t^*)\|_2$ gibt den Betrag der Geschwindigkeit an.

- Ist $c : I \to X$ eine differenzierbare Parametrisierung einer Kurve $C \subset X$ im Banach-Raum X, welche nicht injektiv ist, so bezeichnen wir einen Punkt $x \in X$ mit $x = f(t_1) = f(t_2)$ und $t_1 \neq t_2$ als **Doppelpunkt** der Kurve C.

- Existiert (mindestens) eine stetig differenzierbare Parametrisierung $c : I \to X$ von C, die injektiv ist und $\dot{c}(t) \neq 0$ für alle $t \in I$ erfüllt, so heißt C eine **injektiv immersierte Kurve**.[3] In diesem Fall ist die Menge T_xC aller möglichen Tangentialvektoren am Punkt $x \in C$ ein eindimensionaler Unterraum von X.[4]

- **Definition [rektifizierbare Kurve]**: Eine parametrisierte Kurve $c : I \to X$ in einem Banach-Raum $(X, \| \cdot \|)$ heißt **rektifizierbar**, wenn die Längen aller Sehnenpolygone von c beschränkt sind. Dabei versteht man unter einem Sehnenpolygon zu einer Zerlegung $Z = \{t_0 < t_1 < \cdots < t_k\}$ von I mit Punkten $t_i \in I$ das durch die Strecken von $c(t_{i-1})$ nach $c(t_i)$ gegebene Polygon, und unter der Länge dieses Sehnenpolygons den Wert

$$\sum_{i=1}^{k} \|c(t_i) - c(t_{i-1})\| \ .$$

- **Definition [Bogenlänge]**: Die **Bogenlänge** $L(c)$ einer rektifizierbaren Kurve $c : I \to X$ in einem Banach-Raum X ist das Supremum aller Längen der Sehnenpolygone von c, wobei das Supremum über alle Zerlegungen $Z = \{t_0 < t_t < \cdots < t_k\}$ von I mit Punkten $t_i \in I$ gebildet wird.

- **Satz 1 [Forster 2, §4]**: Jede stetig differenzierbare Parametrisierung $c : [a, b] \to X$ einer Kurve $C \subset X$ in einem Banach-Raum X ist rektifizierbar und besitzt die Bogenlänge
$$L(c) \ = \ \int_b^a \|\dot{c}(t)\| dt. \tag{4.1}$$
Ist $c : I \to X$ eine stetig differenzierbare Parametrisierung und $\varphi : J \to I$ eine C^1-Parametertransformation, so gilt $L(c) = L(c \circ \varphi)$.

- Eine Kurve C in einem Banach-Raum $(X, \|.\|)$ heißt durch eine stetig differenzierbare Parametrisierung $c : I \to X$ **nach Bogenlänge** parametrisiert, falls $\|\dot{c}(t)\| = 1$ für alle $t \in I$ gilt.

- **Definition [Schnittwinkel]**: Zwei reguläre Kurven $C_1, C_2 \subset \mathbb{R}^n$ seien durch $c_1 : I_1 \to \mathbb{R}^n$ und $c_2 : I_2 \to \mathbb{R}^n$ parametrisiert. Existieren $t_1 \in I_1, t_2 \in I_2$ mit $c_1(t_1) = c_2(t_2)$, so ist durch die Gleichung
$$\cos(\theta) \ = \ \frac{\langle \dot{c}_1(t_1), \dot{c}_2(t_2) \rangle}{\|\dot{c}_1(t_1)\|_2 \|\dot{c}_2(t_2)\|_2} \qquad (0 \leq \theta \leq \pi)$$
der **Schnittwinkel** $\theta := \theta(c_1, c_2)$ der parametrisierten Kurven c_1 und c_2 gegeben.

- Sind $\varphi_i : J_i \to I_i \ (i = 1, 2)$ zwei C^1-Parametertransformationen, so gilt
$$\theta(c_1 \circ \varphi_1, c_2 \circ \varphi_2) \ = \ \begin{cases} \theta(c_1, c_2) & \text{falls } \varphi_1, \varphi_2 \text{ beide orientierungstreu oder} \\ & \text{falls } \varphi_1, \varphi_2 \text{ beide orientierungsumkehrend} \\ \pi - \theta(c_1, c_2) & \text{sonst.} \end{cases}$$

[3] Selbst injektiv immersierte Kurven können Punkte besitzen, in deren Nähe die Kurve nicht homöomorph zu einem Intervall ist (vgl. Beispiel 2.10 im Skript aus dem Sommersemester 2010).

[4] Für eine Kurve C können wir immer eine Parametrisierung finden, die für ein Intervall konstant ist und dort Ableitung Null besitzt.

Zusatzaufgabe 4.1: Geben Sie Beispiele für eine Kurve $C \subset \mathbb{R}^2$ an,

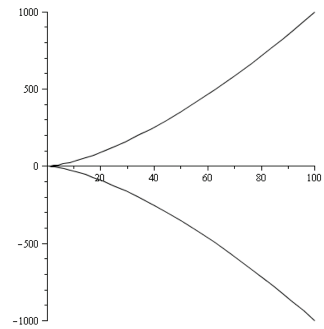
(a) welche singuläre Punkte enhält;

(b) welche keine singulären Punkte enthält (eine sogenannte immersierte Kurve), aber keine injektive Parametrisierung existiert;

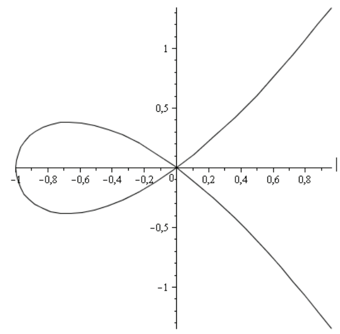
(c) welche injektiv immersiert ist, jedoch mindestens einen Punkt besitzt, in deren Nähe die Kurve nicht homöomorph zu einem Intervall ist.[5]

Lösung zu Zusatzaufgabe 4.1:

(a) (Beispiel 2.8) Die durch $t \mapsto (t^2, t^3)$ parametrisierte **Neilsche Parabel** C hat $(0,0)$ als singulären Punkt. Denn ist $c(t) = (x(t), y(t))$ eine beliebige Parametrisierung mit $c(0) = (0,0)$, dann muss $x'(0) = 0$ gelten, da $x(t)$ bei $t = 0$ ein globales Minimum hat, und wegen $y(t)^2 = x(t)^3$ (ansonsten würde c gar nicht die Kurve C parametrisieren) gilt dann auch $y'(0) = \pm\frac{3}{2}\sqrt{|x(0)|}x'(0) = 0$ (vgl. (4.6) in Forster 2).

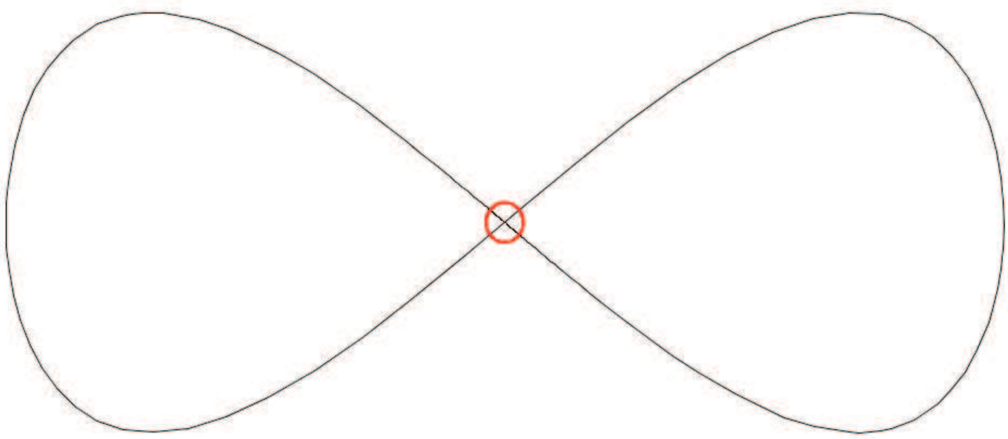


(b) (Beispiel 2.9) Der durch die beliebig oft differenzierbare Kurve $c : \mathbb{R} \to \mathbb{R}^2$ mit $c(t) := (t^2 - 1, t^3 - t)$ parametrisierte **Newtonsche Knoten** C besitzt den Doppelpunkt $c(1) = (0,0) = c(-1)$, bei dem $\dot{c}(1) = (2,2)$ kein Vielfaches von $\dot{c}(-1) = (-2,2)$ ist. (vgl. (4.5) in Forster 2).



(c) Die durch $c(t) := (2\sin(2\arctan(t)), \sin(4\arctan(t)))$ mit $c : \mathbb{R} \to \mathbb{R}^2$ parametrisierte Kurve C sieht aus wie eine liegende Acht, obwohl c injektiv ist und $\dot{c}(t) \neq 0$ für alle $t \in \mathbb{R}$ erfüllt (vgl. Beispiel 2.10 im Skript (Sommer 2010)).

Insbesondere gibt es keinen Homöomorphismus von $C \cap U$ auf ein offenes Intervall I, wenn U eine offene Umgebung von $(0,0)$ ist. Denn $(C \cap U) \setminus \{(0,0)\}$ hat vier Zusammenhangskomponenten, während $I \setminus \{r\}$ ($r \in I$ beliebig) zwei Zusammenhangskomponenten hat, was der Eigenschaft widerspricht, dass ein Homöomorphismus zusammenhängende Mengen auf zusammenhängende Mengen abbildet.



[5]Ist die Parametrisierung $c : I \to C$ nicht nur eine injektive Immersion, sondern sogar ein Homöomorphismus auf ihr mit der Relativtopologie versehenes Bild $C \subset X$, so ist C ein Beispiel für eine eindimensionale Untermannigfaltigkeit von X. Die genaue Definition einer Untermannigfaltigkeit werden wir später kennenlernen.

Zusatzaufgabe 4.2: Zeigen Sie:

(a) Das Ableiten von parametrisierten Kurven ist linear.

(b) Jede Lipschitz-stetige Parametrisierung $c : [a,b] \to X$ auf einem kompakten Intervall $[a,b]$ ist rektifizierbar.

(c) Für eine Parametrisierung $c : [a,b] \to \mathbb{R}^n$ einer gegebenen Kurve nach Bogenlänge ist die Länge der Kurve durch $L(c) = b - a$ gegeben.

Lösung zu Zusatzaufgabe 4.2:

(a) Für Konstanten $\lambda, \mu \in \mathbb{R}$ gilt nach Definition des Grenzwertes

$$\begin{aligned} \frac{d}{dt}(\lambda c(t) + \mu\tilde{c}(t)) &= \lim_{t^* \to t} \frac{(\lambda c(t^*) + \mu\tilde{c}(t^*)) - (\lambda c(t) + \mu\tilde{c}(t))}{t^* - t} \\ &= \lambda \lim_{t^* \to t} \frac{c(t^*) - c(t)}{t^* - t} + \mu \lim_{t^* \to t} \frac{\tilde{c}(t^*) - \tilde{c}(t)}{t^* - t} \;=\; \lambda\dot{c}(t) + \mu\dot{\tilde{c}}(t)\ . \end{aligned}$$

(b) Mit einer Lipschitz-Konstanten L von c gilt

$$\sum_{i=1}^{k} \|c(t_i) - c(t_{i-1})\| \le \sum_{i=1}^{k} L|t_i - t_{i-1}| = L(b-a)$$

für eine beliebige Zerlegung $Z = \{t_0 < t_1 < \cdots < t_k\}$ von I ($t_i \in I$). Damit ist die Länge jedes Sehnenpolygons beschränkt, also c rektifizierbar. Insbesondere folgt

$$L(c) \;:=\; \sup_Z \sum_{i=1}^{k} \|c(t_i) - c(t_{i-1})\| \le L \cdot (b-a)\ .$$

(c) Nach Voraussetzung gilt $\|\dot{c}(t)\| = 1$ in jedem Punkt $t \in [a,b]$ und deswegen

$$L(c) \;=\; \int_a^b \|\dot{c}(t)\|\,dt = \int_a^b 1\,dt = b - a\,.$$

Zusatzaufgabe 4.3: Bestimmen Sie die ...

(a) ... Tangentialvektoren und Kurvenlänge der Zykloide $c(t) = (t - \sin t, 1 - \cos t)$, $t \in [0, 2\pi]$.

(b) ... Bogenlänge der Neilschen Parabel $c(t) = (t^2, t^3)$ für $t \in [-1,1]$.

Lösung zu Zusatzaufgabe 4.3:

(a) Die Tangentialvektoren sind $\dot{c}(t) = (1 - \cos t, \sin t)$ und wegen $(2\sin(\frac{t}{2}))^2 = 2(1 - \cos(t))$ ergibt sich die Kurvenlänge zu

$$L = \int_0^{2\pi} \|\,\dot{c}(t)\,\|_2\,dt \;=\; \int_0^{2\pi} \sqrt{2(1-\cos t)}\,dt \;=\; \int_0^{2\pi} 2\,\sin\left(\tfrac{t}{2}\right)\,dt \;=\; -4\,\cos\left(\tfrac{t}{2}\right)\Big|_0^{2\pi} \;=\; 8.$$

(b) Mit $\dot{c}(t) = (2t, 3t^2)$ ergibt sich aufgrund der Symmetrie für die Bogenlänge

$$L(c) \;=\; \int_1^{-1} \|\dot{c}(t)\|_2 dt \;=\; \frac{2}{18}\int_0^1 18t\sqrt{4+9t^2}\,dt \;=\; \frac{2}{27}\cdot\sqrt{(4+9t^2)^3}\,\Big|_{t=0}^{t=1} \;=\; \frac{2(13\sqrt{13}-8)}{27} \;\approx\; 2,88.$$

Alle Übungsaufgaben sind verfügbar unter: `https://studip.uni-rostock.de/studip/`

Zusatzmaterial zur Analysis II – Serie 5

Niveaumengen, Höhenlinien

- **Definition [Graph]:** Sei $U \subset \mathbb{R}^n$ und $f : U \to \mathbb{R}, (x_1, \ldots, x_n) \mapsto f(x_1, \ldots, x_n)$ eine Abbildung. Der **Graph** von f ist die Menge $\Gamma_f := \{(x, y) \in U \times \mathbb{R} \ : \ y = f(x)\} \subset \mathbb{R}^{n+1}$. Für $n = 2$ ist Γ_f als Fläche über U im dreidimensionalen Raum interpretierbar.

- Sei $U \subset \mathbb{R}^n$ und $f : U \to \mathbb{R}, (x_1, \ldots, x_n) \mapsto f(x_1, \ldots, x_n)$ eine Abbildung. Dann ist f bestimmt durch die Schar $N_f(c)$, $c \in \mathbb{R}$, definiert als die Urbilder
$$N_f(c) \ := \ f^{-1}(\{c\}) \ = \ \{x \in U \ : \ f(x) = c\} \subset \mathbb{R}^n \ ,$$
welche **Niveaumengen** (bzw. im Fall $n = 2$ auch **Höhenlinien**) genannt werden.

Partielle Differenzierbarkeit, Gâteaux-Differenzierbarkeit

- **Definition:** Eine Abbildung $f : X \to Y$ zwischen Banach-Räumen X, Y heißt **im Punkt** $a \in X$ **partiell differenzierbar in Richtung** $h \in X$, falls der Grenzwert
$$\partial_h f(a) := \lim_{t \to 0} \frac{f(a + th) - f(a)}{t} \tag{5.1}$$
in Y existiert, und in diesem Fall bezeichnet man $\partial_h f(a)$ als die **partielle Ableitung von** f **im Punkt** a **in Richtung** h. Ist f im Punkt a in jede Richtung partiell differenzierbar, so nennt man die Funktion f **partiell differenzierbar in** a. Ist die Abbildung
$$G : X \to Y \ , \qquad G : h \mapsto \partial_h f(a) \tag{5.2}$$
linear und stetig (d.h. $G \in L(X, Y)$), dann heißt f **Gâteaux-differenzierbar in** a.

  Bemerkungen:

  (a) $f : X \to Y$ ist genau dann in a in Richtung h partiell differenzierbar, wenn die parametrisierte Kurve $c : \mathbb{R} \to Y$, $c(t) = f(a + th)$, in $t = 0$ differenzierbar ist.

  (b) $f : X \to R$ ist genau dann in a in Richtung h partiell differenzierbar, wenn die Funktion $g : \mathbb{R} \to \mathbb{R}$, definiert durch $t \mapsto f(a + th)$, in 0 differenzierbar ist.

  (c) Ist $f : U \to \mathbb{R}$, $U \subset \mathbb{R}^n$, im Punkt $a \in U$ partiell differenzierbar in die i-te Koordinatenrichtung $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$, so schreiben wir anstelle von $\partial_{e_i} f(a)$ üblicherweise $\frac{\partial f}{\partial x_i}(a)$. Mit der Abbildung $\varphi : \xi \mapsto f(a_1, \ldots, a_{i-1}, \xi, a_{i+1}, \ldots, a_n)$ gilt dann $\frac{\partial f}{\partial x_i}(a) = \varphi'(a_i)$.

  (d) Ist $f : U \to \mathbb{R}$, $U \subset \mathbb{R}^n$ offen, im Punkt $a \in U$ Gâteaux-differenzierbar, so gilt
$$\partial_h f(a) \ = \ \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) h_i \ . \tag{5.3}$$

  (e) **Achtung:** Weder aus der partiellen Differenzierbarkeit noch aus der Gâteaux-Differenzierbarkeit kann i.A. auf die Stetigkeit einer Funktion geschlossen werden (siehe Aufgabenblatt). Ebenso muss die Kettenregel nicht gelten.

- Ist f in jedem Punkt aus U partiell differenzierbar in Richtung h und ist auch die partielle Ableitung $\partial_h f$ als Funktion von U nach Y in jedem Punkt aus U partiell differenzierbar in Richtung k, dann heißt f zweimal partiell differenzierbar und $\partial_k \partial_h f(a)$ wird die **zweite partielle Ableitung im Punkt** $a \in U$ **in Richtung** (h, k) genannt. Rekursiv kann man analog partielle Ableitungen beliebig hoher Ordnungen definieren.[1]

[1] Vorsicht: Im Allgemeinen gilt $\partial_k \partial_h f(a) \neq \partial_h \partial_k f(a)$

- Sei $U \subset X$ eine offene Teilmenge eines Banach-Raumes X und $f : U \to \mathbb{R}$ eine Funktion. Man sagt, f hat in $x^* \in U$ ein **lokales Minimum** (bzw. **lokales Maximum**), wenn es eine Umgebung V von x^* mit $f(x^*) \leq f(x)$ (bzw. $f(x^*) \geq f(x)$) für alle $x \in V$ gibt.

Gradient, Vektorfeld, Divergenz, Rotation, Laplace-Operator: Sei $U \subset \mathbb{R}^n$ offen.

- Für $f : U \to \mathbb{R}$ partiell differenzierbar ist der **Gradient** von f in $x \in U$ der Vektor

$$\operatorname{grad} f(x) \;:=\; \left(\tfrac{\partial}{\partial x_1} f(x), \;\ldots\;, \tfrac{\partial}{\partial x_n} f(x)\right) \;=\; \nabla f(x) \qquad \text{mit} \qquad \nabla \;=\; \left(\tfrac{\partial}{\partial x_1}, \;\ldots,\; \tfrac{\partial}{\partial x_n}\right) .$$

  Dabei kann man ∇ (sprich Nabla) als vektorwertigen Differentialoperator auffassen.

- Für $f, g : U \to \mathbb{R}$ partiell differenzierbar gilt die Produktregel $\boxed{\nabla(fg) \;=\; g\nabla f \;+\; f\nabla g}$.

- Unter einem **Vektorfeld** auf U versteht man eine Abbildung $v : U \to \mathbb{R}^n$.

- Sei $v = (v_1, \ldots, v_n) : U \to \mathbb{R}^n$ ein partiell differenzierbares Vektorfeld (d.h. alle $v_i : U \to \mathbb{R}$ partiell differenzierbar). Die **Divergenz** des Vektorfeldes v heißt dann die Funktion

$$\operatorname{div} v(x) \;:=\; \langle \nabla, v \rangle \;=\; \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\, v_i(x)$$

- Für $f : U \to \mathbb{R}$ partiell differenzierbar und $v : U \to \mathbb{R}^n$ partiell differenzierbar gilt

$$\operatorname{div}(fv) \;=\; \langle \operatorname{grad} f, v \rangle \;+\; f \operatorname{div} v \qquad \text{bzw.} \qquad \langle \nabla, fv \rangle \;=\; \langle \nabla f, v \rangle \;+\; f \langle \nabla, v \rangle \;.$$

- <u>**Satz**</u>: Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ zweimal stetig partiell differenzierbar. Dann gilt

$$\forall x \in U \;\forall i, j = 1, \ldots, n \;:\; \frac{\partial}{\partial x_i}\left(\frac{\partial}{\partial x_j} f(x)\right) \;=\; \frac{\partial}{\partial x_j}\left(\frac{\partial}{\partial x_i} f(x)\right) \;.$$

- Sei $U \subset \mathbb{R}^3$ offen und $v : U \to \mathbb{R}^3$ ein partiell differenzierbares Vektorfeld. Unter der <u>**Rotation**</u> des Vektorfeldes v versteht man das Vektorfeld

$$\operatorname{rot} v \;:=\; \nabla \times v \;=\; \begin{pmatrix} \frac{\partial}{\partial x_2} v_3 - \frac{\partial}{\partial x_3} v_2 \\ \frac{\partial}{\partial x_3} v_1 - \frac{\partial}{\partial x_1} v_3 \\ \frac{\partial}{\partial x_1} v_2 - \frac{\partial}{\partial x_2} v_1 \end{pmatrix}$$

- Sei $U \subset \mathbb{R}^n$ offen. Für eine zweimal stetig partiell differenzierbare Funktion $f : U \to \mathbb{R}$ definiert man

$$\Delta f \;:=\; \operatorname{div} \operatorname{grad} f \;=\; \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2} f \qquad \text{mit dem } \underline{\textbf{Laplace-Operator}} \qquad \Delta \;=\; \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2} \;.$$

- Die Lösungen der <u>**Potentialgleichung**</u> $\Delta f = 0$ werden harmonische Funktionen genannt.

- Sei $I \subset \mathbb{R}$ ein Intervall, das als Zeitintervall interpretiert werde. Für $U \subset \mathbb{R}^n$ offen seien die Koordinaten in $U \times I \subset \mathbb{R}^{n+1}$ mit $(x, t) = (x_1, \ldots, x_n, t)$ bezeichnet. Für Funktionen

$$f : U \times I \to \mathbb{R} \;, \qquad (x, t) \mapsto f(x, t) \;,$$

  heißt dann

$$\Delta f - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} f \;=\; 0$$

  die <u>**Schwingungsgleichung**</u> (oder <u>**Wellengleichung**</u>) und

$$\Delta f - \frac{1}{k} \frac{\partial}{\partial t} f \;=\; 0$$

  die <u>**Wärmeleitungsgleichung**</u>. Dabei wirkt der Laplace-Operator auf f nur als Funktion des Ortes x. Die Konstanten $c > 0$ und $k > 0$ geben die Wellenausbreitungsgeschwindigkeit bzw. Temperaturleitfähigkeit an.

Zusatzaufgabe 5.1:

(a) Zeigen Sie: Für Abbildungen $f, g : X \to Y$, die im Punkt $a \in X$ partiell differenzierbar in Richtung h sind, ist auch die Linearkombination $(\lambda f + \mu g)$ mit $\lambda, \mu \in \mathbb{R}$ im Punkt a partiell differenzierbar in Richtung h mit $\partial_h(\lambda f + \mu g)(a) = \lambda \partial_h f(a) + \mu \partial_h g(a)$.

(b) Zeigen Sie: Für beliebiges $\lambda \in \mathbb{R}$ ist $f : X \to Y$ im Punkt $a \in X$ in Richtung λh partiell differenzierbar ist, falls $f : X \to Y$ im Punkt $a \in X$ in Richtung h partiell differenzierbar ist, und es gilt $\partial_{\lambda h} f(a) = \lambda \partial_h f(a)$.

(c) Seien X, Y, Z Banach-Räume. Zeigen Sie: Eine Abbildung $f = (f_1, f_2) : X \to Y \times Z$ ist genau dann in $a \in X$ partiell differenzierbar in Richtung $h \in X$, wenn die Abbildungen f_1 und f_2 in a in Richtung h partiell differenzierbar sind. In diesem Falle gilt für die partiellen Ableitungen $\partial_h f(a) = (\partial_h f_1(a), \partial_h f_2(a))$.

(d) Sei $U \subset X$ eine offene Teilmenge eines Banach-Raumes X und seien $x, \tilde{x} \in U$ Punkte, für die die Strecke $C := \{x + t(\tilde{x} - x) \mid t \in [0,1]\}$ in U enthalten ist. Zeigen Sie: Ist die Abbildung $f : X \to \mathbb{R}$ in jedem Punkt der Strecke C partiell differenzierbar in Richtung $\tilde{x} - x$, dann gibt es ein $\theta \in]0,1[$ mit $f(\tilde{x}) - f(x) = \partial_{(\tilde{x}-x)} f(x + \theta(\tilde{x} - x))$.

(e) Zeigen Sie: Besitzt eine Funktion $f : U \to \mathbb{R}$ auf einer offenen Teilmenge $U \subset X$ eines Banach-Raumes X in $x^* \in U$ ein lokales Extremum und ist f in x^* partiell differenzierbar in Richtung h, dann gilt $\partial_h f(x^*) = 0$.

(f) Zeigen Sie: Es bezeichne e_i den i-ten Einheitsvektor. Zeigen Sie, dass die 1-Norm $\|x\|_1$ auf dem $\mathbb{R}^n$ in den Punkten e_i bei $i \neq j$ nicht in die Richtung e_j partiell differenzierbar ist.

Lösung zu Zusatzaufgabe 5.1:

(a) Unter den gegebenen Voraussetzungen gilt für beliebige $\lambda, \mu \in \mathbb{R}$

$$\begin{aligned}\partial_h(\lambda f + \mu g)(a) &= \lim_{t\to 0} \frac{(\lambda f(a+th) + \mu g(a+th)) - (\lambda f(a) + \mu g(a))}{t} \\ &= \lambda \lim_{t\to 0} \frac{f(a+th) - f(a)}{t} + \mu \lim_{t\to 0} \frac{g(a+th)) - g(a)}{t} \quad = \lambda\, \partial_h f(a) + \mu\, \partial_h g(a)\,,\end{aligned}$$

was die Behauptung zeigt.

(b) Während für $\lambda = 0$ die Aussage klar ist, folgt diese für $\lambda \neq 0$ unmittelbar aus

$$\partial_{\lambda h} f(a) = \lim_{t\to 0} \frac{f(a + t\lambda h) - f(a)}{t} = \lambda \lim_{t\to 0} \frac{f(a + t\lambda h) - f(a)}{\lambda t} = \lambda \lim_{s\to 0} \frac{f(a + sh) - f(a)}{s} = \lambda\, \partial_h f(a)\,.$$

(c) Nach Satz 1 (Forster II, §2) existiert der Grenzwert $\lim_{t\to 0} \frac{f(a+th)-f(a)}{t}$ genau dann, wenn die Grenzwerte $\lim_{t\to 0} \frac{f_1(a+th)-f_1(a)}{t}$ und $\lim_{t\to 0} \frac{f_2(a+th)-f_2(a)}{t}$ existieren, und mit den entsprechenden Bezeichnungen für die Grenzwerte gilt $\partial_h f(a) = (\partial_h f_1(a), \partial_h f_2(a))$.

(d) Nach Voraussetzung ist die Funktion $\varphi : [0,1] \to \mathbb{R}$, $\varphi(t) := f(x + t(\tilde{x} - x))$, differenzierbar. Somit gibt es nach dem Mittelwertsatz[2] ein $\theta \in]0,1[$ mit $\varphi(1) - \varphi(0) = \frac{\varphi(1)-\varphi(0)}{1-0} = \varphi'(\theta)$, also mit $f(\tilde{x}) - f(x) = \partial_{(\tilde{x}-x)} f(x + \theta(\tilde{x} - x))$.

(e) Die Funktion $\varphi(t) := f(x^* + th)$ auf $\mathbb{R}$ besitzt in $t = 0$ ein lokales Exremum und ist dort differenzierbar, also folgt aus dem notwendigen Kriterium für lokale Extrema[3] demnach $0 = \varphi'(0) = \partial_h f(x^*)$.

(f) Wegen $\lim_{t\searrow 0} \frac{\|e_i+te_j\|_1-\|e_i\|_1}{t} = \lim_{t\searrow 0} \frac{1+t-1}{t} = 1$ und $\lim_{t\nearrow 0} \frac{\|e_i+te_j\|_1-\|e_i\|_1}{t} = \lim_{t\nearrow 0} \frac{1-t-1}{t} = -1$ kann die partielle Ableitung $\partial_{e_j} f(e_i)$ für die reellwertige Funktion $f(x) := \|x\|_1$ nicht existieren.

[2] vgl. Corollar 1 zu Satz 2, Forster I, §16

[3] vgl. Satz 1, Forster I, §16

Zusatzaufgabe 5.2:

(a) Sei $f:\mathbb{R}^2\to\mathbb{R}$ durch $f(0,0):=0$ und $f(x,y):=\frac{x^2y}{x^2+y^2}$ sonst definiert. Bestimmen Sie die partiellen Ableitungen $\partial_h f$ von f im Punkt $(0,0)$ in jede Richtung $h\in\mathbb{R}^2$.

(b) Beweisen Sie, dass f in $(0,0)$ stetig ist.

(c) Ist f Gâteaux-differenzierbar in $(0,0)$?

Lösung zu Zusatzaufgabe 5.2:

(a) Mit $f(t\,h)=t\,f(h)$ für alle $t\in\mathbb{R}$ und alle $h\in\mathbb{R}^2$ ($h=(h_1,h_2)$) folgt

$$\partial_h f(0,0) \;=\; \lim_{t\to 0}\frac{f(th)-f(0,0)}{t} \;=\; \lim_{t\to 0}\frac{f(th)}{t} \;=\; \lim_{t\to 0}\frac{\frac{(th_1)^2th_2}{(th_1)^2+(th_2)^2}}{t} \;=\; \lim_{t\to 0}\frac{h_1^2h_2}{h_1^2+h_2^2} \;=\; f(h).$$

(b) Der Ungleichung

$$|f(x,y)| \;\le\; \frac{x^2|y|}{x^2+y^2} \;\le\; |y|$$

wegen gilt $\lim\limits_{(x,y)\to 0} f(x,y)=0$, was offensichtlich die Stetigkeit von f in $(0,0)$ impliziert.

(c) Die Abbildung $h\mapsto\partial_h f(0,0)$ ist in diesem Fall nichts anderes als die Abbildung $h\mapsto f(h)$, und diese ist aber nicht linear, wie die Rechnung $f(1,1)=\frac{1}{2}\neq 0=f(1,0)+f(0,1)$ bestätigt. Also ist f nicht Gâteaux-differenzierbar.

Zusatzaufgabe 5.3:

(a) Sei $(X,\langle .,.\rangle)$ ein Hilbertraum und $f(x):=\|x\|=\sqrt{\langle x,x\rangle}$ die Abbildung, welche einem $x\in X$ seine Euklidische Norm zuordnet. Ist $f:X\to\mathbb{R}$ in Punkten $a\neq 0$ partiell differenzierbar in jede Richtung $h\in X$? Ist sie dort sogar Gâteaux-differenzierbar?

(b) Zeigen Sie: Die durch

$$f(x) \;:=\; \begin{cases}\frac{\prod_{i=1}^n x_i}{\|x\|_2^{2n}} & \text{für } x\neq 0\,,\\ 0 & \text{für } x=0\end{cases} \tag{5.4}$$

definierte Funktion $f:\mathbb{R}^n\to\mathbb{R}$ ist in Null für beliebiges $i=1,\ldots,n,$ in die Koordinatenrichtung e_i partiell differenzierbar, aber dort nicht stetig.

Lösung zu Zusatzaufgabe 5.3:

(a) Die Funktion $t\mapsto f(a+th)=\sqrt{\langle a+th,a+th\rangle}=(\|a\|^2+2t\langle a,h\rangle+t^2\|h\|^2)^{1/2}$ ist bei $a\neq 0$ nach den üblichen Differentiationsregeln aus der Analysis I in $t=0$ differenzierbar mit Ableitung

$$\partial_h f(a)=\frac{2\langle a,h\rangle}{2\,(\|a\|^2)^{1/2}}=\frac{\langle a,h\rangle}{\|a\|}\quad.$$

Offensichtlich ist $h\mapsto\partial_h f(a)=\frac{\langle a,h\rangle}{\|a\|}$ linear, und die Stetigkeit dieser linearen Abbildung folgt (vgl. Satz 10, Forster II, §2) aus der Cauchy-Schwarzschen Ungleichung $|\langle a,h\rangle|\le\|a\|\|h\|$, nach welcher die lineare Abbildung $h\mapsto\partial_h f(a)$ die Operatornorm $\frac{\|a\|}{\|a\|}=1<\infty$ besitzt. Also ist f Gâteaux-differenzierbar auf $X\setminus\{0\}$, da $\partial_h f(a)$ bei $a\neq 0$ für jedes $h\in X$ existiert sowie linear und stetig in h ist.

(b) Die durch (5.4) definierte Funktion $f:\mathbb{R}^n\to\mathbb{R}$, ist nicht nur in allen Punkten $a\neq 0$ partiell differenzierbar, sondern auch im Nullpunkt in die Koordinatenrichtungen e_i partiell differenzierbar, denn wegen $f(te_i)=0$ für alle $t\in\mathbb{R}$ gilt

$$\frac{\partial f}{\partial x_i}(0) \;=\; \lim_{\substack{t\to 0\\ t\neq 0}}\frac{f(te_i)-0}{t}=0\quad.$$

Aber f ist nicht stetig im Nullpunkt, denn z.B. existiert der Grenzwert von $f(t,t,\ldots,t)=\frac{t^n}{(nt^2)^n}$ für $t\to 0$ nicht und ist somit insbesondere nicht gleich $f(0)=0$.

Zusatzaufgabe 5.4:

Untersuchen Sie, ob die durch $$f(x,y) := \begin{cases} \frac{x^2y^2}{x^2+y^2}, & (x,y) \neq (0,0) \\ 0, & (x,y) = (0,0) \end{cases}$$

auf $\mathbb{R}^2$ definierte reellwertige Funktion (a) stetig ist, ob (b) die partiellen Ableitungen in die Koordinatenrichtungen jeweils existieren und ob (c) Gâteaux-Differenzierbarkeit vorliegt.

Lösung zu Zusatzaufgabe 5.4:

(a) Bei der Stetigkeit ist nur der Punkt $(0,0)$ zu untersuchen, da in allen anderen Punkten f als Komposition stetiger Funktionen stetig ist. Für alle $(x,y) \neq (0,0)$ gilt nun

$$|f(x,y) - f(0,0)| \;=\; \left|\tfrac{x^2y^2}{x^2+y^2}\right| \;=\; x^2 \cdot \tfrac{y^2}{x^2+y^2} \;\leq\; x^2$$

und analog

$$|f(x,y) - f(0,0)| \;=\; \left|\tfrac{x^2y^2}{x^2+y^2}\right| \;=\; y^2 \cdot \tfrac{x^2}{x^2+y^2} \;\leq\; y^2,$$

woraus für $(x,y) \to 0$ sofort $f(x,y) \to 0$, also die Stetigkeit von f in $(0,0)$ folgt.

(b) Für die partiellen Ableitungen in Punkten $(x,y) \neq (0,0)$ erhalten wir

$$\frac{\partial f(x,y)}{\partial x} \;=\; \frac{2xy^2(x^2+y^2) - x^2y^2 \cdot 2x}{(x^2+y^2)^2} \;=\; \frac{2xy^4}{(x^2+y^2)^2}$$

und

$$\frac{\partial f(x,y)}{\partial y} \;=\; \frac{2x^2y(x^2+y^2) - x^2y^2 \cdot 2y}{(x^2+y^2)^2} \;=\; \frac{2x^4y}{(x^2+y^2)^2}\,.$$

In $(x,y) = (0,0)$ folgt für alle $x,y \in \mathbb{R}$ aus $f(x,0) = 0$ und $f(0,y) = 0$ nach Definition

$$\tfrac{\partial f(0,0)}{\partial x} \;=\; \lim_{x\to 0} \tfrac{f(x,0)-0}{x} \;=\; 0 \qquad \text{sowie} \qquad \tfrac{\partial f(0,0)}{\partial y} \;=\; \lim_{y\to 0} \tfrac{f(0,y)-0}{x} \;=\; 0\,.$$

(c) Darüber hinaus existieren die partiellen Ableitungen in eine beliebige Richtung $h \in \mathbb{R}^2$ wegen

$$\partial_h f(0,0) \;=\; \lim_{t\to 0} \frac{f(th) - f(0,0)}{t} \;=\; \lim_{t\to 0} \frac{f(th)}{t} \;=\; \lim_{t\to 0} \frac{\frac{(th_1)^2(th_2)^2}{(th_1)^2+(th_2)^2}}{t} \;=\; \lim_{t\to 0} t\frac{h_1^2h_2^2}{h_1^2+h_2^2} \;=\; 0\,,$$

womit die Abbildung $h \mapsto \partial_h f(0,0)$ als konstante Nullabbildung trivialerweise linear und stetig ist, also f in $(x,y) = (0,0)$ Gâteaux-differenzierbar ist.
In allen anderen Punkten ist f ebenfalls Gâteaux-differenzierbar, denn die partiellen Ableitungen in eine beliebige Richtung $h \in \mathbb{R}^2$ sind gegeben durch

$$\begin{aligned}
\partial_h f(x,y) &= \lim_{t\to 0} \frac{f(x+th_1, y+th_2) - f(x,y)}{t} \;=\; \lim_{t\to 0} \frac{\frac{(x+th_1)^2(y+th_2)^2}{(x+th_1)^2+(y+th_2)^2} - \frac{x^2y^2}{x^2+y^2}}{t} \\
&= \lim_{t\to 0} \frac{(x+th_1)^2(y+th_2)^2[x^2+y^2] - x^2y^2[(x+th_1)^2+(y+th_2)^2]}{t[x^2+y^2]\cdot[(x+th_1)^2+(y+th_2)^2]} \\
&= \;\ldots\; = h_1 \cdot \frac{2xy^4}{(x^2+y^2)^2} \;+\; h_2 \cdot \frac{2x^4y}{(x^2+y^2)^2} \;=\; \begin{pmatrix} \frac{2xy^4}{(x^2+y^2)^2} & \frac{2x^4y}{(x^2+y^2)^2} \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}\,,
\end{aligned}$$

womit die Abbildung $h \mapsto \partial_h f(x,y)$ wie behauptet linear und stetig ist.

Zusatzaufgabe 5.5:

(a) Sei $r : \mathbb{R}^n \to \mathbb{R}$ definiert durch $r(x) := \|x\|_2 = \sqrt{\sum_{k=1}^{n} x_k^2}$ für $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Welche Niveaumengen $N_r(c) = \{x \in \mathbb{R}^n \;:\; r(x) = c\}$ sind nicht leer? Was beschreiben sie?

(b) Zeigen Sie, dass r in allen $x \in \mathbb{R}^n \setminus \{0\}$ partiell differenzierbar ist und berechnen Sie die partiellen Ableitungen. Was ist demzufolge der Gradient von $r(x)$?

(c) Fassen Sie den Gradienten von $r(x)$ als Vektorfeld $F : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}^n$ auf und berechnen Sie die Divergenz dieses Vektorfeldes. Berechnen Sie dazu vorher den Gradienten von $\frac{1}{r(x)}$.

(d) Sei $f : \mathbb{R}_+^* \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion.
Berechnen Sie Δg für $g := f \circ r : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$.

(e) Zeigen Sie, dass $F(x) := \frac{1}{(r(x))^{n-2}}$ für $n \geq 3$ die Potentialgleichung $\Delta F = 0$ löst.

Lösung zu Zusatzaufgabe 5.5:

(a) Es ist $N_r(0) = \{0\}$. Für $c < 0$ gilt $N_r(c) = \emptyset$ und für $c > 0$ beschreiben die Niveaumengen $N_r(c)$ Sphären (für $n = 2$ Kreise und für $n = 3$ Kugeloberflächen) vom Radius c.

(b) Für jedes $i = 1, \ldots, n$ ist $r(x) = \sqrt{x_i^2 + \sum_{\substack{k=1\\k\neq i}}^n x_k^2}$, so dass für $x \neq 0$ nach der Kettenregel

$$\frac{\partial}{\partial x_i} r(x) = \frac{\partial}{\partial x_i} \left(x_i^2 + \sum_{\substack{k=1\\k\neq i}}^n x_k^2 \right)^{\frac{1}{2}} = \frac{1}{2} \left(x_i^2 + \sum_{\substack{k=1\\k\neq i}}^n x_k^2 \right)^{-\frac{1}{2}} \cdot 2x_i = \frac{x_i}{r(x)}.$$

Also gilt offenbar $\operatorname{grad} r(x) = \nabla r(x) = \frac{x}{r(x)}$.

(c) Analog zu (b) ist für $x \neq 0$ nach der Kettenregel

$$\frac{\partial}{\partial x_i} \frac{1}{r(x)} = \frac{\partial}{\partial x_i} \left(x_i^2 + \sum_{\substack{k=1\\k\neq i}}^n x_k^2 \right)^{-\frac{1}{2}} = -\frac{1}{2} \left(x_i^2 + \sum_{\substack{k=1\\k\neq i}}^n x_k^2 \right)^{-\frac{3}{2}} \cdot 2x_i = -\frac{x_i}{(r(x))^3}$$

und somit $\nabla \frac{1}{r(x)} = \frac{-x}{(r(x))^3}$. Weiterhin können wir die Identität $\operatorname{Id} : \mathbb{R}^n \to \mathbb{R}^n,\ x \mapsto x$, offenbar als spezielles Vektorfeld auffassen, welches die Divergenz

$$\operatorname{div} \operatorname{Id} = \sum_{i=1}^n \frac{\partial}{\partial x_i} x_i = \sum_{i=1}^n 1 = n$$

besitzt. Mit $\langle x, x \rangle = (r(x))^2$ und den Rechenregeln für die Divergenz folgt nun

$$\operatorname{div} \operatorname{grad} r = \operatorname{div} \left(\frac{1}{r} x \right) = \left\langle \operatorname{grad} \frac{1}{r}, x \right\rangle + \frac{1}{r} \operatorname{div} x = -\frac{\langle x, x \rangle}{r^3} + \frac{n}{r} = \frac{n-1}{r}.$$

(d) Nach der Kettenregel folgt zunächst $\operatorname{grad} g = \operatorname{grad} f(r) = f'(r) \operatorname{grad} r = f'(r) \frac{x}{r}$. Dann ist mit der Definition von Δg und den Rechenregeln der Divergenz

$$\begin{aligned} \Delta g = \operatorname{div} \operatorname{grad} g = \operatorname{div} \left(f'(r) \frac{x}{r} \right) &= \left\langle \operatorname{grad} f'(r), \frac{x}{r} \right\rangle + f'(r) \operatorname{div} \frac{x}{r} \\ &= f''(r) \left\langle \frac{x}{r}, \frac{x}{r} \right\rangle + f'(r) \frac{n-1}{r} = f''(r) + f'(r) \frac{n-1}{r}. \end{aligned}$$

(e) Wegen $F = f \circ r$ betrachten wir hier $f(r) = \frac{1}{r^{n-2}}$ mit den Ableitungen $f'(r) = \frac{2-n}{r^{n-1}}$ und $f''(r) = \frac{(2-n)(1-n)}{r^n}$. Nach (d) folgt somit

$$\Delta F = f''(r) + f'(r) \frac{n-1}{r} = \frac{(2-n)(1-n)}{r^n} + \frac{2-n}{r^{n-1}} \cdot \frac{n-1}{r} = 0.$$

Zusatzmaterial zur Analysis II – Serie 6

Differenzierbarkeit, Kettenregel, Stetige Differenzierbarkeit, Schrankensatz

- **Definition [(total/Fréchet-)differenzierbar]:** Seien $(X, \|.\|_X), (Y, \|.\|_Y)$ Banach-Räume, $U \subset X$ offen. Eine Abbildung $f : U \to Y$ heißt **im Punkt** $a \in U$ **differenzierbar**, falls ein $A \in L(X, Y)$ (d.h., eine lineare und stetige Abbildung $A : X \to Y$) existiert, so dass
$$\lim_{h \to 0} \frac{f(a+h) - f(a) - Ah}{\|h\|} = 0 .$$
In diesem Fall heißt A die **Ableitung von** f **im Punkt** a und wird mit $df(a) := A$ symbolisiert.[1]

- **Satz:**[2] Seien $(X, \|.\|_X), (Y, \|.\|_Y)$ Banach-Räume, $U \subset X$ offen. Ist $f : U \to Y$ in $a \in U$ differenzierbar, dann ist f in a in jede Richtung $h \in X$ partiell differenzierbar mit $df(a)h = \partial_h f(a)$, also f in a auch Gâteaux-differenzierbar.

- **Satz 1(a) [vgl. Forster, §6]:** Seien $(X, \|.\|_X), (Y, \|.\|_Y)$ Banach-Räume, sei $f : U \to Y$ eine Abbildung auf einer offenen Teilmenge $U \subset X$. Ist f in einem Punkt $a \in U$ differenzierbar, dann ist f auch stetig[3] in a.

- **Bemerkung/Satz 1(b):** Ist $U \subset \mathbb{R}^m$ offen und $f : U \to \mathbb{R}^n$ im Punkt $a \in U$ differenzierbar, kann $df(a) \in L(\mathbb{R}^m, \mathbb{R}^n)$ bezüglich der kanonischen Basen wegen $\langle df(a)e_j, e_i\rangle = \frac{\partial f_i}{\partial x_j}(a)$ mit $j = 1, \dots, m$ und $i = 1, \dots, n$ durch die $(m \times n)$-Matrix
$$Jf(a) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_m}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(a) & \dots & \frac{\partial f_n}{\partial x_m}(a) \end{pmatrix} \tag{6.1}$$
repräsentiert werden. Diese wird dann als die **Jacobi-Matrix** bzw. die **Funktionalmatrix** von f im Punkt a bezeichnet. Speziell für eine in a differenzierbare Funktion $f : U \to \mathbb{R}$ definiert man den **Gradienten** $\operatorname{grad} f(a)$ als Spaltenvektor mittels des Euklidischen Skalarproduktes $\langle \cdot, \cdot \rangle$ durch $df(a) = \langle \operatorname{grad} f(a), \cdot \rangle$.

- **Lemma:**[4] Seien $(X, \|.\|_X), (Y, \|.\|_Y), (Z, \|.\|_Z)$ Banach-Räume, $U \subset X$ offene Teilmenge. Eine Abbildung $f = (f_1, f_2) : U \to Y \times Z$ ist genau dann im Punkt $a \in U$ differenzierbar, wenn die Abbildungen $f_1 : U \to Y$ und $f_2 : U \to Z$ im Punkt a differenzierbar sind. In diesem Fall gilt $df(a) = (df_1(a), df_2(a))$.

- **Satz 3 [Kettenregel]:** Seien $(X, \|.\|_X), (Y, \|.\|_Y), (Z, \|.\|_Z)$ Banach-Räume und $U \subset X$, $V \subset Y$ offen. Ist $f : U \to V$ differenzierbar in $a \in U$ mit Ableitung $df(a) \in L(X, Y)$ und $g : V \to Z$ differenzierbar in $b = f(a)$ mit Ableitung $dg(b) \in L(Y, Z)$, dann ist $g \circ f$ differenzierbar in a mit Ableitung[5]
$$d(g \circ f)(a) = dg(f(a)) \circ df(a) \qquad \text{(Kettenregel).} \tag{6.2}$$

---

[1] In der Literatur wird die Differenzierbarkeit im oben definierten Sinn auch oft als totale oder auch Fréchet-Differenzierbarkeit und die Ableitung A entsprechend als das totale oder Fréchet-Differential bezeichnet.

[2] zum Beweis siehe Satz 2.24 im Skript zur Analysis II vom Sommer 2009; ein Spezialfall ist Satz 4, Forster §6

[3] Achtung: Partielle Differenzierbarkeit oder Gâteaux-Differenzierbarkeit reichen im Allgemeinen nicht aus, um auf Stetigkeit schließen zu können. Vergleiche Aufgabe 5.2 (a) und (b).

[4] zum Beweis siehe Lemma 2.27 im Skript zur Analysis II vom Sommersemester 2009

[5] Insbesondere gilt bei $X = \mathbb{R}^k$, $Y = \mathbb{R}^m$, $Z = \mathbb{R}^n$ für die Jacobi-Matrizen $J(g \circ f)(a) = Jg(f(a)) \cdot Jf(a)$.

- **Definition [stetig differenzierbar]:** Eine Abbildung $f : U \to Y$ mit $U \subset X$ offen und X, Y Banach-Räumen heißt **stetig differenzierbar**, wenn f auf U differenzierbar ist und die Ableitung $df : U \to L(X, Y)$ stetig ist.

- **Satz:**[6] Seien $(X, \|.\|_X), (Y, \|.\|_Y)$ Banach-Räume, $U \subset X$ offen. Eine Abbildung $f : U \to Y$ ist genau dann stetig differenzierbar, wenn sie Gâteaux-differenzierbar mit stetiger Gâteaux-Ableitung $U \ni x \mapsto (h \mapsto \partial_h f(x)) \in L(X, Y)$ ist.

- **Korollar [vgl. Satz 2, Forster §6]:**[7] Existieren für $f : U \to \mathbb{R}^n$ mit $U \subset \mathbb{R}^m$ offen die partiellen Abbildungen $\frac{\partial f_j}{\partial x_i}$ in jedem Punkt von U und ist $x \mapsto \frac{\partial f_j}{\partial x_i}(x)$ stetig für jedes $i = 1, \ldots, m$, $j = 1, \ldots, n$, so ist f stetig differenzierbar.

- Der Mittelwertsatz gilt in der folgenden Form (vgl. Satz 5, Forster §6 oder Satz 2.20):

  Sei $U \subset X$ eine offene Teilmenge eines Banach-Raumes $(X, \|.\|_X)$ und seien $x, \tilde{x} \in U$ Punkte, für die die Strecke $C := \{x + t(\tilde{x} - x) \,|\, t \in [0, 1]\}$ in U enthalten ist. Ist die Abbildung $f : X \to \mathbb{R}$ in jedem Punkt der Strecke C partiell differenzierbar in Richtung $\tilde{x} - x$, dann gibt es ein $\theta \in (0, 1)$ mit $f(\tilde{x}) - f(x) = \partial_{(\tilde{x}-x)} f(x + \theta(\tilde{x} - x))$.

- **Schrankensatz (vgl. Corollar zu Satz 5, Forster §6 oder Satz 2.36)**: Seien $(X, \|.\|_X)$, $(Y, \|.\|_Y)$ Banach, $U \subset X$ offen und konvex. Ist $f : U \to Y$ stetig differenzierbar, so gilt

  $$\forall x, \tilde{x} \in U \quad : \quad \|f(\tilde{x}) - f(x)\|_Y \;\leq\; \left(\max_{t \in [0,1]} \|df(x + t(\tilde{x} - x))\|_{L(X,Y)} \right) \|\tilde{x} - x\|_X \;.$$

  Insbesondere ist jede stetig differenzierbare Abbildung lokal Lipschitz-stetig, d.h. es gibt zu jedem Punkt eine Umgebung, auf der f Lipschitz-stetig ist (Korollar 2.37).

- **Definition:** Seien $(X, \|.\|_X)$, $(Y, \|.\|_Y)$ Banach, $U \subset X$ offen. Ist die Abbildung $df : U \to L(X, Y)$ einer stetig differenzierbaren Abbildung $f : U \to Y$ selbst wieder stetig differenzierbar, so heißt f **zweimal stetig differenzierbar** mit $d^2 f : U \to L(X, L(X, Y))$. Induktiv definieren wir so **k-mal stetig differenzierbare Abbildungen**[8] und $C^k(U, Y)$ als den Vektorraum der k-mal stetig differenzierbaren Abbildungen $f : U \to Y$.

- **Satz von Schwarz (Satz 2.39):** Seien X, Y Banach, $U \subset X$ offen, $f \in C^2(U, Y)$. Für alle $x \in U$ gilt

  $$\forall h, \tilde{h} \in X \quad : \quad d^2 f(x)(h, \tilde{h}) = d^2 f(x)(\tilde{h}, h) \;.$$

Zusatzaufgabe 6.1:

(a) Bestimmen Sie alle Punkte $(x, y) \in \mathbb{R}^2$, in denen die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$, definiert durch $f(x) = |xy|$, (total) differenzierbar ist.

(b) Bestimmen Sie alle Punkte $(x, y) \in \mathbb{R}^2$, in denen die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$, definiert durch $f(x) = x \cdot |xy|$, (total) differenzierbar ist.

Lösung zu Zusatzaufgabe 6.1:

(a) In der Nähe jedes Punktes $(x, y) \in \mathbb{R}^2 \setminus \{(x, y) \in \mathbb{R}^2 \mid xy = 0\}$ entspricht die Funktion $f(x) = |xy|$ lokal dem Polynom xy oder $-xy$, welches differenzierbar mit $df(x, y) = \pm \begin{pmatrix} y & x \end{pmatrix}$ ist (die partiellen Ableitungen sind stetig und Satz 2, Forster §6)).

In Punkten $(x, 0)$ mit $x \neq 0$ existieren die partiellen Ableitungen $\partial_{(0,1)} f(x, 0)$ wegen

$$\lim_{t \searrow 0} \tfrac{f(x,t) - f(x,0)}{t} = \lim_{t \searrow 0} \tfrac{|xt| - 0}{t} = |x| \lim_{t \searrow 0} \tfrac{t}{t} = |x| \neq -|x| = |x| \lim_{t \nearrow 0} \tfrac{-t}{t} = \ldots = \lim_{t \nearrow 0} \tfrac{f(x,t) - f(x,0)}{t}$$

[6] zum Beweis siehe Satz 2.33 im Skript zur Analysis II vom Sommersemester 2009

[7] zum Beweis siehe Korollar 2.34 im Skript zur Analysis II vom Sommersemester 2009

[8] Da man Abbildungen $B \in L(X, L(X, Y))$ durch $(x, \tilde{x}) \mapsto B(x)(\tilde{x})$ mit bilinearen stetigen Abbildungen $X \times X \to Y$ identifizieren kann, folgt induktiv, dass die k-te Ableitung eine (nach Satz 2.39 symmetrische) k-lineare Abbildung $d^k f(x) : X \times \ldots \times X \to Y$ ist.

nicht, so dass $f(x) = |xy|$ (beispielsweise nach Satz 1, Forster §6) in diesen Punkten nicht differenzierbar sein kann. Analog existieren die partiellen Ableitungen $\partial_{(1,0)} f(0, y)$ wegen

$$\lim_{t\searrow 0} \tfrac{f(t,y)-f(0,y)}{t} = \lim_{t\searrow 0} \tfrac{|ty|-0}{t} = |y| \lim_{t\searrow 0} \tfrac{t}{t} = |y| \neq -|y| = |y| \lim_{t\nearrow 0} \tfrac{-t}{t} = \ldots = \lim_{t\nearrow 0} \tfrac{f(t,y)-f(0,y)}{t}$$

nicht in Punkten $(0, y)$ mit $y \neq 0$, womit ebenfalls folgt, dass $f(x) = |xy|$ in diesen Punkten nicht differenzierbar sein kann.

Dagegen existieren die partiellen Ableitungen im Punkt $(0, 0)$ wegen

$$f_x(0,0) = \partial_{(1,0)} f(0,0) = \lim_{t\to\infty} \tfrac{f(t,0)-f(0,0)}{t} = 0$$

und analog $f_y(0, 0) = 0$, so dass $\begin{pmatrix} 0 & 0 \end{pmatrix} : \mathbb{R}^2 \to \mathbb{R}$ ein Kandidat für das Fréchet-Differential $df(0)$ ist und aufgrund von $\sqrt{h_1^2 + h_2^2} \geq \max\{|h_1|, |h_2|\}$ und folglich

$$\lim_{\|h\|\to 0} \frac{f(h) - f(0) - df(0)h}{\|h\|} = \lim_{\|h\|\to 0} \frac{|h_1 h_2| - 0 - 0}{\sqrt{h_1^2 + h_2^2}} \leq \lim_{\|h\|\to 0} \max\{|h_1|, |h_2|\} = 0$$

tatsächlich $f(x) = |xy|$ in $(0, 0)$ total differenzierbar ist. Somit ist die Menge der Punkte, in denen $f(x) = |xy|$ total differenzierbar ist, genau $\{(0, 0)\} \cup \mathbb{R}^2 \setminus \{(x, y) \in \mathbb{R}^2 \mid xy = 0\}$.

(b) Analog zu (a) ist die Funktion in allen Punkten $(x, y) \in \mathbb{R}^2$ mit $xy \neq 0$ differenzierbar (da dort die partiellen Ableitungen existieren und stetig sind). Desweiteren existieren die partiellen Ableitungen in allen Punkten $(0, y)$ wegen

$$\lim_{t\to 0} \tfrac{f(t,y)-f(0,y)}{t} = \lim_{t\to 0} \tfrac{t\cdot|ty|-0}{t} = \lim_{t\to 0} |ty| = 0 \quad \text{und} \quad \lim_{t\to 0} \tfrac{f(0,y+t)-f(0,y)}{t} = \lim_{t\to 0} \tfrac{0-0}{t} = 0,$$

so dass wegen $\left| \frac{h_1\cdot|h_1(y+h_2)|}{\sqrt{h_1^2+h_2^2}} \right| \leq |h_1(y + h_2)|$ dort die Fréchet-Differenzierbarkeit folgt.

Dagegen existieren in Punkten $(x, 0)$ mit $x \neq 0$ die partiellen Ableitungen $\partial_{(0,1)} f(x, 0)$ wegen

$$\lim_{t\searrow 0} \tfrac{f(x,t)-f(x,0)}{t} = x \cdot |x| \lim_{t\searrow 0} \tfrac{t}{t} = x \cdot |x| \neq -x \cdot |x| = x \cdot |x| \lim_{t\nearrow 0} \tfrac{-t}{t} = \lim_{t\nearrow 0} \tfrac{f(x,t)-f(x,0)}{t}$$

nicht, so dass $f(x) = x \cdot |xy|$ (beispielsweise nach Satz 1, Forster §6) in diesen Punkten nicht differenzierbar sein kann. Somit ist die Menge der Punkte, in denen $f(x) = x \cdot |xy|$ total differenzierbar ist, genau $\mathbb{R}^2 \setminus \{(x, 0) \in \mathbb{R}^2 \mid x \neq 0\}$.

Zusatzaufgabe 6.2:

(a) Existieren die partiellen Ableitungen der Funktion $f : \mathbb{R}^2 \to \mathbb{R}$, die durch

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

definiert ist, im Punkt $(0, 0)$ in die Richtungen $(1, 0)$, $(0, 1)$ und $(1, 1)$? Ist f in $(0, 0)$ stetig? Ist f in $(0, 0)$ differenzierbar?

(b) Geben Sie ein Beispiel für eine zweimal partiell differenzierbare Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ an, für die $\partial_k \partial_h f(a) \neq \partial_h \partial_k f(a)$ an (mindestens) einem Punkt $a \in \mathbb{R}^2$ gilt.

(c) Bestimmen Sie für $g(t) = (\sin(t)+\cos(t), t-t^2)$ und $h(x, y) = \ln(1+x^2+y^2)$ die Ableitung von $(h \circ g)(t)$ im Punkt $t = 0$ mittels der Kettenregel.

Lösung zu Zusatzaufgabe 6.2:

(a) Es gilt

$$\partial_h f(0,0) := \lim_{t\to 0} \frac{f(th)}{t} = \lim_{t\to 0} \frac{2t^2 h_1 h_2}{t^3(h_1^2 + h_2^2)} = \lim_{t\to 0} \frac{2h_1 h_2}{t(h_1^2 + h_2^2)}.$$

Demnach existieren die partiellen Ableitungen im Punkt $(0,0)$ sowohl in Richtung $h_x = (1,0)$ als auch in Richtung $h_y = (0,1)$:

$$\begin{aligned}
\lim_{t\to 0} \frac{f(th_x)}{t} &= \lim_{t\to 0} \frac{2\cdot 1\cdot 0}{t(1^2+0^2)} = 0,\\
\lim_{t\to 0} \frac{f(th_y)}{t} &= \lim_{t\to 0} \frac{2\cdot 0\cdot 1}{t(0^2+1^2)} = 0.
\end{aligned}$$

Jedoch existiert die partielle Ableitung nicht in Richtung $h_{xy} = (1,1)$:

$$\lim_{t\to 0} \frac{f(th_{xy})}{t} = \lim_{t\to 0} \frac{2\cdot 1\cdot 1}{t(1^2+1^2)} = \lim_{t\to 0} \frac{1}{t} \quad \text{existiert nicht.}$$

Die Funktion ist in $(0,0)$ nicht stetig, denn beispielsweise ist

$$\lim_{t\to 0} f(th_{xy}) = \lim_{t\to 0} \frac{2t^2\cdot 1\cdot 1}{t^2(1^2+1^2)} = 1 \neq 0 = f(0,0).$$

Da die Funktion in $(0,0)$ nicht stetig ist, kann sie dort nicht differenzierbar sein (Satz 2).

(b) (Beispiel 2.21) Die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) := \frac{x^3y-xy^3}{x^2+y^2}$ für $(x,y) \neq (0,0)$ und $f(0,0) := 0$, ist Gâteaux-differenzierbar mit partiellen Ableitungen

$$\frac{\partial f}{\partial x}(x,y) = \frac{(x^2+y^2)(3x^2y-y^3)-2x(x^3y-xy^3)}{(x^2+y^2)^2} = \frac{(x^4+4x^2y^2-y^4)y}{(x^2+y^2)^2}$$

und

$$\frac{\partial f}{\partial y}(x,y) = \frac{(x^2+y^2)(x^3-3xy^2)-2y(x^3y-xy^3)}{(x^2+y^2)^2} = \frac{(x^4-4x^2y^2-y^4)x}{(x^2+y^2)^2}$$

in Koordinatenrichtungen für $(x,y) \neq (0,0)$ sowie $\frac{\partial f}{\partial x}(0,0) = 0 = \frac{\partial f}{\partial y}(0,0)$. Die gemischten zweiten partiellen Ableitung im Nullpunkt existieren, sie sind aber wegen

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) = \lim_{t\to 0} \frac{\frac{\partial f}{\partial x}(0,t)-0}{t} = \lim_{t\to 0} \frac{\frac{(0^4+4\cdot 0^2t^2-t^4)t}{(0^2+t^2)^2}-0}{t} = -1$$

und

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = \lim_{t\to 0} \frac{\frac{\partial f}{\partial y}(t,0)-0}{t} = \lim_{t\to 0} \frac{\frac{(t^4-4t^20^2-0^4)t}{(t^2+0^2)^2}-0}{t} = 1$$

voneinander verschieden.

(c) Es gilt $dg(t) = \dot{g}(t) = \begin{pmatrix} \cos(t)-\sin(t) \\ 1-2t \end{pmatrix}$, also im Ursprung somit $dg(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Weiterhin gilt

$$dh(x,y) = \left(\frac{2x}{1+x^2+y^2}, \frac{2y}{1+x^2+y^2}\right).$$

Somit ist Punkt $g(0) = (1,0)$ die Ableitung genau $dh(g(0)) = dh(1,0) = \begin{pmatrix} 1 & 0 \end{pmatrix}$. Insgesamt ergibt sich nach der Kettenregel

$$d(h\circ g)(0) = dh(g(0))\cdot dg(0) = \begin{pmatrix} 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1 .$$

Zusatzmaterial zur Analysis II – Serie 7

Taylor-Formel, Taylor-Polynom, Hessematrix

- **Bezeichnungen:** Für ein n-tupel $(\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$ und $x \in \mathbb{R}^n$ definiert man üblicherweise

$$|\alpha| \;:=\; \|\alpha\|_1 \;:=\; \sum_{k=1}^{n} |\alpha_k| \,, \qquad \alpha! \;:=\; \prod_{k=1}^{n} (\alpha_k!) \,, \qquad x^\alpha \;:=\; \prod_{k=1}^{n} x^{\alpha_k} \,.$$

  Da die Differentiationsreihenfolge bei $|\alpha|$-mal stetig differenzierbaren Funktionen nach dem Satz von Schwarz (vgl. Satz 2.39 im Skript SoSe 2009) keine Rolle spielt, betrachten wir

$$D^\alpha f \;:=\; \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \ldots\, \partial x_n^{\alpha_n}}$$

- **Satz 1 [Forster §7]**: Sei $U \subset \mathbb{R}^n$ offen, $f \in C^k(U, \mathbb{R})$ und $x \in U, \xi \in \mathbb{R}^n$ derart, dass $\{x + t\xi \mid t \in [0,1]\} \subset U$. Dann ist $g : [0,1] \to \mathbb{R}, g(t) = f(x + t\xi)$ k-mal stetig differenzierbar mit

$$\frac{d^k g}{dt^k}(t) \;=\; \sum_{k=|\alpha|} \frac{k!}{\alpha!} (D^\alpha f)(x + t\xi) \xi^\alpha$$

- **Satz 2 [Taylorformel, Forster II, §7]**: Sei $U \subset \mathbb{R}^n$ offen, $f \in C^{k+1}(U, \mathbb{R})$ und $x \in U, \xi \in \mathbb{R}^n$ derart, dass $\{x + t\xi \mid t \in [0,1]\} \subset U$. Dann existiert ein $\theta \in [0,1]$, so dass

$$f(x + \xi) \;=\; \sum_{|\alpha| \le k} \frac{D^\alpha f(x)}{\alpha!} \xi^\alpha \;+\; \sum_{|\alpha| = k+1} \frac{D^\alpha f(x + \theta\xi)}{\alpha!} \xi^\alpha \,.$$

  Nach **Corollar 1 [Forster II, §7]** konvergiert der zweite Term gegen Null für $\|\xi\| \to 0$.

- (Definition 2.40) Seien X, Y Banach, $U \subset X$ offen, $f \in C^k(U, Y)$, $a \in X$, $d^0 f(a) := f(a)$. Das **Taylorpolynom k-ten Grades von f im Entwicklungspunkt** a ist definiert durch

$$\boxed{T_k f(x; a) := \sum_{r=0}^{k} \tfrac{1}{r!} d^r f(a)(x - a)^r} \quad \text{mit} \quad \boxed{(x - a)^r \;:=\; \begin{cases} ((x - a), \cdots, (x - a)) \in X^r, & r > 0, \\ 1\,, & r = 0. \end{cases}}$$

- Im Fall $f \in C^k(U, \mathbb{R})$ mit $U \subset \mathbb{R}^n$ offen ist die k-te Ableitung in Richtung h^k durch gegeben

$$\boxed{d^k f(x) h^k \;=\; \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} \frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_k}}(x)\, h_{i_1} \ldots h_{i_k}} \tag{7.1}$$

  Für $k = 2$ gilt wegen $df(a)h = \langle \operatorname{grad} f(a), h \rangle$ und $d^2 f(a)(h, h) \;=\; \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\, h_i h_j$ mit der (nach Satz 2.39 symmetrischen) **Hessematrix** $\operatorname{Hess} f(a) := \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)_{i,j=1}^{n}$ die prägnante Formel (vgl. Corollar 2 zu Satz 2, Forster II, §7)

$$\boxed{T_2 f(x; a) \;=\; f(a) + \langle \operatorname{grad} f(a), x - a \rangle \;+\; \tfrac{1}{2} \langle \operatorname{Hess} f(a)(x - a), (x - a) \rangle}\,. \tag{7.2}$$

Lokale Extrema

- Ist $f : U \to \mathbb{R}$ mit $U \subset X$ offen und $(X, \|.\|_X)$ Banach-Raum partiell differenzierbar in $a \in U$ und gilt $\partial_h f(a) = 0$ für alle Richtungen $h \in X$, so heißt a **stationärer Punkt von** f.

- Sei $(X, \|.\|)$ Banach, $U \subset X$ offen und $f : U \to \mathbb{R}$ eine Funktion. Man sagt, f besitzt in $x^* \in U$ ein **lokales Minimum** (bzw. ein **lokales Maximum**), wenn es eine Umgebung V von x^* mit $f(x^*) \leq f(x)$ (bzw. $f(x^*) \geq f(x)$) für alle $x \in V$ gibt.[1]
- Ein stationärer Punkt $a \in U$ wird insbesondere als **Sattelpunkt** von f bezeichnet, wenn zu jedem $\delta > 0$ Punkte $x, \tilde{x} \in B_\delta(a) \cap U$ mit $f(x) > f(a)$ und $f(\tilde{x}) < f(a)$ existieren.
- **Satz [Notwendiges Kriterium für lokale Extrema]:** Besitzt $f : U \to \mathbb{R}$ auf einer offenen Teilmenge $U \subset X$ eines Banach-Raumes X in $x^* \in U$ ein lokales Extremum und ist f in x^* partiell differenzierbar, dann ist x^* ein stationärer Punkt[2] von f (vgl. Satz 2.22).
- (Definition 2.45) Eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ beziehungsweise die ihr zugeordnete quadratische Form $\mathbb{R}^n \ni h \mapsto \langle Ah, h\rangle \in \mathbb{R}$ heißt
  - **positiv (semi-)definit**, falls $\langle Ah, h\rangle > 0$ (bzw. $\langle Ah, h\rangle \geq 0$) für alle $h \in \mathbb{R}^n \setminus \{0\}$ gilt.
  - **negativ (semi-)definit**, falls $\langle Ah, h\rangle < 0$ (bzw. $\langle Ah, h\rangle \leq 0$) für alle $h \in \mathbb{R}^n \setminus \{0\}$ gilt.
  - **indefinit**, falls es Vektoren $h, \tilde{h} \in \mathbb{R}^n$ mit $\langle Ah, h\rangle > 0$ und $\langle A\tilde{h}, \tilde{h}\rangle < 0$ gibt.
- (Lemma 2.46) Für eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ gilt
  - A ist positiv (negativ) definit $\iff$ Jeder Eigenwert von A ist positiv (negativ).
  - A ist indefinit $\iff$ A besitzt sowohl positive als auch negative Eigenwerte.
  - A ist positiv (negativ) semidefinit $\iff$ Alle EW von A sind nicht-negativ (nicht-positiv).
- **Satz [Hauptminoren - Forster II, §7]**: Eine reelle symmetrische $n \times n$-Matrix A ist genau dann positiv definit, wenn alle Hauptminoren positiv sind, d.h., wenn gilt
$$\forall s = 1, \dots, n \;:\; \det\left((a_{j,k})_{j,k=1}^s\right) \;>\; 0\,.$$
- **Satz 4 [Forster II, §7]**: (vgl. auch Satz 2.47 im Skript SoSe 2009)
Sei $f \in C^2(U, \mathbb{R})$, $U \subset \mathbb{R}^n$ offen, $a \in U$ ein stationärer Punkt (d.h. hier $\operatorname{grad} f(a) = 0$).
  (a) Ist $\operatorname{Hess} f(a)$ positiv definit, so besitzt f in a ein lokales Minimum.
  (b) Ist $\operatorname{Hess} f(a)$ negativ definit, so besitzt f in a ein lokales Maximum.
  (c) Ist $\operatorname{Hess} f(a)$ indefitiv, hat f in a kein lokales Extremum (sondern einen Sattelpunkt).

Banachscher Fixpunktsatz (vgl. Satz 1, Forster II, §8)

Vor.: (i) (M, d) ist ein vollständiger metrischer Raum.
(ii) $f : M \to M$ ist eine (Selbst)abbildung.
(iii) f ist eine Kontraktion, d.h., $\exists L < 1\ \forall x, y \in M \;:\; d(f(x), f(y)) \leq L \cdot d(x, y)$.

Beh.: (i) Es gibt einen eindeutigen Punkt $x^* \in M$ mit $f(x^*) = x^*$ (dieser heißt **Fixpunkt**).
(ii) Zu jedem beliebigen Startwert $x_0 \in M$ konvergiert die durch $x_{k+1} := f(x_k)$, $k \geq 0$, rekursiv definierte Folge gegen x^* für $k \to \infty$.

Zusatzaufgabe 7.1:

(a) Zeigen Sie, dass für eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ jeder Eigenwert reell ist.

(b) Ist die Hesse-Matrix für ein $f \in C^2(\mathbb{R}^n, \mathbb{R})$ an jedem Punkt symmetrisch?

(c) Bestimmen Sie die Nullstellen von $\operatorname{charpol}(A) := \det(A - \lambda I)$ für $A := \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 1 & 0 \end{pmatrix}$.

[1] In beiden Fällen wird auch von einem lokalen Extremum von f in $x^* \in U$ gesprochen.

[2] Im Fall $X = \mathbb{R}^n$ gilt insbesondere $\operatorname{grad} f(x^*) = 0 \iff \forall h \in \mathbb{R}^n : 0 = \partial_h f(x^*) = \langle \operatorname{grad} f(x^*), h\rangle$. Vorsicht: Ist $K \subset X$ beliebig und $x^* \in \partial K \cap K$ lokales Extremum von $f : K \to \mathbb{R}$, gilt nicht notwendigerweise $\partial_h f(x^*) = 0$.

(d) Besitzt $f(x,y,z) = \frac{x^2}{2} + (2x+y+9)z + y^2 + 3$ lokale Extrema?

(e) Bestimmen Sie die lokalen und globalen Extrema von $g : \mathbb{R}^2 \to \mathbb{R},\ g(x,y) := x^4 - 8x^2 + y^2 + 16$.

(f) Bestimmen Sie $T_2 f((x,y,z);(0,0,0))$ und $T_2 f((x,y,z);(-4,-1,2))$ mit f aus (d).

Lösung zu Zusatzaufgabe 7.1:

(a) Ähnlich dem Skalarprodukt auf $\mathbb{R}^n$ können wir auf $\mathbb{C}^n$ die Sesquilinearform (Hermite'sches Skalarprodukt) $\mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C},\ (x,y) \mapsto \langle x,y\rangle_H := \bar{y}^T x = \sum_{k=1}^{n} x_k \bar{y}_k$ betrachten, welche die leicht überprüfbaren Eigenschaften

(i) $\forall \alpha, \beta \in \mathbb{C} \forall x,y,z \in \mathbb{C}^n \ : \ \langle \alpha x + \beta y, z\rangle_H \ = \ \alpha\langle x,z\rangle_H + \beta\langle y,z\rangle_H$ (komplex-linear),

(ii) $\forall x,y \in \mathbb{C}^n \ : \ \langle x,y\rangle_H \ = \ \overline{\langle y,x\rangle_H}$ (hermite'sch),

(iii) $\forall x \in \mathbb{C}^n \ : \ \langle x,x\rangle_H \geq 0$ und $\langle x,x\rangle_H = 0 \iff x = 0$ (positiv definit)

besitzt. Insbesondere liefert $\|x\|_2 := \sqrt{\langle x,x\rangle_H}$ eine Norm auf $\mathbb{C}^n$. Ist $\lambda \in \mathbb{C}$ nun ein Eigenwert der reellen symmetrischen $n \times n$-Matrix A und $x \in \mathbb{C}^n$ ein zugehöriger Eigenvektor mit $\|x\|_2^2 = \langle x,x\rangle_H = 1$, dann gilt wie behauptet $\lambda \in \mathbb{R}$ aufgrund von $A = A^T$ (Symmetrie), $\bar{A} = A$ (Reellwertigkeit) und daher

$$\begin{aligned} \lambda &= \lambda\langle x,x\rangle_H \overset{\text{(i)}}{=} \langle \lambda x, x\rangle_H = \langle Ax, x\rangle_H = \bar{x}^T A x = \bar{x}^T \bar{A}^T x \\ &= \overline{Ax}^T x = \langle x, Ax\rangle_H = \langle x, \lambda x\rangle_H \overset{\text{(i),(ii)}}{=} \overline{\lambda\langle x,x\rangle_H} = \bar{\lambda} \,. \end{aligned}$$

(b) Für ein $f \in C^2(\mathbb{R}^n, \mathbb{R})$ gilt nach dem Satz von Schwarz (Satz 2.39) an jedem Punkt $a \in \mathbb{R}^n$ die Gleichung $\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a)$ und somit für die Hesse-Matrix

$$(\operatorname{Hess} f(a))^T \ = \ \left(\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right)_{i,j=1}^{n}\right)^T \ = \ \left(\frac{\partial^2 f}{\partial x_j \partial x_i}(a)\right)_{i,j=1}^{n} \ = \ \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right)_{i,j=1}^{n} \ = \ \operatorname{Hess} f(a) \,.$$

(c) Die Nullstellen von $\operatorname{charpol}(A) : \lambda \mapsto \det(A - \lambda I_3)$ erhalten wir aus $\det(A - \lambda I_3) = 0$. Entsprechend

$$\begin{aligned} \det(A - \lambda I_3) &= \begin{vmatrix} 1-\lambda & 0 & 2 \\ 0 & 2-\lambda & 1 \\ 2 & 1 & -\lambda \end{vmatrix} \ = \ (1-\lambda)\begin{vmatrix} 2-\lambda & 1 \\ 1 & -\lambda \end{vmatrix} + 2\begin{vmatrix} 0 & 2 \\ 2-\lambda & 1 \end{vmatrix} \\ &= -\lambda(1-\lambda)(2-\lambda) - (1-\lambda) - 4(2-\lambda) \ = \ (\lambda^2 - \lambda)(2-\lambda) + 5\lambda - 9 \\ &= 3\lambda^2 - \lambda^3 + 3\lambda - 9 \ = \ (\lambda - 3)(-\lambda^2 + 3) \end{aligned}$$

ergeben sich nun für die Eigenwerte von A unmittelbar $\lambda_1 = 3$, $\lambda_2 = \sqrt{3}$ und $\lambda_3 = -\sqrt{3}$.

(d) Wegen $df(x,y,z) = (x + 2z, z + 2y, 2x + y + 9)$ ist $P = (-4,-1,2)$ einziger stationärer Punkt von f. Da $\operatorname{Hess} f(x,y,z) = A$ in diesem Fall konstant und – nach (c) – indefinit ist, besitzt f in P nur einen Sattelpunkt. Nach Satz 2.22 kann f somit keine lokalen Extrema besitzen, da es keine weiteren stationären Punkte gibt.

(e) Die Ableitung von $g(x,y) := x^4 - 8x^2 + y^2 + 16$ ist $dg(x,y) = (4x^3 - 16x, 2y)$ und somit sind $(0,0)$, $(2,0)$ und $(-2,0)$ die stationären Punkte. Wegen $\operatorname{Hess} g(x,y) = \begin{pmatrix} 12x^2 - 16 & 0 \\ 0 & 2 \end{pmatrix}$ besitzt $\operatorname{Hess} g(0,0)$ die Eigenwerte -16 und 2, so dass bei $(0,0)$ ein Sattelpunkt von g vorliegt, während für $(\pm 2, 0)$ die Hesse-Matrix die Eigenwerte 32 und 2 besitzt und f somit in $(\pm 2, 0)$ nach Satz 2.47 lokale Minima besitzt. Die lokalen sind auch die globalen Minima, da der Wert im Sattelpunkt $g(0,0) = 16$ größer als der Wert $g(\pm 2, 0) = 0$ ist und $g(x,y) \to +\infty$ für $\|(x,y)\| \to +\infty$ wegen $g(x,y) = (x^2-4)^2 + y^2$ gilt.

(f) Nach (7.2) ergeben sich (aufgrund der Eindeutigkeit nicht überraschend)

$$\begin{aligned} T_2 f((x,y,z);(0,0,0)) &= f(0,0,0) + \langle \operatorname{grad} f(0,0,0), (x,y,z)^T\rangle \ + \ \tfrac{1}{2}\left\langle A(x,y,z)^T, (x,y,z)^T\right\rangle \\ &= 3 \ + \ 9z \ + \ \tfrac{1}{2}\left(x^2 + 2y^2 + 4xz + 2yz\right) \ = \ f(x,y,z) \end{aligned}$$

und analog $T_2 f((x,y,z);(-4,-1,2)) = f(x,y,z)$.

Zusatzaufgabe 7.2: Betrachten Sie die Iteration $x_{n+1} := f(x_n)$ mit $f : \mathbb{R} \to \mathbb{R}$, $f(x) := ax + b$.

(a) Für welche $a, b \in \mathbb{R}$ ist f eine Selbstabbildung des Intervalls $[0,1]$?

(b) Für welche $a, b \in \mathbb{R}$ ist f eine Kontraktion ?

(c) Was besagt der Banachsche Fixpunktsatz für f auf $[0,1]$?

(d) Geben Sie ein Gegenbeispiel dafür an, dass $|f(x) - f(y)| < |x - y|$ nicht ausreicht, um die Existenz eines Fixpunktes zu garantieren.

Lösung zu Zusatzaufgabe 7.2:

(a) Da $f(x) := ax + b$ immer zwischen $f(0) = b$ und $f(1) = a + b$ liegt, ist f genau dann eine Selbstabbildung von $[0,1]$, wenn b und $a + b$ zwischen 0 und 1 liegen.

(b) f ist wegen $|f(x) - f(y)| = |a|\,|x - y|$ genau dann eine Kontraktion, wenn $|a| < 1$ gilt.

(c) Der Banachsche Fixpunktsatz besagt, dass für $0 \le b, a + b \le 1$ und $|a| < 1$ die Abbildung f auf $[0,1]$ genau einen Fixpunkt hat, und dass die Iteration $x_{n+1} := f(x_n)$ für jeden Startwert $x_0 \in [0,1]$ gegen diesen Fixpunkt konvergiert. In dem hier behandelten speziellen Fall lässt sich der Fixpunkt $\frac{b}{1-a}$ sogar per Hand ausrechnen.

(d) Beispielsweise erfüllt die Abbildung $f(x) = x + \frac{\pi}{2} - \arctan(x)$ wegen $f'(x) = 1 - \frac{1}{x^2+1} < 1$ und des Mittelwertsatzes $|f(x) - f(y)| = |f'(\xi)|\,|x - y| < |x - y|$ die geforderte Ungleichung, besitzt jedoch wegen $f(x) = x \iff \arctan(x) = \frac{\pi}{2}$ und $\tan\left(\frac{\pi}{2}\right) = \pm\infty$ keinen Fixpunkt in $\mathbb{R}$.

Zusatzaufgabe 7.3:

(a) Zeigen Sie, dass mit der Matrix $A := \begin{pmatrix} 1 & -\sqrt{2} \\ -\sqrt{2} & 0 \end{pmatrix}$ die Ungleichung $\|Ax\|_2 \le 2\|x\|_2$ für alle $x \in \mathbb{R}^2$ erfüllt ist.

(b) Beweisen Sie, dass die durch $f(x) := \frac{1}{4}(Ax + b)$ mittels der oben angegebenen Matrix A und dem Vektor $b := (1, \sqrt{3})^T$ definierte Abbildung auf der Teilmenge $D := \{(x,y)^T \mid x^2 + y^2 \le 1\}$ der Euklidischen Ebene $(\mathbb{R}^2, \|\cdot\|_2)$ einen eindeutigen Fixpunkt besitzt.

Lösung zu Zusatzaufgabe 7.3:

(a) Wegen $0 \le (x + \sqrt{2}y)^2 = x^2 + 2\sqrt{2}xy + 2y^2 \iff -2\sqrt{2}xy \le x^2 + 2y^2$ gilt

$$\|A(x,y)\|_2^2 = (x - \sqrt{2}y)^2 + (-\sqrt{2}x)^2 = 3x^2 - 2\sqrt{2}xy + 2y^2 \le 4(x^2 + y^2) = 4\|(x,y)\|_2^2$$

für alle $(x,y) \in \mathbb{R}^2$. Mittels Wurzelziehen folgt nun die Behauptung.

(b) Da die Kreisscheibe $D \subset \mathbb{R}^2$ eine abgeschlossene Teilmenge des vollständigen metrischen Raumes $(\mathbb{R}^2, d_{\|\cdot\|_2})$ ist, ist $(D, d_{\|\cdot\|_2|_D})$ selbst ein vollständiger metrischer Raum. Also hat man nur noch zu prüfen, ob f eine kontrahierende Selbstabbildung auf D ist. Die Abbildung f bildet die Kreisscheibe tatsächlich in sich ab, denn für jedes $x \in D$ gilt

$$\|f(x)\|_2 \le \frac{1}{4}(\|Ax\|_2 + \|b\|_2) \le \frac{1}{4}(2\|x\|_2 + \|b\|_2) \le \frac{1}{4}(2+2) = 1\,,$$

also $f(x) \in D$. Darüberhinaus bildet f gemäß

$$\|f(x) - f(y)\|_2 = \frac{1}{4}\|Ax - Ay\| = \frac{1}{4}\|A(x-y)\| \le \frac{1}{2}\|x - y\|$$

für alle $x, y \in D$ eine Kontraktion. Also ist der Banachsche Fixpunktsatz anwendbar und liefert die Existenz eines eindeutigen Fixpunktes in D.

Zusatzmaterial zur Analysis II – Serie 8

Lineare Ausgleichsrechnung

- Ist $m > n$, so besitzt das Gleichungssystem $Ax = b$ mit $A \in \mathbb{R}^{m\times n}$ und einem Vektor $b \in \mathbb{R}^m$ i.A. keine Lösung, da wir „mehr Gleichungen als Unbekannte“ haben. In diesem Fall kann man i.A. nur noch ein $x^* \in \mathbb{R}^n$ finden, welches das **Minimierungsproblem** $\|Ax - b\|_2 \stackrel{!}{=} \min$ in $\mathbb{R}^n$ löst, oder anders gesagt, für welches die Summe der Quadrate der Komponenten des Vektors $Ax - b$ mit $x \in \mathbb{R}^n$ kleinstmöglich ist. Und $x^* \in \mathbb{R}^n$ löst das oben angegebene lineare Ausgleichsproblems genau dann, wenn es die $\boxed{\text{Normalengleichungen } A^T Ax = A^T b}$ löst.
- Ein Anwendungsbeispiel für ein lineares Ausgleichsproblems ist z.B. die lineare diskrete Approximation. Es seien die (etwa physikalischen) Messwerte (t_i, y_i) für $i = 1, \ldots, m$ gegeben. Bei linearer Approximation werden für gegebene Funktionen $f_1, \ldots, f_n : \mathbb{R} \to \mathbb{R}$ die Koeffizienten $x_j \in \mathbb{R}$, $j = 1, \ldots, n$, einer Ausgleichskurve
$$g(t) \;=\; \sum_{j=1}^{n} x_j f_j(t)$$
gesucht, für welche die quadratische Abweichung $\sum\limits_{i=1}^{m} (g(t_i) - y_i)^2$ minimal wird. Dabei sind also die Koeffizienten $x_j \in \mathbb{R}$ so zu bestimmen, dass mit der Matrix
$$A := \begin{pmatrix} f_1(t_1) & f_2(t_1) & \cdots & f_n(t_1) \\ f_1(t_2) & f_2(t_2) & \cdots & f_n(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(t_m) & f_2(t_m) & \cdots & f_n(t_m) \end{pmatrix}$$
und dem Vektor $b := (y_1, \ldots, y_m)^T$ die Funktion $F(x) := \|Ax - b\|_2^2$ minimal wird.

Diffeomorphismen, Satz über lokale Umkehrbarkeit, Satz über implizite Funktionen

- (Definition 2.53) Seien X, Y Banach, $U \subset X$, $V \subset Y$ offen. Eine Abbildung $f : U \to V$ heißt **C^k-invertierbar**/ ein **C^k-Diffeomorphismus** von U auf V, wenn $f \in C^k(U, V)$ bijektiv ist und zusätzlich $f^{-1} \in C^k(V, U)$ gilt.
- (Definition 2.54) Seien X, Y Banach, $U \subset X$ offen. Eine Abbildung $f : U \to Y$ heißt **lokal C^k-invertierbar bei** $x \in U$, wenn es offene Umgebungen $U' \subset U$ von x und $V' \subset Y$ von $f(x)$ gibt, so dass $f : U' \to V'$ ein C^k-Diffeomorphismus ist.
- (Lemma 2.55) Seien X, Y Banach, $U \subset X$ offen. Ist $f : U \to Y$ lokal C^k-invertierbar bei $x \in U$, so ist die stetige lineare Abbildung $df(x) \in L(X, Y)$ invertierbar[1] mit Inversem $df^{-1}(f(x)) \in L(Y, X)$.
- (Satz 2.57/Satz 3 (Forster II, §8) – **Satz über lokale Umkehrbarkeit/Invertierbarkeit**) Seien $(X, \|.\|_X), (Y, \|.\|_Y)$ Banach, $U \subset X$ offen. Genau dann ist eine Abbildung $f \in C^k(U, Y)$ lokal C^k-invertierbar bei $x \in U$, wenn $df(x) \in L(X, Y)$ invertierbar ist.
- (Korollar 2.58) Seien X, Y Banach, $U \subset X$ offen. Weiter besitze ein $f \in C^1(U, Y)$ die Eigenschaft, dass die Ableitung $df(x)$ in jedem Punkt $x \in U$ invertierbar ist. Dann ist $f(U)$ offen. Falls f zusätzlich injektiv ist, dann ist f sogar ein Diffeomorphismus von U auf $f(U)$.

[1] Insbesondere gilt bei $X = \mathbb{R}^m$ und $Y = \mathbb{R}^n$ auch $m = n$.

- (Satz 2.61/Satz 2 (Forster II, §8) – **Satz über implizite Funktionen/lokale Auflösbarkeit**) Seien $(X,\|.\|_X),(Y,\|.\|_Y),(Z,\|.\|_Z)$ Banach, $W\subset X\times Y$ offen und $f\in C^k(W,Z)$. Löst zu vorgegebenem $z^*\in Z$ der Punkt $(x^*,y^*)\in W$ die Gleichung $f(x^*,y^*)=z^*$ und ist die Einschränkung $df(x^*,y^*)|_Y =: d_Y f(x^*,y^*)\in L(Y,Z)$ invertierbar, dann gibt es offene Umgebungen U^* von x^* und V^* von y^* sowie ein $g\in C^k(U^*,V^*)$ mit $f(x,y)=z^* \iff y=g(x)$ für alle $(x,y)\in U^*\times V^*$.

Zusatzaufgabe 8.1:

(a) Sei $m>n$, $A\in\mathbb{R}^{m\times n}$ eine Matrix vollen Ranges und $b\in\mathbb{R}^m$. Zeigen Sie: Dann löst der Vektor $x\in\mathbb{R}^n$ das lineare Ausgleichsproblem $\|Ax-b\|_2 \stackrel{!}{=} \min$ genau dann, wenn er die Normalengleichungen $A^TAx=A^Tb$ erfüllt.

(b) Zeigen Sie, dass die Abbildung $\Phi:\mathbb{R}^2\to\mathbb{R}^2,\ \Phi(x,y):=((x-y)^2,2xy)$ in der Nähe jedes Punktes $(x,y)\in\mathbb{R}^2$ mit $|x|\neq|y|$ eine differenzierbare Umkehrabbildung besitzt. Beweisen Sie, dass Φ ein Diffeomorphismus von $A:=\{(x,y)\mid 0<y<x\}\subset\mathbb{R}^2$ auf $B:=(0,\infty)^2$ ist.

Lösung zu Zusatzaufgabe 8.1:

(a) Die Funktion $f(x):=\|Ax-b\|_2^2$ besitzt wegen $f(x)=\langle Ax-b,Ax-b\rangle$ die Ableitung $df(x)h= 2\langle Ax-b,Ah\rangle$, denn

(i) die affin-lineare Abbildung $u:\mathbb{R}^n\to\mathbb{R}^m,\ x\mapsto Ax-b$ besitzt wegen

$$\lim_{h\to 0}\tfrac{u(x+h)-u(x)-Ah}{\|h\|} \;=\; \lim_{h\to 0}\tfrac{A(x+h)-b-(Ax-b)-Ah}{\|h\|} \;=\; \lim_{h\to 0}\tfrac{0}{\|h\|} \;=\; 0$$

in jedem beliebigen Punkt die Ableitung $du(x)=A\in L(\mathbb{R}^n,\mathbb{R}^m)$.

(ii) die in jedem Punkt $\neq 0$ differenzierbare Funktion $r:\mathbb{R}^m\to\mathbb{R},\ x\mapsto r(x):=\|x\|_2$ besitzt nach ZA 5.5 (b) die Ableitung $dr(x)h = \langle\tfrac{x}{r(x)},h\rangle$.

(iii) es gilt $f(x)=(s\circ r\circ u)(x)$ mit $s:\mathbb{R}\to\mathbb{R},\ r\mapsto r^2$; also folgt nach Kettenregel

$$\begin{aligned} df(x)h &= ds((r\circ u)(x))\cdot dr(u(x))\cdot du(x)h \;=\; ds((r\circ u)(x))\cdot dr(u(x))Ah\\ &= ds((r\circ u)(x))\cdot\left\langle\frac{u(x)}{r(u(x))},Ah\right\rangle \;=\; 2r(u(x))\cdot\left\langle\frac{u(x)}{r(u(x))},Ah\right\rangle \;=\; 2\,\langle u(x),Ah\rangle\end{aligned}$$

Die notwendige Bedingung $df(x)h=0$ für alle $h\in\mathbb{R}^n$ ist dann äquivalent zu $A^T(Ax-b)=0$. Darüber hinaus ist die zweite Ableitung $\operatorname{Hess} f(x) = A^TA$ positiv definit, da A vollen Rang hat, also ist die Lösung von $A^TAx=A^Tb$ auch wirklich ein Minimum, welche eindeutig durch $x=(A^TA)^{-1}A^Tb$ bestimmt ist.

(b) Nach Satz 2.57 ist $\Phi:\mathbb{R}^2\to\mathbb{R}^2$ in einem Punkt $\xi\in\mathbb{R}^2$ lokal (C^1-)invertierbar, wenn Φ stetig differenzierbar ist und $d\Phi(\xi)$ als lineare Funktion von $\mathbb{R}^2$ nach $\mathbb{R}^2$ invertierbar ist. Die Ableitung $d\Phi(x,y)=\begin{pmatrix}2(x-y) & -2(x-y)\\ 2y & 2x\end{pmatrix}$ ist wegen

$$\det(d\Phi(x,y)) \;=\; 4x(x-y)+4y(x-y) \;=\; 4(x+y)(x-y) \;=\; 4(x^2-y^2) \;=\; 0 \quad\iff\quad |x|=|y|$$

jedoch nur in den Punkten $(x,y)\in\mathbb{R}^2$ mit $|x|=|y|$ singulär. Tatsächlich ist Φ sogar ein Diffeomorphismus von $\{(x,y)\,|\,0<y<x\}$ auf $(0,\infty)\times(0,\infty)$ mit Umkehrabbildung

$$\Psi: B\to A\,,\qquad \Psi(u,v) \;=\; \left(\tfrac12\left(\sqrt{u+2v}+\sqrt{u}\right)\quad \tfrac12\left(\sqrt{u+2v}-\sqrt{u}\right)\right)\,,$$

denn $(x-y)^2=u>0, 2xy=v>0$ ist für $(x,y)\in\mathbb{R}^2$ mit $x>y>0$ wegen $u+2v=(x+y)^2$ und $x=\frac12((x+y)+(x-y))$, $y=\frac12((x+y)-(x-y))$, äquivalent zu

$$x \;=\; \tfrac12\left(\sqrt{u+2v}+\sqrt{u}\right) \qquad\text{und}\qquad y \;=\; \tfrac12\left(\sqrt{u+2v}-\sqrt{u}\right)\,.$$

Zusatzaufgabe 8.2:

(a) Zeigen Sie, dass man für genügend nahe bei 1 liegende x, y, z das Gleichungssystem

$$\begin{cases} -2x^2 + y^2 + z^2 = 0 \\ x^2 + \mathrm{e}^{y-1} - 2y = 0 \end{cases}$$

durch C^∞-Funktionen $y = \varphi(x)$, $z = \psi(x)$ nach y und z auflösen kann.

(b) Zeigen Sie, dass in einer Umgebung des Punktes $(2, 5)$ das Gleichungssystem

$$\begin{cases} x^2 + uy + \mathrm{e}^v = 0 \\ 2x + u^2 - uv = 5 \end{cases}$$

durch eine C^1-Abbildung $(x, y) \mapsto \big(u(x, y), v(x, y)\big)$ mit $u(2, 5) = -1$ und $v(2, 5) = 0$ aufgelöst werden kann, und berechne deren Ableitung in diesem Punkt.

Lösung zu Zusatzaufgabe 8.2:

(a) Die beliebig oft stetig differenzierbare Funktion

$$F : \mathbb{R}^3 \to \mathbb{R}^2\ ,\quad F(x, y, z) := (-2x^2 + y^2 + z^2, x^2 + \exp(y - 1) - 2y)$$

erfüllt $F(x, y, z) = 0 \in \mathbb{R}^2$ für $(x, y, z) = (1, 1, 1)$ und besitzt die Ableitung

$$dF(x,y,z) \;=\; \begin{pmatrix} -4x & 2y & 2z \\ 2x & \exp(y-1)-2 & 0 \end{pmatrix} \quad\Longrightarrow\quad dF(1,1,1) \;=\; \begin{pmatrix} -4 & 2 & 2 \\ 2 & -1 & 0 \end{pmatrix} .$$

Die zu den Variablen (y, z) gehörige quadratische Untermatrix $\begin{pmatrix} 2 & 2 \\ -1 & 0 \end{pmatrix}$ ist invertierbar. Also existieren nach dem Satz über implizite Funktionen (Satz 2.61/Satz 2 - Forster II, §8) eine Umgebung U von $x = 1$, eine Umgebung V von $(y, z) = (1, 1)$ und eine vektorwertige C^∞-Funktion $g = (\varphi, \psi) : U \to V$ mit $\varphi(1) = 1$, $\psi(1) = 1$ und $F(x, \varphi(x), \psi(x)) = 0$ für alle $x \in U$.

(b) Die (sogar beliebig oft) stetig differenzierbare Funktion

$$F : \mathbb{R}^4 \to \mathbb{R}^2\ ,\quad F(x, y, u, v) := (x^2 + uy + \exp(v), 2x + u^2 - uv)$$

erfüllt $F(x, y, u, v) = (0, 5) \in \mathbb{R}^2$ für $(x, y, u, v) = (2, 5, -1, 0)$ und besitzt die Ableitung

$$dF(x,y,u,v) \;=\; \begin{pmatrix} 2x & u & y & \mathrm{e}^v \\ 2 & 0 & 2u-v & -u \end{pmatrix} \quad\Longrightarrow\quad dF(2,5,-1,0) \;=\; \begin{pmatrix} 4 & -1 & 5 & 1 \\ 2 & 0 & -2 & 1 \end{pmatrix} .$$

Die zu den Variablen (u, v) gehörige quadratische Untermatrix $\begin{pmatrix} 5 & 1 \\ -2 & 1 \end{pmatrix}$ ist invertierbar. Also existieren nach dem Satz über implizite Funktionen (Satz 2.61/Satz 2 - Forster II, §8) eine Umgebung U von $(x, y) = (2, 5)$, eine Umgebung V von $(u, v) = (-1, 0)$ und eine vektorwertige (mindestens) C^1-Abbildung $g : (x, y) \mapsto \big(u(x, y), v(x, y)\big)$ mit $u(2, 5) = -1$, $v(2, 5) = 0$ und $F(x, y, u(x, y), v(x, y)) = (0, 5)$ für alle $(x, y) \in U$.
Aus der Gleichung $F \circ H = 0$ folgt mit $F \circ H : \mathbb{R}^2 \to \mathbb{R}^2$ für $H : \mathbb{R}^2 \to \mathbb{R}^4$, $(x, y) \mapsto (x, y, g(x, y))$ nach der Kettenregel offenbar

$$dF(x, y, u(x, y), v(x, y)) \cdot \begin{pmatrix} I_2 \\ dg(x, y) \end{pmatrix} \;=\; 0_2 \in \mathbb{R}^{2\times 2}\ ,$$

also am Punkt $(x, y) = (2, 5)$ demnach

$$\begin{pmatrix} 4 & -1 & 5 & 1 \\ 2 & 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} I_2 \\ dg(2,5) \end{pmatrix} = 0_2 \Longrightarrow dg(2,5) \;=\; -\begin{pmatrix} 5 & 1 \\ -2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 4 & -1 \\ 2 & 0 \end{pmatrix} \;=\; -\frac{1}{7}\begin{pmatrix} 2 & -1 \\ 18 & -2 \end{pmatrix}$$

Zusatzaufgabe 8.3:

(a) Sei y eine Funktion, welche die Gleichung $F(x,y) := x^2 + y^2 - 1 = 0$ löst, d.h. welche $F(x,y(x)) = 0$ erfüllt. Zeigen Sie mittels Kettenregel, dass dann $y' = -\frac{x}{y}$ bei $y \neq 0$ gilt.

(b) Seien $f : \mathbb{R}^3 \to \mathbb{R}$ und $g : \mathbb{R}^2 \to \mathbb{R}$ differenzierbar und $F : \mathbb{R}^2 \to \mathbb{R}$ definiert durch $F(x,y) := f(x,y,g(x,y))$. Drücken Sie die Ableitung von F durch die partiellen Ableitungen von f und g aus. Drücken Sie die partiellen Ableitungen von g durch die von f aus, für den Fall, dass $F = 0$ gelte.

(c) Die Gleichung $z^3 + z + xy = 1$ besitzt für jedes $(x,y) \in \mathbb{R}^2$ genau eine reelle Lösung $g(x,y)$. Zeigen Sie, dass $g : \mathbb{R}^2 \to \mathbb{R}$ differenzierbar ist. Berechnen Sie anschließend $dg(1,1)$ und untersuchen Sie g auf stationäre Punkte. Besitzt g lokale Extrema ?

Lösung zu Zusatzaufgabe 8.3:

(a) Die Gleichung $F(x,y) = 0$ ist äquivalent zu $x^2 + y^2 = 1$. Für genügend kleine $x\,(|x| < 1)$ können wir dies nach $y = y(x) = \pm\sqrt{1-x^2}$ umstellen. Aus dieser expliziten Darstellung von y ergibt sich $y'(x) = \pm\frac{-x}{\sqrt{1-x^2}} = -\frac{x}{y}$. Alternativ können wir nach der Kettenregel die Gleichung $F(x,y(x)) = 0$ differenzieren und erhalten die Gleichung $dF(x,y(x)) = 2x + 2y(x)y'(x) = 0$. Unter der Annahme $y(x) \neq 0$ folgt nun ebenso $y'(x) = -\frac{x}{y(x)}$.

(b) Zunächst gilt $F = f \circ h : \mathbb{R}^2 \to \mathbb{R}$ mit einer Hilfsfunktion $h : \mathbb{R}^2 \to \mathbb{R}^3,\ h(x,y) = (x,y,g(x,y))$. Somit folgt nun nach der Kettenregel

$$\begin{aligned} dF(x,y) &= df(h(x,y)) \cdot dh(x,y) \\ &= \begin{pmatrix} \frac{\partial}{\partial x} f(h(x,y)) & \frac{\partial}{\partial y} f(h(x,y)) & \frac{\partial}{\partial z} f(h(x,y)) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\partial}{\partial x} g(x,y) & \frac{\partial}{\partial y} g(x,y) \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial}{\partial x} f(h(x,y)) + \frac{\partial}{\partial z} f(h(x,y)) \frac{\partial}{\partial x} g(x,y) & \frac{\partial}{\partial y} f(h(x,y)) + \frac{\partial}{\partial z} f(h(x,y)) \frac{\partial}{\partial y} g(x,y) \end{pmatrix}. \end{aligned}$$

Für den Fall $F = 0$ ist somit auch die Ableitung $dF(x,y) \in L(\mathbb{R}^2, \mathbb{R})$ die Nullabbildung, so dass nach Umstellen für die partiellen Ableitungen von g folgt, dass

$$\frac{\partial}{\partial x} g(x,y) = -\frac{\frac{\partial}{\partial x} f(x,y,g(x,y))}{\frac{\partial}{\partial z} f(x,y,g(x,y))}, \qquad \frac{\partial}{\partial y} g(x,y) = -\frac{\frac{\partial}{\partial y} f(x,y,g(x,y))}{\frac{\partial}{\partial z} f(x,y,g(x,y))}.$$

(c) Die (sogar beliebig oft) stetig differenzierbare Funktion $F : \mathbb{R}^3 \to \mathbb{R}$, $F(x,y,z) := z^3 + z + xy$ besitzt die Ableitung $dF(x,y,z) = \begin{pmatrix} y & x & 3z^2+1 \end{pmatrix} \implies dF(1,1,0) = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$. Die zur Variable z gehörige quadratische Untermatrix $\begin{pmatrix} 3z^2+1 \end{pmatrix}$ ist wegen $\forall z \in \mathbb{R} : 3z^2 + 1 \geq 1 > 0$ invertierbar. Also existieren zu jedem $(x,y) \in \mathbb{R}^2$ und dem nach Voraussetzung eindeutig bestimmten z nach dem Satz über implizite Funktionen (Satz 2.61/Satz 2 - Forster II, §8) eine Umgebung U von (x,y), eine Umgebung V von z und eine vektorwertige C^1-Funktion $g : U \to V$ mit $g(x,y) = z$ und $F(x,x,g(x,y)) = 0$ für alle $(x,y) \in U$.
Aus der Gleichung $F \circ H = 0$ folgt mit $F \circ H : \mathbb{R}^2 \to \mathbb{R}$ für $H : \mathbb{R}^2 \to \mathbb{R}^3,\ (x,y) \mapsto (x,y,g(x,y))$ nach der Kettenregel offenbar

$$dF(x,y,g(x,y)) \cdot \begin{pmatrix} I_2 \\ dg(x,y) \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix} \implies dg(x,y) = -\frac{1}{3(g(x,y))^2+1} \begin{pmatrix} y & x \end{pmatrix}$$

also wegen $z(z^2+1) + 1 \cdot 1 = 1 \iff z = 0$ demnach $dg(1,1) = \begin{pmatrix} -1 & -1 \end{pmatrix}$. Der einzige stationäre Punkt von g ist somit $(0,0)$, an welchem g jedoch einen Sattelpunkt besitzt (die entsprechende Hesse-Matrix $\operatorname{Hess} g(0,0)$ besitzt die Gestalt $\begin{pmatrix} 0 & -a \\ -a & 0 \end{pmatrix}$, ist also indefinit).

Alle Übungsaufgaben sind verfügbar unter: `https://studip.uni-rostock.de/studip/`

Zusatzmaterial zur Analysis II – Serie 9

Untermannigfaltigkeiten, Extrema unter NB (Lagrange-Multiplikator-Methode)

- (Definition 2.65) Sei $f : \mathbb{R}^n \to \mathbb{R}^m$ eine differenzierbare Abbildung. Ein Punkt $x \in \mathbb{R}^n$ heißt **regulärer Punkt von** f, wenn $df(x)$ surjektiv ist (insbesondere muss dann $n \geq m$ sein). Ein Wert $z^* \in \mathbb{R}^m$ heißt **regulärer Wert von** f, wenn jedes x mit $f(x) = z^*$ ein regulärer Punkt ist, d.h., wenn das Urbild $f^{-1}(\{z^*\})$ nur aus regulären Punkten besteht.

- **Satz 2.66/Satz 2 (Forster II, §9)** – (Äquivalente Charakterisierungen von Untermannigfaltigkeiten): Für eine nichtleere Teilmenge $M \subset \mathbb{R}^n$ sind äquivalent:

  (a) **(Beschreibung durch Gleichungen)** M ist lokal Urbild eines regulären Wertes einer C^k-Funktion vom $\mathbb{R}^n$ in den $\mathbb{R}^{n-d}$, d.h. für jedes $x \in M$ existiert eine Umgebung $U \subset \mathbb{R}^n$, eine C^k-Funktion $f : U \to \mathbb{R}^{n-d}$ und ein regulärer Wert $z^* \in \mathbb{R}^{n-d}$ von f mit $M \cap U = f^{-1}(\{z^*\})$.

  (b) **(Darstellung als Graph einer C^k-Funktion)** M ist lokal der Graph einer C^k-Funktion auf dem $\mathbb{R}^d$, d.h. für jeden Punkt $a \in M$ gibt es eine Zerlegung $\mathbb{R}^n \cong X \times Y$ in einen d-dimensionalen linearen Unterraum X und einen $(n-d)$-dimensionalen linearen Unterraum Y von $\mathbb{R}^n$ sowie bei $a = (x^*, y^*)$ offene Umgebungen $U^* \subset X$ von x^*, $V^* \subset Y$ von y^* und eine C^k-Abbildung $g : U^* \to V^*$ mit $M \cap (U^* \times V^*) = \text{Graph}(g)$, wobei $\text{Graph}(g) := \{(x, g(x)) | x \in U^*\}$ den Graphen von g bezeichnet.

  (c) **(Transformation in eine Ebene)** M ist lokal C^k-diffeomorph zum $\mathbb{R}^d$, d.h. für jedes $x \in M$ gibt es eine offene Umgebung $U \subset \mathbb{R}^n$ und einen C^k-Diffeomorphismus $\Phi : U \to V$ auf eine offene Teilmenge $V \subset \mathbb{R}^n$ mit $\phi(M \cap U) = \mathbb{R}^d \times \{0\}^{n-d} \cap V$ („Φ biegt die d-dimensionale Fläche M gerade“).

- **Definition 2.67**: Eine Teilmenge $M \subset \mathbb{R}^n$, die eine (und somit jede) der Eigenschaften aus Satz 2.66 besitzt, heißt **d-dimensionale C^k-Untermannigfaltigkeit des $\mathbb{R}^n$**. Die Zahl $n - d$ wird gelegentlich als **Codimension** der Untermannigfaltigkeit bezeichnet.

- **Definition 2.72/Satz 2.73**: Sei M eine d-dimensionale C^k-Untermannigfaltigkeit des $\mathbb{R}^n$. Die Gesamtheit aller Tangentialvektoren T_aM, welche aus allen möglichen Ableitungen $\dot{c}(0)$ von parametrisierten stetig differenzierbaren Kurven $c :]-\varepsilon, \varepsilon[\to M$ besteht, die innerhalb von M durch den Punkt $c(0) = a$ verlaufen, ist ein d-dimensionaler Unterraum des $\mathbb{R}^n$ und heißt **Tangentialraum von M im Punkt a**. Ist M lokal bei $a \in M$ das Urbild eines regulären Wertes $z^* \in \mathbb{R}^{n-d}$ unter der C^k-Abbildung $f : \mathbb{R}^n \to \mathbb{R}^{n-d}$, dann gilt $T_aM = \text{Kern}(df(a))$. (Die Zeilen von $df(a)$ bilden dann eine Basis des $n - d$ dimensionalen Vektorraumes N_aM, welcher aus der Gesamtheit aller Normalenvektoren an M im Punkt a besteht[1]).

- (Satz 2.76 – **Lagrange'sche Multiplikator-Regel**) Sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ differenzierbar und $g = (g_1, \ldots, g_k) \in C^1(U, \mathbb{R}^k)$. Liegt in x^* ein lokales Extremum von f unter den Nebenbedingungen $g = 0$ vor und ist $dg(x^*) \in L(\mathbb{R}^n, \mathbb{R}^k)$ surjektiv, d.h. sind die Gradienten $\operatorname{grad} g_1(x^*), \ldots, \operatorname{grad} g_k(x^*)$ linear unabhängig, dann gibt es einen Vektor $\lambda = (\lambda_1, \ldots, \lambda_k) \in \mathbb{R}^k$ mit

$$\operatorname{grad} f(x^*) = \sum_{i=1}^{k} \lambda_i \operatorname{grad} g_i(x^*) \ . \tag{9.1}$$

  Dies ist äquivalent dazu, dass $\operatorname{grad} f(x^*)$ ein Normalenvektor ist, also in N_aM liegt.[2]

[1] vergleiche auch Satz 3, Forster II, §9
[2] vergleiche auch Satz 4, Forster II, §9

Zusatzaufgabe 9.1:

(a) In der Nähe welcher Punkte ist die C^1-Abbildung $\Phi(x,y) := (x^2+y^2, x+y^2)$ ein lokaler Diffeomorphismus?

(b) Zeigen Sie, dass das Urbild $M := \{(x,y,z) \,|\, g(x,y,z) = 1\}$ des Wertes 1 unter der Funktion $g(x,y,z) := x^2+y^2+yz+z^2$ eine zweidimensionale C^∞-Untermannigfaltigkeit des $\mathbb{R}^3$ ist. Ist M kompakt?

(c) Bestimmen Sie die globalen Extrema von $f(x,y,z) := x^2+y^2+z^2$ unter der Nebenbedingung $g(x,y,z) = 1$ mit g aus (b).

Lösung zu Zusatzaufgabe 9.1:

(a) Es gilt $d\Phi(x,y) = \begin{pmatrix} 2x & 2y \\ 1 & 2y \end{pmatrix}$ und $d\Phi(x,y)$ ist invertierbar in den Punkten (x,y) mit $\det(d\Phi(x,y)) = 4xy - 2y \neq 0$. Also ist Φ in der nähe der Punkte (x,y) mit $x \neq \frac{1}{2}$ und $y \neq 0$ jeweils ein lokaler Diffeomorphismus.

(b) Der Wert 1 ist ein regulärer Wert von g, denn $dg(x,y,z) = (2x, 2y+z, y+2z)$ verschwindet wegen

$$\det \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} = 8-2 = 6 \neq 0$$

(also aufgrund der Regularität dieser Matrix) nur im Punkt $(0,0,0)$, ist also nur dort nicht surjektiv. Da aber $g(0,0,0) = 0 \neq 1$ gilt, folgt somit $(0,0,0) \notin M$. Also ist das Urbild M der Eins unter $g : \mathbb{R}^3 \to \mathbb{R}^{3-2}$ nach Satz 2.66 eine zweidimensionale C^∞-Untermannigfaltigkeit.

Die Teilmenge $M \subset \mathbb{R}^3$ ist sogar kompakt, denn sie ist abgeschlossen als Urbild der einelementigen (und daher trivialerweise abgeschlossenen) Menge $\{1\}$ unter der stetigen Funktion g, und beschränkt, da $2g(x,y,z) = 2x^2+y^2+(y+z)^2+z^2 > 4$ für $\max\{|x|,|y|,|z|\} > 2$ ist und demnach $M \subset [-2,2]^3$ gilt.

(c) Zunächst ist zu erwähnen, dass die als Polynom trivialerweise stetige Funktion f auf dem – wie in (b) gezeigt – Kompaktum $M := g^{-1}(\{1\})$ tatsächlich sowohl ihr Maximum als auch ihr Minimum annehmen muss. Desweiteren können wir aufgrund der Surjektivität von $dg(x,y,z)$ auf M den Satz 2.76 bzw. Satz 4, Forster II, §9 anwenden. Es sind somit Lösungen von (9.1) unter der Nebenbedingung $g = 1$ gesucht, was in diesem Fall auf das Gleichungssystem

$$\begin{aligned} 2x &= \lambda 2x \\ 2y &= \lambda(2y+z) \\ 2z &= \lambda(y+2z) \\ x^2+y^2+yz+z^2 &= 1 \end{aligned}$$

führt. Aus der ersten Zeile erhält man somit die Fälle $\lambda = 1$ oder $x = 0$.

- Im Fall $\lambda = 1$ erhält man aus der dritten Zeile $y = 0$ und somit aus der zweiten Zeile $z = 0$, also mit der vierten Zeile $(\pm 1, 0, 0)$ als Kandidaten für ein lokales Extremum.
- Im Fall $x = 0$ ergibt sich aus der zweiten und dritten Zeile $2(\lambda-1)y + \lambda z = 0$ und $\lambda y + 2(\lambda-1)z = 0$, also das lineare Gleichungssystem

  $$\begin{pmatrix} 2(\lambda-1) & \lambda \\ \lambda & 2(\lambda-1) \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

  Eine von $(y,z) = (0,0)$ verschiedene Lösung dieser beiden Gleichungen existiert nur dann, wenn die zugehörige Matrix singulär ist, d.h.

  $$4(\lambda-1)^2 - \lambda^2 = 3\lambda^2 - 8\lambda + 4 = 0$$

oder äquivalenterweise $\lambda = \frac{4}{3} \pm \frac{2}{3}$ gilt. Dann ist aber $z = -y$ oder $z = y$ und somit folgt aus der vierten Zeile $y^2 = 1$ oder $3y^2 = 1$. Kandidaten für Extrema sind also $(0, 1, -1)$, $(0, -1, 1)$, $\left(0, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$ und $\left(0, -\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}\right)$.

Da $f(\pm 1, 0, 0) = 1$, $f(0, 1, -1) = 2 = f(0, -1, 1)$ und $f\left(0, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) = \frac{2}{3} = f\left(0, -\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}\right)$ gilt, liegen in $\left(0, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$ und $\left(0, -\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}\right)$ globale Minima während in $(0, 1, -1)$ und $(0, -1, 1)$ globale Maxima liegen.

Zusatzaufgabe 9.2:

(a) Bestimmen Sie die Punkte auf der Ellipse $x^2 + xy + y^2 = 3$, die den größten bzw. kleinsten Abstand vom Ursprung $(0, 0)$ haben.

(b) Auf der Sphäre $S^{n-1} = \{x \in \mathbb{R}^n : \|x\|_2 = 1\} \subset \mathbb{R}^n$ sei $f : S^{n-1} \to \mathbb{R}$ durch $f(x) = \prod_{j=1}^{n} x_j^2$ gegeben. Zeigen Sie, dass $\max_{x \in S^{n-1}} f(x) = n^{-n}$ gilt.

Lösung zu Zusatzaufgabe 9.2:

(a) Der Abstand eines Punktes $(x, y) \in \mathbb{R}^2$ vom Nullpunkt ist gegeben durch

$$d_2((x, y), (0, 0)) = \|(x, y) - (0, 0)\|_2 = \|(x, y)\|_2 = \sqrt{x^2 + y^2} \ .$$

Es genügt jedoch, die Funktion $f(x, y) = x^2 + y^2$ unter der Nebenbedingung

$$g(x) = x^2 + xy + y^2 - 3 = 0$$

zu maximieren/minimieren. Zunächst halten wir fest, dass es sich bei $M = g^{-1}(\{0\})$ als Urbild einer abgeschlossenen (da einelementigen) Menge unter einer stetigen (als Polynom sogar beliebig oft stetig differenzierbaren) Funktion wieder um eine abgeschlossene Menge handelt. Da außerdem

$$x^2 + xy + y^2 = \frac{1}{2}(x + y)^2 + \frac{1}{2}(x^2 + y^2) \geq \frac{1}{2}(x^2 + y^2) \geq \frac{9}{2} > 3$$

für alle Punkte $(x, y) \notin B_3((0, 0)) := \{(x, y) \in \mathbb{R}^2 \ : \ d_2((x, y), (0, 0)) \ < \ 3\}$ erfüllt ist, folgt $M \subset B_3((0, 0))$, also auch die Beschränktheit von M. Nach dem Satz von Heine-Borel ist M demnach kompakt.

Da $dg(x, y) = \begin{pmatrix} 2x + y & x + 2y \end{pmatrix}$ wegen $\det \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} = 3 \neq 0$ surjektiv für alle $(x, y) \neq (0, 0)$ ist und $(0, 0) \notin M$ gilt, handelt es sich bei M um eine sogar kompakte Untermannigfaltigkeit des $\mathbb{R}^2$ (der Klasse C^∞). Als stetige (da als Polynom sogar beliebig oft stetig differenzierbare) Funktion muss f auf dem Kompaktum M sowohl ihr Minimum als auch ihr Maximum annehmen. Für jeden Punkt $(x^*, y^*) \in M$, an dem f ein Extremum besitzt, existiert dann (da alle Voraussetzungen von Satz 2.76 erfüllt sind) in diesem Fall ein sogenannter Lagrange-Multiplikator $\lambda \in \mathbb{R}^1$, welcher $df(x^*, y^*) = \lambda dg(x^*, y^*)$ löst. Daher ist nun das (nichtlineare) Gleichungssystem

$$2x = \lambda(2x + y) \ , \qquad 2y = \lambda(x + 2y) \ , \qquad x^2 + xy + y^2 = 3$$

zu lösen. Nun lässt sich das Teilsystem der ersten beiden Gleichungen als (in x und y) lineares Gleichungssystem mit Parameter λ auffassen. Da $(0, 0) \notin M$, interessieren nur Parameter λ, für welche die Matrix $\begin{pmatrix} 2\lambda - 2 & \lambda \\ \lambda & 2\lambda - 2 \end{pmatrix}$ singulär ist. Die Determinante $4(\lambda - 1)^2 - \lambda^2$ verschwindet nun genau bei $\lambda = \frac{4}{3} \pm \frac{2}{3}$, d.h., es muss $\lambda = 2$ oder $\lambda = \frac{2}{3}$ gelten.

Im ersten Fall ist $y = -x$, so dass die Nebenbedingung $(\sqrt{3}, -\sqrt{3})$ und $(-\sqrt{3}, \sqrt{3})$ als Kandidaten für Extrema liefert. Der Abstand vom Ursprung ist hier jeweils $\sqrt{6}$.

Im zweiten Fall ist $x = y$, so dass die Nebenbedingung $(1, 1)$ und $(-1, -1)$ als weitere Kandidaten für Extrema liefert, welche den Abstand $\sqrt{2}$ vom Ursprung haben.

(b) **Variante A:** Zunächst halten wir fest, dass $f(x) \geq 0$ ist, also ohne Einschränkung nur Punkte betrachtet werden können, für die keine Komponente verschwindet (da sonst nur das Minimum $f(x) = 0$ erwischt würde). Wir maximieren also die nichtnegative Funktion $f(x)$ unter der Nebenbedingung $g(x) = \|x\|_2^2 - 1 = 0$ mittels Lagrange-Multiplikator-Methode, denn $g : \mathbb{R}^n \to \mathbb{R}$ ist stetig differenzierbar und $dg(x) = 2x^T$ ist für alle $x \in S^{n-1}$ surjektiv (was in diesem Fall ungleich der Nullabbildung heißt). Wir suchen also die stationären Punkte der Funktion $F : \mathbb{R}^{n+1} \to \mathbb{R}$, definiert durch $F(x, \lambda) = f(x) - \lambda g(x)$. Inbesondere folgt an diesen Punkten

$$\frac{\partial F}{\partial x_i}(x, \lambda) = 0 \qquad \iff \qquad 2x_i \left(\prod_{\substack{j=1 \\ j \neq i}}^{n} x_j^2 - \lambda \right) = 0 \qquad (i = 1, \dots, n) . \tag{9.2}$$

Wie erwähnt, interessiert der Fall nicht, dass $x_i = 0$ für mindestens ein i ist. Also muss

$$\lambda = \prod_{\substack{j=1 \\ j \neq i}}^{n} x_j^2 \neq 0 \qquad (i = 1, \dots, n)$$

gelten. Dies ist jedoch nur möglich, wenn alle Komponenten (betragsmäßig) gleich sind. Aus der Nebenbedingung ergibt sich dann

$$1 = \sum_{k=1}^{n} x_k^2 = n x_i^2 \qquad \Longrightarrow \qquad x_i = \pm \frac{1}{\sqrt{n}} \qquad (i = 1, \dots, n)$$

und somit wie behauptet $f(x) = n^{-n}$. Aufgrund der Kompaktheit von S^{n-1} muss eine Maximalstelle existieren, an welcher nach Satz 2.76 Bedingung (9.2) erfüllt sein muss. Mittels Ausschlusskriterium können wir nun $\max_{x \in S^{n-1}} f(x) = n^{-n}$ schlussfolgern.

Variante B: Für beliebige Zahlen $a_k \geq 0$ ($k = 1, \dots, n$) lässt sich per Induktion die Ungleichung vom arithmetischen-geometrischen Mittel

$$\sqrt[n]{\prod_{k=1}^{n} a_k} \leq \frac{1}{n} \sum_{k=1}^{n} a_k . \tag{9.3}$$

beweisen:

- Für $n = 2$ ergibt sich $0 \leq (a-b)^2 \iff 2ab \leq a^2 + b^2 \iff ab \leq \left(\frac{a+b}{2}\right)^2$, womit der Induktionsanfang gesetzt ist. Gelte nun die Ungleichung für n und beliebige $a_k > 0, k = 1, \dots, n$ (im Fall $a_k = 0$ für ein k ist nichts zu zeigen). O.B.d.A. sei nun $a_{n+1} \geq \max_{k=1,\dots,n} a_k$ (andernfalls ordne die Elemente um). Dann gelten

  $$a_{n+1} \geq \alpha := \frac{1}{n} \sum_{k=1}^{n} a_k \qquad \text{und} \qquad x := \frac{a_{n+1} - \alpha}{(n+1)\alpha} \geq 0 ,$$

  so dass nun über die Anwendung der Bernoulli-Ungleichung nacheinander

  $$\begin{aligned} \left(\tfrac{1}{n+1} \sum_{k=1}^{n+1} a_k \right)^{n+1} &= \alpha^{n+1} \left(\tfrac{1}{(n+1)\alpha} (n\alpha + a_{n+1}) \right)^{n+1} = \alpha^{n+1} (1+x)^{n+1} \\ &\overset{\text{Bernoulli}}{\geq} \alpha^{n+1} (1 + (n+1)x) = \alpha^{n+1} \left(1 + \tfrac{a_{n+1} - \alpha}{\alpha} \right) = \alpha^n a_{n+1} \overset{\text{IV}}{\geq} \prod_{k=1}^{n+1} a_k \end{aligned}$$

  folgt, was äquivalent zur Behauptung im Fall $n+1$ ist.

Für $x \in S^{n-1}$ gilt wegen $\|x\|_2^2 = 1$ somit $f(x) = \prod_{j=1}^{n} x_j^2 \overset{(9.3)}{\leq} \left(\frac{1}{n} \sum_{k=1}^{n} x_k^2 \right)^n = \frac{1}{n^n}$, wobei die Gleichheit beispielsweise für $x = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 1 & \dots & 1 \end{pmatrix}^T$ angenommen wird.

Zusatzmaterial zur Analysis II – Serie 10

Hinreichendes Kriterium für lokale Extrema unter Nebenbedingungen

- **Satz 2.77:** Sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ und $g = (g_1, \dots, g_k) : U \to \mathbb{R}^k$ $(k < n)$ jeweils zweimal stetig differenzierbar. Desweiteren sei $M = g^{-1}(\{0\})$. Besitzt $dg(x^*)$ für einen Punkt x^* maximalen Rang, gilt mit einem Vektor λ^* die Gleichung $dF(x^*, \lambda^*) = 0$ mit $F : U \times \mathbb{R}^k \to \mathbb{R}$, durch $F(x, \lambda) := f(x) - \sum_{i=1}^{k} \lambda_i g_i(x)$ definiert,[1] und ist die Einschränkung der Hesse-Matrix $\operatorname{Hess} F_{\lambda^*}(x^*)$ mit $F_{\lambda^*} : x \mapsto F(x, \lambda^*)$ (d.h., bei festem λ^* wird nur nach x abgeleitet) auf den Tangentialraum $T_{x^*}M = \operatorname{Ker}(dg(x^*))$ positiv bzw. negativ definit, so liegt in x^* ein lokales Minimum bzw. Maximum von f unter den Nebenbedingungen $g = 0$.

Parameterabhängige Integrale, Euler-Lagrange-Gleichung

- **Satz 1.66**: Sei X ein metrischer Raum und $f : X \times [a, b] \to \mathbb{R}$ eine stetige Funktion. Dann ist die Funktion $F : X \to \mathbb{R}$, $F(x) := \int_a^b f(x, t)\, dt$, wohldefiniert und stetig.[2]

- **Satz 2.50**:[3] Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge und $f : U \times [a, b] \to \mathbb{R}$ eine stetige Funktion $f(x, t)$, deren partielle Ableitungen $\frac{\partial f}{\partial x_i}(x, t)$ nach den Variablen x_i, $i = 1, \dots, n$, als Funktionen auf $U \times [a, b]$ stetig sind. Dann ist die Funktion $F : U \to \mathbb{R}$, $F(x) := \int_a^b f(x, t)\, dt$ stetig differenzierbar mit partiellen Ableitungen $\frac{\partial F}{\partial x_i}(x) := \int_a^b \frac{\partial f}{\partial x_i}(x, t)\, dt$.

- **Satz 2.51**: Ist $L : \mathbb{R}^2 \to \mathbb{R}$, $(q, v) \mapsto L(q, v)$ stetig differenzierbar und $u \in C^1([a, b], \mathbb{R})$ so, dass $t \mapsto \frac{\partial L}{\partial v}(u(t), u'(t))$ stetig differenzierbar ist, dann gilt die **Euler-Lagrange-Gleichung**

$$\frac{d}{dt}\frac{\partial L}{\partial v}(u(t), u'(t)) = \frac{\partial L}{\partial q}(u(t), u'(t)) \tag{10.1}$$

  auf (a, b) genau dann, wenn u ein stationärer Punkt von $I(u) := \int_a^b L(u(t), u'(t))\, dt$ ist, d.h., wenn $\partial_h I(u) = 0$ für alle $h \in C^1([a, b], \mathbb{R})$ mit $h(a) = 0 = h(b)$ gilt.[4]

- **Bemerkung:** Der vorangegangene Satz lässt sich wie folgt verallgemeinern:
  Sei $L : [a, b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $(t, q_1, \dots, q_n, v_1, \dots, v_n)$ eine zweimal stetig differenzierbare Funktion. Desweiteren sei das lineare Funktional $I : C^2([a, b], \mathbb{R}^n) \to \mathbb{R}$ definiert durch

$$I(\varphi) \ := \ \int_a^b L(t, \varphi_1(t), \dots, \varphi_n(t), \dot\varphi_1(t), \dots, \dot\varphi_n(t))dt \ .$$

  Für ein $u \in C^2([a, b], \mathbb{R}^n)$ mit $I(u) = \inf_{\varphi \in C^2([a,b],\mathbb{R}^n)} I(\varphi)$ gelten dann die **Euler-Lagrange-Gleichungen**

$$\frac{d}{dt}\frac{\partial L}{\partial v_k}(t, u(t), \dot u(t)) \ = \ \frac{\partial L}{\partial q_k}(t, u(t), \dot u(t)) \qquad (k = 1, \dots, n) \ . \tag{10.2}$$

[1] Aufgrund der Ableitung nach λ gilt dann insbesondere auch $g(x^*) = 0$, also $x^* \in M$.
[2] vgl. Satz 1, Forster II, §10
[3] vgl. auch Satz 2, Forster II, §10
[4] Dies ist äquivalent zu Satz 4, Forster II, §10, denn stationäre Punkte sind Kandidaten für $I(u) = \inf_v I(v)$.

Kapitel L: Lebesgue-Maß, Abstrakte Maße

- **Definition 1.1**: Die durch $\lambda^*(A) := \inf\left\{\sum_{k=1}^{\infty}(b_k - a_k) \ : \ \bigcup_{k=1}^{\infty}]a_k, b_k[\ \supset \ A\right\}$ definierte Funktion $\lambda^* : \mathcal{P}(\mathbb{R}) \to \mathbb{R} \cup \{\infty\}$ heißt **äußeres Lebesgue-Maß** einer Menge $A \subset \mathbb{R}$.

- **Lemma 1.2**: Das äußere Maß besitzt die folgenden Eigenschaften

  (i) $\forall A, B \in \mathcal{P}(\mathbb{R}) \ : \ \big(A \subset B \implies \lambda^*(A) \ \le \ \lambda^*(B)\big)$. (ii) $\forall x \in \mathbb{R} \ : \ \lambda^*(\{x\}) = 0$.

  (iii) $\forall a, b \in \mathbb{R} \ : \ \big(a < b \implies \lambda^*([a,b]) = \lambda^*([a,b[) = \lambda^*(]a,b]) = \lambda^*(]a,b[) = b - a\big)$.

  (iv) $\forall A \in \mathcal{P}(\mathbb{R}) \forall x \in \mathbb{R} \ : \ \lambda^*(A) \ = \ \lambda^*(A + x)$ (Translationsinvarianz).

  (v) $\forall (A_k)_{k\in\mathbb{N}}, A_k \in \mathcal{P}(\mathbb{R}) \ : \ \lambda^*\Big(\bigcup_{k\in\mathbb{N}} A_k\Big) \ \le \ \sum_{k\in\mathbb{N}} \lambda^*(A_k)$ (σ-Subadditivität).

- **Definition 1.3**: Eine Teilmenge $A \in \mathcal{P}(\mathbb{R})$ heißt **(Lebesgue-)messbar** genau dann, wenn $\lambda^*(I) \ge \lambda^*(A \cap I) + \lambda^*(I \setminus A)$ für jedes beschränkte Intervall $I \subset \mathbb{R}$. Die Menge aller messbaren Mengen (ist eine σ-Algebra gemäß Definition 2.1, Theorem 1.4) wird mit $\mathcal{M}(\mathbb{R})$ bezeichnet. Die Funktion $\lambda : \mathcal{M}(\mathbb{R}) \to \mathbb{R} \cup \{\infty\}$, $\lambda : M \mapsto \lambda^*(M)$, heißt **Lebesgue-Maß**.

- **Theorem 1.4**: (i) Alle Intervalle sind Elemente aus $\mathcal{M}(\mathbb{R})$.
  (ii) Für jede Folge $(M_k)_{k\in\mathbb{N}}$ aus $\mathcal{M}(\mathbb{R})$ liegen $M_1 \setminus M_2$, $\bigcap_{k\in\mathbb{N}} M_k$, $\bigcup_{k\in\mathbb{N}} M_k$ in $\mathcal{M}(\mathbb{R})$.
  Sind zusätzlich alle M_k paarweise disjunkt (d.h., es gilt $M_k \cap M_l = \emptyset$ für $k \ne l$), so gilt
  $$\lambda\Big(\bigcup_{k\in\mathbb{N}} M_k\Big) \ = \ \sum_{k\in\mathbb{N}} \lambda(M_k) \qquad (\sigma\text{-Additivität}).$$

- **Definition 1.5**: Für ein Intervall $I := \bigtimes_{k=1}^{n} I^{(k)} \subset \mathbb{R}^n$ sei $\mathrm{vol}_n(I) := \prod_{k=1}^{n}(b_k - a_k)$ das n-dimensionale Volumen, wobei $a_k = \inf I^{(k)}$, $b_k = \sup I^{(k)}$ bezeichnen. Desweiteren heißt
  $$\lambda_n^* : \mathcal{P}(\mathbb{R}^n) \to \mathbb{R} \cup \{\infty\}, \ A \mapsto \lambda_n^*(A) := \inf\left\{\sum_{k\in\mathbb{N}} \mathrm{vol}(I_k) \ : \ \bigcup_{k\in\mathbb{N}} I_k \supset A, \ I_k \subset \mathbb{R}^n \text{offenes Intervall}\right\}$$
  das n-dimensionale **äußere Maß**, analog zu Definition 1.3 seien $\mathcal{M}(\mathbb{R}^n)$ und λ_n definiert.

- **Theorem 1.6**: Das äußere Lebesgue-Maß auf $\mathcal{P}(\mathbb{R}^n)$ ist σ-subadditiv.
- **Theorem 1.7**: Die Aussagen aus Theorem 1.4 (ii) gelten für $\mathcal{M}(\mathbb{R}^n)$ bzw. λ_n.
- **Theorem 1.8**: Für beschränkte Intervalle $I \subset \mathbb{R}^n$ stimmt $\lambda_n^*(I)$ mit $\mathrm{vol}_n(I)$ überein.
- **Theorem 1.9**: (i) Ist $A \subset \mathbb{R}^n$ offen, folgt $A \in \mathcal{M}(\mathbb{R}^n)$. (ii) Aus $\lambda_n^*(A) = 0$ folgt $A \in \mathcal{M}(\mathbb{R}^n)$.
- **Theorem 1.10**: Für $A \subset \mathbb{R}^n$ gilt $\lambda^*(A) \ = \ \inf\{\lambda(G) \mid G \subset \mathbb{R}^n \text{ offen }, A \subset G\}$.
- **Theorem 1.11**: Für eine Teilmenge $M \subset \mathbb{R}^n$ sind äquivalent: (i) M ist Lebesgue-messbar.

  (ii) Für jedes beschränkte Intervall I gilt $\lambda_n^*(I) \ge \lambda_n^*(I \cap M) + \lambda_n^*(I \setminus M)$.

  (iii) Zu jedem $\varepsilon > 0$ existiert eine offene Obermenge $G \supset M$ mit $\lambda_n^*(G \setminus M) < \varepsilon$.

- **Definition 2.1**: Sei $\Omega \ne \emptyset$ beliebig. Eine Teilmenge $\mathcal{A} \subset \mathcal{P}(\Omega)$ heißt **σ-Algebra**, falls

  (a) $\Omega \in \mathcal{A}$ (b) $\forall A \in \mathcal{A} \ : \ \Omega \setminus A \in \mathcal{A}$ (c) $\forall (A_k)_{k\in\mathbb{N}}, A_k \in \mathcal{A} \ : \ \bigcup_{k\in\mathbb{N}} A_k \in \mathcal{A}$

Zusatzaufgabe 10.1:

(a) Zeigen Sie, dass die Menge der orthogonalen Matrizen der Größe 3×3 eine dreidimensionale kompakte Untermannigfaltigkeit der Menge aller Matrizen der Größe 3×3 sind (wenn wir diese mit dem $\mathbb{R}^9$ identifizieren).

(b) Bestimmen Sie die lokalen Extrema von $f : M \to \mathbb{R}$, $f(x,y) := 3x^2 + 7y^2$, auf der nichtkompakten Untermannigfaltigkeit $M = \{(x,y) \in \mathbb{R}^2 \ : \ x^2 - xy - y^2 = 55\}$.

Lösung zu Zusatzaufgabe 10.1:

(a) Die Menge $O(3) = \{A \in M(3\times 3,\mathbb{R}) : A^{\mathrm{T}}A = I_3\}$ lässt sich (nach Identifizierung von $M(3\times 3,\mathbb{R})$ und $\mathbb{R}^9$) auch schreiben als $M := \{x \in \mathbb{R}^9 : g(x) = 0 \in \mathbb{R}^6\}$ mit der differenzierbaren Funktion $g : \mathbb{R}^9 \to \mathbb{R}^6$ definiert durch

$$g_1(x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9) = \begin{pmatrix} x_1^2+x_2^2+x_3^2-1 \\ x_4^2+x_5^2+x_6^2-1 \\ x_7^2+x_8^2+x_9^2-1 \\ x_1x_4+x_2x_5+x_3x_6 \\ x_1x_7+x_2x_8+x_3x_9 \\ x_4x_7+x_5x_8+x_6x_9 \end{pmatrix}.$$

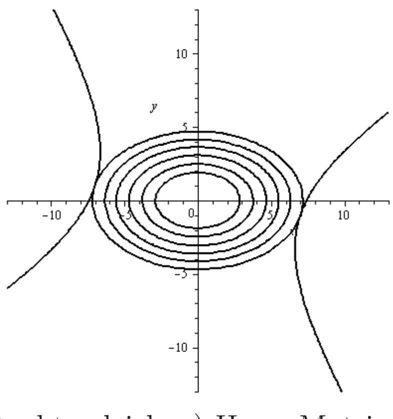
Deren Ableitung ist dann

$$dg(x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9) = \begin{pmatrix} 2x_1 & 2x_2 & 2x_3 & & & & & & \\ & & & 2x_4 & 2x_5 & 2x_6 & & & \\ & & & & & & 2x_7 & 2x_8 & 2x_9 \\ x_4 & x_5 & x_6 & x_1 & x_2 & x_3 & & & \\ x_7 & x_8 & x_9 & & & & x_1 & x_2 & x_3 \\ & & & x_7 & x_8 & x_9 & x_4 & x_5 & x_9 \end{pmatrix}.$$

Um die Surjektivität zu zeigen, müssen wir die lineare Unabhängigkeit der Zeilen zeigen. Für die Zeilen $a_k, k = 1,\ldots,6$ gelte mit $\lambda \in \mathbb{R}^6$ nun $\sum \lambda_k a_k = 0 \in \mathbb{R}^9$. Aufgrund der linearen Unabhängigkeit der Teilvektoren (x_1,x_2,x_3), (x_4,x_5,x_6) und (x_7,x_8,x_9) folgen (aus den ersten drei Zeilen des Gleichungssystems $\sum \lambda_k a_k = 0 \in \mathbb{R}^9$) sofort $\lambda_1 = \lambda_4 = \lambda_5 = 0$. Analog ergeben sich (aus den Zeilen vier bis sechs des Gleichungssystems $\sum \lambda_k a_k = 0 \in \mathbb{R}^9$) dann $\lambda_2 = \lambda_6 = 0$ und schließlich ebenfalls $\lambda_3 = 0$. Damit ist $M = g^{-1}(\{0\})$ das Urbild eines regulären Wertes einer C^∞-Funktion von $\mathbb{R}^9 \to \mathbb{R}^{9-3}$ und demzufolge nach Satz 2.66 eine dreidimensionale C^∞-Untermannigfaltigkeit des $\mathbb{R}^9$. Als Urbild einer abgeschlossenen (da einelementigen Menge) unter einer stetigen Funktion ist M selbst auch abgeschlossen. Da für alle $x \in M$ ebenso $\|x\|_2^2 = g_1(x)+1+g_2(x)+1+g_3(x)+1 = 3$ gilt, folgt auch die Beschränktheit, also nach dem Satz von Heine-Borel (Satz 1.58) die Kompaktheit von M.

(b) Zunächst bemerke man, dass M eine eindimensionale C^∞-Untermannigfaltigkeit ist, denn die Ableitung $dg(x,y) = \begin{pmatrix}2x-y & -x-2y\end{pmatrix}$ von $g(x,y) := x^2-xy-y^2-55$ ist nur im Punkt $(0,0)$ nicht surjektiv, welcher jedoch nicht in M liegt. Somit sind nach Satz 2.76 die stationären Punkte von $F(x,y,\lambda) = f(x,y) - \lambda g(x,y)$ Kandidaten für lokale Extrema von $f : M \to \mathbb{R}$. Da sich aus $dF(x,y,\lambda) = \begin{pmatrix}0 & 0 & 0\end{pmatrix}$ insbesondere für x und y das lineare Gleichungssystem $6x = \lambda(2x-y)$, $14y = \lambda(-x-2y)$ ergibt, welches nur für $\lambda_1 = -6$ oder $\lambda_2 = \frac{14}{5}$ singulär ist (und damit eine Lösung $\neq (0,0)$ besitzen kann), ergibt sich im Fall $\lambda = -6$ die Lösungsmenge $\{(x,y) \in M : 3x = y\} = \emptyset$ und für $\lambda = \frac{14}{5}$ die Lösungsmenge

$$\{(x,y) \in M : x = -7y\} = \{(7,-1),(-7,1)\},$$



wobei M genau an diesen Punkten die Niveaumenge

$$N_{154}f := \{(x,y) \in \mathbb{R}^2 : f(x,y) = 154\}$$

tangiert. Die entsprechenden Tangentialräume sind

$$\begin{aligned} T_{(7,-1)} &= \operatorname{Ker}\left(\begin{pmatrix}15 & -5\end{pmatrix}\right) = \operatorname{span}\left\{\begin{pmatrix}1\\3\end{pmatrix}\right\} \\ &= \operatorname{Ker}\left(\begin{pmatrix}-15 & 5\end{pmatrix}\right) = T_{(-7,1)}. \end{aligned}$$

Die Einschränkung der (in diesem Fall konstanten, für beide Punkte gleichen) Hesse-Matrix

$$\operatorname{Hess} F_{\frac{14}{5}}(x,y) = \operatorname{Hess} f(x,y) - \frac{14}{5}\operatorname{Hess} g(x,y) = \begin{pmatrix}6 & 0\\0 & 14\end{pmatrix} - \frac{14}{5}\begin{pmatrix}2 & -1\\-1 & -2\end{pmatrix} = \frac{2}{5}\begin{pmatrix}1 & 7\\7 & 49\end{pmatrix}$$

auf den Tangentialraum $T_{(7,-1)} = T_{(-7,1)}$ ist an beiden Punkten positiv definit, denn es gilt

$$c\begin{pmatrix}1 & 3\end{pmatrix}\frac{2}{5}\begin{pmatrix}1 & 7\\ 7 & 49\end{pmatrix}c\begin{pmatrix}1\\ 3\end{pmatrix} \;=\; \frac{2c^2}{5}(1+2\cdot 21+49\cdot 9) \;=\; \frac{968}{5}c^2 \;>\; 0$$

für alle $c \neq 0$, d.h., mit $H = \operatorname{Hess} F_{\frac{14}{5}}(x,y)$ gilt $v^T Hv > 0$ für alle $v \in T_{(7,-1)} = T_{(-7,1)}$ mit $v \neq 0$. Nach Satz 2.77 besitzt $f : M \to \mathbb{R}$ lokale Minima bei $(7,-1)$ und $(-7,1)$.

Zusatzaufgabe 10.2:

(a) Sei $g : \mathbb{R}^2 \to \mathbb{R}$ die Funktion $\qquad u(t,x) := \begin{cases} \frac{tx^3}{(x^2+t^2)^2}\,, & \text{falls } (t,x) \neq (0,0)\,,\\ 0\,, & \text{falls } (t,x) = (0,0)\,. \end{cases}$

Zeigen Sie, dass die Integrale $f(x) := \int_0^1 u(t,x)dt$ und $g(x) := \int_0^1 \frac{\partial}{\partial x}u(t,x)dt$ für jedes $x \in \mathbb{R}$ wohldefiniert sind, die Funktion $f : \mathbb{R} \to \mathbb{R}$ differenzierbar ist, jedoch $f'(0) \neq g(0)$ gilt.

(b) (Hamiltonsches Prinzip kleinster Wirkung)
Der Zustand eines physikalischen Systems wird üblicherweise durch von der Zeit abhängige Größen $t, \varphi_1(t), \dots, \varphi_n(t)$ beschrieben. Beispielsweise die Bewegung eines Massenpunktes sowie seine Geschwindigkeit zur Zeit t fassen wir im Allgemeinen als (genügend oft stetig differenzierbare) vektorwertige Funktionen $x : [a,b] \to \mathbb{R}^3,\ t \mapsto x(t) := (x_1(t), x_2(t), x_3(t))$ und $v : [a,b] \to \mathbb{R}^3,\ t \mapsto v(t) := \dot{x}(t)$ auf. Nun interessiert man sich für das Minimum des mit Hilfe der sogenannten Lagrange-Funktion $L(t,\varphi(t),\varphi'(t))$ gebildeten sogenannten Wirkungsintegrales $S(\varphi) = \int_a^b L(t,\varphi(t),\varphi'(t))dt$. Kandidaten für Minimalpunkte sind – falls L nicht explizit von der Zeit abhängt – genau die Lösungen der Euler-Lagrange-Gleichungen (10.2).

(i) Geben Sie die Euler-Lagrange-Gleichung bezüglich des Funktionales $S(\varphi)$ für mechanische Systeme an, bei denen die Bewegung unter dem Einfluss der kinetische Energie $T = \frac{1}{2}m\|v\|_2^2$ (mit Masse m) und eines vom Ort abhängigen Potentials $U : \mathbb{R}^3 \to \mathbb{R}$, $x \mapsto U(x)$ betrachtet wird und bei denen für die Lagrange-Funktion $L = T - U$ gilt.

(ii) Zeigen Sie, dass aus (i) die Konstanz der Gesamtenergie $E = T + U$ folgt.

Lösung zu Zusatzaufgabe 10.2:

(a) Das erste Integral ist für alle $x \in \mathbb{R}$ wohldefiniert wegen $f(0) = \int_0^1 u(t,0)dt = \int_0^1 0dt = 0$ einerseits und

$$f(x) \;=\; \int_0^1 u(t,x)dt \;=\; \int_0^1 \frac{tx^3}{(x^2+t^2)^2}dt \;=\; -\frac{x^3}{2(x^2+t^2)}\bigg|_{t=0}^{t=1} \;=\; \frac{x}{2} - \frac{x^3}{2(x^2+1)} \;=\; \frac{x}{2(x^2+1)}$$

für alle $x \neq 0$ andererseits. Mit $\frac{\partial}{\partial x}u(0,0) = \lim\limits_{h\to 0}\frac{u(0,h)-u(0,0)}{h} = 0$ und da die Quotientenregel

$$\frac{\partial}{\partial x}u(t,x) \;=\; \frac{\partial}{\partial x}\Big(\frac{tx^3}{(x^2+t^2)^2}\Big) \;=\; \frac{3tx^2(x^2+t^2)^2 - 4tx^4(x^2+t^2)}{(x^2+t^2)^4} \;=\; \frac{tx^2(3t^2-x^2)}{(x^2+t^2)^3}$$

für alle $(t,x) \neq (0,0)$ liefert, folgt für alle $x \in \mathbb{R}$ die Stetigkeit von $t \mapsto \frac{\partial}{\partial x}u(t,x)$ und somit auch die Wohldefiniert von $g : \mathbb{R} \to \mathbb{R}$. Insbesondere ergibt sich $g(0) = 0$ wegen $\frac{\partial}{\partial x}u(t,0) \equiv 0$. Andererseits ist die Funktion $f : \mathbb{R} \to \mathbb{R}$ differenzierbar mit Ableitung

$$f'(x) \;=\; \begin{cases} \frac{1}{2}\left(\frac{x}{(x^2+1)}\right)' \;=\; \frac{1}{2}\frac{(x^2+1)-2x^2}{(x^2+1)^2} \;=\; \frac{1-x^2}{2(x^2+1)}\,, & \text{falls } x \neq 0\,,\\ \lim\limits_{h\to 0}\frac{f(h)-f(0)}{h} \;=\; \lim\limits_{h\to 0}\frac{\frac{h}{2(h^2+1)}-0}{h} \;=\; \frac{1}{2}\,, & \text{falls } x = 0\,. \end{cases}$$

(b) Es ist die stetig differenzierbare (und nicht mehr explizit von t abhängige) Funktion

$$L : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R},\ (q,v) \mapsto L(q,v) \;:=\; T(v) - U(q) \;=\; \tfrac{1}{2}m(v_1^2+v_2^2+v_3^2) - U(q_1,q_2,q_3)$$

zu betrachten. Demnach ergeben sich $\frac{\partial L}{\partial q_k} = -\frac{\partial U}{\partial q_k}$ und $\frac{\partial L}{\partial v_k} = mv_k$, also $\frac{d}{dt}\frac{\partial L}{\partial v_k} = m\dot{v}_k = m\ddot{x}_k$.

(i) Die Euler-Lagrange-Gleichungen (10.2) sind zusammengefasst $m\ddot{x}(t) = -\nabla U(x(t))$.

(ii) Die Funktion $E : [a,b] \to \mathbb{R},\ t \mapsto \frac{1}{2}m\|\dot{x}(t)\|_2^2 + U(x(t))$ besitzt nach Kettenregel die Ableitung $\dot{E}(t) = m\langle \dot{x}(t), \ddot{x}(t)\rangle + \langle \nabla U(x(t)), \dot{x}(t)\rangle = \langle m\ddot{x} + \nabla U(x(t)), \dot{x}(t)\rangle$, welche wegen der nach (i) gültigen Beziehung $m\ddot{x}(t) = -\nabla U(x(t))$ verschwindet. Somit muss E (nach dem Hauptsatz der Differential- und Integralrechnung) konstant sein.

Zusatzaufgabe 10.3:

(a) Sind $\mathbb{N}, \mathbb{Q}$ und $\mathbb{R}$ Lebesgue-messbar ? Falls ja, bestimmen Sie jeweils das Lebesgue-Maß.

(b) Finden wir zu jeder nichtleeren Menge Ω eine σ-Algebra ?

(c) Geben Sie die maximale und die minimale σ-Algebra zu einer Menge $\Omega \neq \emptyset$ an.

Lösung zu Zusatzaufgabe 10.3:

(a) Zunächst bemerken wir, dass jede einelementige Menge $\{c\}$ nach Lemma 1.2 (ii) das äußere Maß Null besitzt. Zusammen mit Lemma 1.2 (iii) erhalten wir für jedes offene Intervall $]a,b[$ ($a < b$) und jede einelementige Menge $\{c\}$ demnach im Fall $]a,b[\cap\{c\} = \{c\}$ einerseits

$$\lambda^*(]a,b[) \;=\; b-a \;=\; b-c+c-a \;=\; \lambda^*(]a,c[\cup]c,b[) \;=\; \lambda^*(]a,b[\setminus\{c\}) \;+\; \lambda^*(]a,b[\cap\{c\})$$

und andernfalls wegen $\lambda^*(\emptyset) = 0$ ebenso

$$\lambda^*(]a,b[) \;=\; \lambda^*(]a,b[\setminus\{c\}) \;=\; \lambda^*(]a,b[\setminus\{c\}) \;+\; \lambda^*(]a,b[\cap\{c\}) \,.$$

Da dies ebenso für die restlichen Arten von beschränkten Intervallen gilt, folgt nach Definition die Lebesgue-Messbarkeit einelementiger Mengen. Da letztere also in der σ-Algebra $\mathcal{M}(\mathbb{R})$ liegen, tun dies – aufgrund der dritten Eigenschaft einer σ-Algebra (bzw. nach Theorem 1.4 (ii)) – auch die abzählbar unendlichen Mengen

$$\mathbb{N} \;=\; \bigcup_{k\in\mathbb{N}}\{n\} \qquad \text{und} \qquad \mathbb{Q} \;=\; \{0\} \cup \bigcup_{n\in\mathbb{N}}\left(\bigcup_{m\in\mathbb{N}}\left\{\frac{m}{n}\right\}\cup\left\{-\frac{m}{n}\right\}\right) \,,$$

welche aufgrund der σ-Subadditivität aus Lemma 1.2 (v) ebenso äußeres Lebesgue-Maß Null besitzen. Da das Lebesgue-Maß λ genau die Einschränkung des äußeren Maßes λ^* auf die σ-Algebra $\mathcal{M}(\mathbb{R})$ ist, folgen sofort $\lambda(\mathbb{N}) = \lambda^*(\mathbb{N}) = 0$ und $\lambda(\mathbb{Q}) = \lambda^*(\mathbb{Q}) = 0$.

Desweiteren lässt sich die Menge der reellen Zahlen als abzählbare Vereinigung

$$\mathbb{R} \;=\; \{0\} \bigcup_{n\in\mathbb{N}}]n-1,n] \cup [-n, 1-n[$$

mit paarweise disjunkten Mengen darstellen. Nach Theorem 1.4 (i) sind diese Mengen sämtlich Lebesgue-messbar, also in der σ-Algebra $\mathcal{M}(\mathbb{R})$. Aufgrund der dritten Eigenschaft einer σ-Algebra (bzw. nach Theorem 1.4 (ii)) ist demnach auch $\mathbb{R}$ in der σ-Algebra $\mathcal{M}(\mathbb{R})$, also Lebesgue-messbar. Mit der nach Theorem 1.4 (ii) gültigen σ-Additivität erhalten wir demnach

$$\begin{aligned}\lambda(\mathbb{R}) \;&=\; \lambda(\{0\}) \;+\; \sum_{n\in\mathbb{N}}\lambda(]n-1,n]\cup[-n,1-n[) \;=\; 0 \;+\; \sum_{n\in\mathbb{N}}(\lambda(]n-1,n]) + \lambda([-n,1-n[)) \\ &=\; \sum_{n\in\mathbb{N}}((n-(n-1)) + (1-n-(-n))) \;=\; \sum_{n\in\mathbb{N}}2 \;=\; \infty \,.\end{aligned}$$

(b) Für eine beliebige Menge $\Omega \neq \emptyset$ erfüllt stets die Potenzmenge $\mathcal{P}(\Omega) = \{A \mid A \subset \Omega\}$ alle drei Eigenschaften einer σ-Algebra.

(c) Die minimale σ-Algebra ist genau $\mathcal{A}_{\min} = \{\emptyset, \Omega\}$ und die maximale $\mathcal{A}_{\max} = \mathcal{P}(\Omega)$.

Zusatzaufgabe 10.4:

(a) Sei $\Omega = \{1,2,3,4\}$. Bestimmen Sie $\sigma(\mathfrak{E})$ für $\mathfrak{E} := \{\emptyset, \Omega, \{1,2\}\}$ und $\mathfrak{E} := \{\emptyset, \{1,2,4\}, \{3\}\}$.

(b) Zeigen Sie: Ist $\mathcal{A}$ eine σ-algebra auf einer Menge Ω und $E \subset \Omega$, dann ist

$$\mathcal{A} \sqcap E : \ = \{ A \cap E \mid A \in \mathcal{A} \} \qquad \text{(die **Spur-}\sigma\text{-Algebra**)}$$

eine σ-Algebra auf E. Für $E \in \mathcal{A}$ folgt $\mathcal{A} \sqcap E \ = \ \{ A \mid A \in \mathcal{A} \text{ und } A \subset E \}$.

Lösung zu Zusatzaufgabe 10.4:

(a) Zu $\Omega = \{1,2,3,4\}$ ist

- $\sigma(\{\emptyset, \Omega, \{1,2\}\}) \ = \ \{\emptyset, \Omega, \{1,2\}, \{3,4\}\}$, denn sie enthält die leere Menge, zu jedem Element auch das Komplement, und alle abzählbaren Vereinigungen $\bigcup_{n\in\mathbb{N}} A_n$, denn es können nur die Fälle auftreten:
  - $A_{n_0} = \Omega$ für ein n_0, also $\bigcup\limits_{n\in\mathbb{N}} A_n = \Omega$;
  - $A_{n_0} = \{1,2\}$ für ein n_0, also $\bigcup\limits_{n\in\mathbb{N}} A_n = \begin{cases} \Omega & \text{falls } A_{n_0'} = \Omega \text{ für ein } n_0', \\ \Omega & \text{falls } A_{n_0'} = \{3,4\} \text{ für ein } n_0', \\ \{1,2\} & \text{sonst;} \end{cases}$
  - $A_{n_0} = \{3,4\}$ für ein n_0, also $\bigcup\limits_{n\in\mathbb{N}} A_n = \begin{cases} \Omega & \text{falls } A_{n_0'} = \Omega \text{ für ein } n_0', \\ \Omega & \text{falls } A_{n_0'} = \{1,2\} \text{ für ein } n_0', \\ \{3,4\} & \text{sonst;} \end{cases}$
  - $A_n = \emptyset$ für alle $n \in \mathbb{N}$, also $\bigcup\limits_{n\in\mathbb{N}} A_n = \emptyset$.
- $\sigma(\{\emptyset, \{1,2,4\}, \{3\}\}) \ = \ \{\emptyset, \Omega, \{1,2,4\}, \{3\}\}$, denn sie enthält die leere Menge, zu jedem Element auch das Komplement, und alle abzählbaren Vereinigungen $\bigcup_{n\in\mathbb{N}} A_n$, denn es können nur die folgenden Fälle auftreten:
  - $A_{n_0} = \Omega$ für ein n_0, also $\bigcup\limits_{n\in\mathbb{N}} A_n = \Omega$;
  - $A_{n_0} = \{1,2,4\}$ für ein n_0, also $\bigcup\limits_{n\in\mathbb{N}} A_n = \begin{cases} \Omega & \text{falls } A_{n_0'} = \Omega \text{ für ein } n_0', \\ \Omega & \text{falls } A_{n_0'} = \{3\} \text{ für ein } n_0', \\ \{1,2,4\} & \text{sonst;} \end{cases}$
  - $A_{n_0} = \{3\}$ für ein n_0, also $\bigcup\limits_{n\in\mathbb{N}} A_n = \begin{cases} \Omega & \text{falls } A_{n_0'} = \Omega \text{ für ein } n_0', \\ \Omega & \text{falls } A_{n_0'} = \{1,2,4\} \text{ für ein } n_0', \\ \{3\} & \text{sonst;} \end{cases}$
  - $A_n = \emptyset$ für alle $n \in \mathbb{N}$, also $\bigcup\limits_{n\in\mathbb{N}} A_n = \emptyset$.

(b) Es ist zu zeigen, dass $\mathcal{A} \sqcap E$ die drei Eigenschaften einer σ-Algebra auf E besitzt.

(i) Wegen $\Omega \in \mathcal{A}$ und $\Omega \cap E = E$ gilt auch $E \in \mathcal{A} \sqcap E$.

(ii) Sei $B \in \mathcal{A} \sqcap E$. Dann existiert ein $A \in \mathcal{A}$ mit $B = A \cap E$. Da $\mathcal{A}$ eine σ-Algebra auf Ω ist, folgt auch $\Omega \setminus A \in \mathcal{A}$ und weiter $\Omega \setminus A \cap E \in \mathcal{A} \sqcap E$. Mit $E \subset \Omega$ ist jedoch $\Omega \setminus A \cap E = (\Omega \cap E) \setminus (A \cap E) = E \setminus B$, also $E \setminus B \in \mathcal{A} \sqcap E$.

(iii) Sei nun $(B_k)_{k\in\mathbb{N}}$ eine beliebige Folge aus $\mathcal{A} \sqcap E$, dann gibt es (mindestens) eine entsprechende Folge $(A_k)_{k\in\mathbb{N}}$ aus $\mathcal{A}$, so dass $B_k = A_k \cap E$. Da $\mathcal{A}$ eine σ-Algebra auf Ω ist, liegt auch die Vereinigung $\bigcup_{k\in\mathbb{N}} A_k$ wieder in $\mathcal{A}$ und somit folgt

$$\bigcup_{k\in\mathbb{N}} B_k \ = \ \bigcup_{k\in\mathbb{N}} (A_k \cap E) \ = \ \bigcup_{k\in\mathbb{N}} A_k \cap E \in \mathcal{A} \sqcap E \,.$$

Falls $E \in \mathcal{A}$, liegen alle Schnitte $A \cap E$ für $A \in \mathcal{A}$ wieder in $\mathcal{A}$, da $\mathcal{A}$ eine σ-Algebra ist.

Alle Übungsaufgaben sind verfügbar unter: `https://studip.uni-rostock.de/studip/`

Zusatzmaterial zur Analysis II – Serie 11

Kapitel L: Abstrakte Maße

- **Definition 2.1**: Sei $\Omega \neq \emptyset$ beliebig. Eine Teilmenge $\mathcal{A} \subset \mathcal{P}(\Omega)$ heißt σ**-Algebra**, falls

  (a) $\Omega \in \mathcal{A}$ (b) $\forall A \in \mathcal{A} \; : \; \Omega \setminus A \in \mathcal{A}$ (c) $\forall (A_k)_{k\in\mathbb{N}}, A_k \in \mathcal{A} \; : \; \bigcup\limits_{k\in\mathbb{N}} A_k \in \mathcal{A}$

  Das Paar $(\Omega, \mathcal{A})$ nennen wir einen **Messraum** und jedes $A \in \mathcal{A}$ heißt **messbar**. Zu einer beliebigen Teilmenge $\mathfrak{E} \subset \mathcal{P}(\Omega)$ bezeichnet $\sigma(\mathfrak{E})$ die kleinste $\mathfrak{E}$ enthaltende σ-Algebra und wird **die von $\mathfrak{E}$ erzeugte σ-Algebra** genannt. Dabei gilt
  $$\sigma(\mathfrak{E}) = \bigcap_{\substack{\mathfrak{E} \subset \mathcal{A} \\ \mathcal{A}\ \sigma-\text{Algebra in } \Omega}} \mathcal{A}$$

- **Definition 2.2**: Sei $X := (M, \mathcal{T})$ ein topologischer Raum. Dann nennen wir $\mathcal{B}(\mathcal{T}) := \sigma(\mathcal{T})$ die **Borel-σ-Algebra** von X. Ihre Elemente werden **Borelmengen** genannt.

- **Definition 2.3**: Sei $(\Omega, \mathcal{A})$ ein Messraum. Die Funktion $\mu : \mathcal{A} \to [0, \infty]$ heißt **Maß**, falls

  (a) $\mu(\emptyset) = 0$

  (b) Für jede Folge A_k paarweise disjunkter Mengen aus $\mathcal{A}$ gilt: $\mu\Big(\bigcup\limits_{k\in\mathbb{N}} A_k\Big) = \sum\limits_{k\in\mathbb{N}} \mu(A_k)$.

  Das Tripel $(\Omega, \mathcal{A}, \mu)$ nennen wir einen Maßraum. Wir sagen weiterhin, der Maßraum

  - ist **endlich**, falls $\mu(\Omega) < \infty$ ist.
  - ist σ**-endlich**, falls es eine Folge $A_k \in \mathcal{A}$ mit $\mu(A_k) < \infty$ und $\Omega = \bigcup\limits_{k\in\mathbb{N}} A_k$ gibt.
  - ist ein **Wahrscheinlichkeitsraum**, falls $\mu(\Omega) = 1$.
  - ist **vollständig**, falls $\forall B \in \mathcal{A}\ \forall A \subset \Omega \; : \; (\mu(B) = 0 \wedge A \subset B \implies A \in \mathcal{A})$.

  Alle $A \in \mathcal{A}$ mit $\mu(A) = 0$ werden μ**-Nullmengen** genannt.

- **Lemma 2.4 [Eigenschaften eines Maßes]**: Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Dann gelten:

  (a) Für $(A_k)_{k\in\mathbb{N}}$ mit $A_k \in \mathcal{A}$ und $A_1 \subset A_2 \subset \ldots$ gilt $\mu\Big(\bigcup\limits_{k\in\mathbb{N}} A_k\Big) = \lim\limits_{k\to\infty} \mu(A_k)$.

  (b) Für $(A_k)_{k\in\mathbb{N}}$ mit $A_k \in \mathcal{A}, \mu(A_1) < \infty$ und $A_1 \supset A_2 \supset \ldots$ gilt $\mu\Big(\bigcap\limits_{k\in\mathbb{N}} A_k\Big) = \lim\limits_{k\to\infty} \mu(A_k)$.

  (c) Für $(A_k)_{k\in\mathbb{N}}$ mit $A_k \in \mathcal{A}$ gilt $\mu\Big(\bigcup\limits_{k\in\mathbb{N}} A_k\Big) \leq \sum\limits_{k\in\mathbb{N}} \mu(A_k)$.

Zusatzaufgabe 11.1: Sei $\Omega \neq \emptyset$ beliebig.

Zeigen Sie, dass $\mathcal{A} := \{A \subset X \mid A \text{ abzählbar oder } \Omega \setminus A \text{ abzählbar } \}$ eine σ-Algebra ist.
Hinweis: Die leere Menge ist abzählbar.

Lösung zu Zusatzaufgabe 11.1:

Es sind die drei Eigenschaften einer σ-Algebra zu überprüfen. Es gilt $\emptyset \in \mathcal{A}$, da sie abzählbar ist. Für $A \in \mathcal{A}$ beliebig gilt A abzählbar oder $\Omega \setminus A$ abzählbar, also offensichtlich auch $B := \Omega \setminus A$ abzählbar oder $\Omega \setminus B = A$ abzählbar und damit auch $B \in \mathcal{A}$, also $\Omega \setminus A \in \mathcal{A}$. Sei nun $(A_k)_{k\in\mathbb{N}}$ eine beliebige Folge aus $\mathcal{A}$. Sind nun alle A_k abzählbar, dann auch ihre abzählbare Vereinigung, also $\bigcup A_k \in \mathcal{A}$. Andernfalls existiert ein k_0, so dass $\Omega \setminus A_{k_0}$ abzählbar ist. Da Teilmengen abzählbarer Mengen selbst abzählbar sind, folgt nun auch die Abzählbarkeit von $\Omega \setminus (\bigcup A_k) \subset \Omega \setminus A_{k_0}$ uns somit ebenfalls $\bigcup A_k \in \mathcal{A}$.

Zusatzaufgabe 11.2:

(a) Sei $(\Omega, \mathcal{A})$ ein beliebiger Messraum. Zeigen Sie:

(i) Endliche Vereinigungen und endliche Durchschnitte messbarer Mengen sind messbar.

(ii) Sind A und B messbar, dann ist auch $A \setminus B$ messbar.

(iii) Abzählbare Durchschnitte messbarer Mengen sind wieder messbar.

(b) Gegeben sei der topologische Raum $X := (\mathbb{R}, \mathcal{T})$ mit $\mathcal{T} := \mathcal{T}_{d_{|\cdot|}}$ und $\mathcal{B}(\mathcal{T}) := \sigma(\mathcal{T})$ die entsprechende Borel-σ-Algebra von X. Zeigen oder widerlegen Sie: $\mathcal{B}(\mathcal{T})$ enthält

(i) alle abgeschlossenen Teilmengen von X.

(ii) alle abzählbaren Durchschnitte von offenen Mengen (diese heißen G_δ-Mengen).

(iii) alle abzählbaren Vereinigungen von abgeschlossenen Mengen (diese heißen F_σ-Mengen).

(iv) alle abzählbaren Vereinigungen von G_δ-Mengen.

(v) alle abzählbaren Durchschnitte von F_σ-Mengen.

Lösung zu Zusatzaufgabe 11.2:

(a) (i) Wegen $A \cup \emptyset = A$ für eine beliebige Menge $A \in \mathcal{A}$ können wir eine endliche Menge $\{A_1, \ldots, A_p\}$ mit $A_k \in \mathcal{A}$, $k = 1, \ldots, p$, durch $A_k := \emptyset \in \mathcal{A}$ für $k = p+1, p+2, \ldots$ zu einer abzählbaren Menge (A_k) ergänzen und erhalten mit der Definition einer σ-Algebra

$$\bigcup_{k=1}^{p} A_k = \bigcup_{k\in\mathbb{N}} A_k \in \mathcal{A} . \tag{11.1}$$

Nach Definition einer σ-Algebra ist mit einer Menge $A \in \mathcal{A}$ auch $\Omega \setminus A \in \mathcal{A}$. Daher folgt für eine endliche Menge $\{A_1, \ldots, A_p\}$ mit $A_k \in \mathcal{A}$, $k = 1, \ldots, p$ mit (11.1) dann

$$\bigcap_{k=1}^{p} A_k = \underbrace{\Omega \setminus \left(\bigcup_{k=1}^{p} \underbrace{(\Omega \setminus A_k)}_{\in\mathcal{A}} \right)}_{\in\mathcal{A}} \in \mathcal{A} . \tag{11.2}$$

(ii) Die Behauptung folgt sofort mit (i) wegen $A \setminus B = A \cap (\Omega \setminus B) \in \mathcal{A}$.

(iii) Nach Definition einer σ-Algebra ist mit einer Menge $A \in \mathcal{A}$ auch $\Omega \setminus A \in \mathcal{A}$ und liegt für jede Folge $(A_n)_{n\in\mathbb{N}}$ von Mengen aus $\mathcal{A}$ auch $\bigcup\limits_{n\in\mathbb{N}} A_n$ in $\mathcal{A}$. Für jede Folge $(A_n)_{n\in\mathbb{N}}$ von Mengen aus $\mathcal{A}$ folgt somit

$$\bigcap_{n\in\mathbb{N}} A_n = \underbrace{\Omega \setminus \left(\bigcup_{n\in\mathbb{N}} \underbrace{(\Omega \setminus A_n)}_{\in\mathcal{A}} \right)}_{\in\mathcal{A}} \in \mathcal{A} . \tag{11.3}$$

(b) (i) Da die Borel-σ-Algebra genau von den offenen Mengen erzeugt ist, sind alle offenen Mengen in $\mathcal{B}(\mathcal{T})$ enthalten. Da mit jedem $A \in \mathcal{B}(\mathcal{T})$ aufgrund der Eigenschaften einer σ-Algebra auch $\mathbb{R} \setminus A$ in $\mathcal{B}(\mathcal{T})$ enthalten sein muss und da die abgeschlossenen Mengen genau die Komplemente der offenen Mengen sind, folgt die Behauptung.

(ii) Ist eine direkte Folgerung aus (a)(iii) bzw. (11.3), da offene Mengen Borelmengen sind.

(iii) Ist eine direkte Folgerung aus (i) und der Eigenschaft einer σ-Algebra.

(iv) Da nach (ii) bereits alle G_δ-Mengen in $\mathcal{B}(\mathcal{T})$ liegen, folgt die Behauptung aus den Eigenschaften einer σ-Algebra.

(v) Da nach (iii) bereits alle F_σ-Mengen in $\mathcal{B}(\mathcal{T})$ liegen, folgt die Behauptung aus (a)(iii).

Zusatzaufgabe 11.3:

(a) Sei $(\Omega, \mathcal{A}, \mu)$ ein beliebiger Maßraum. Zeigen Sie:

  (i) μ ist additiv, d.h., für paarweise disjunkte $A_k \in \mathcal{A},\ k = 1, \ldots, n$ gilt $\mu\Big(\bigcup\limits_{k=1}^{n} A_k\Big) = \sum\limits_{k=1}^{n} \mu(A_k)$.

  (ii) μ ist isoton, d.h., $\forall A, B \in \mathcal{A}\ :\ (A \subset B \implies \mu(A) \leq \mu(B))$

  (iii) Ist $(\Omega, \mathcal{A}, \mu)$ endlich, dann besitzt jede messbare Menge endliches Maß.

(b) Sei $(\Omega, \mathcal{A})$ ein beliebiger Messraum mit $\mathcal{A} \neq \mathcal{P}(\Omega)$. Geben Sie ein Beispiel für ein Maß μ an, so dass $(\Omega, \mathcal{A}, \mu)$ endlich, aber nicht vollständig ist.

(c) Zeigen Sie: Sind μ und ν zwei endliche Maße auf dem Messraum $(\Omega, \mathcal{A})$ und a, b nichtnegative reelle Zahlen, dann ist auch $\lambda := a\mu + b\nu$ ein endliches Maß auf $(\Omega, \mathcal{A})$.

(d) Wann wird das Maß λ aus (c) zu einem Wahrscheinlichkeitsmaß?

Lösung zu Zusatzaufgabe 11.3:

(a) (i) Wegen $A \cup \emptyset = A$ und $A \cap \emptyset = \emptyset$ sowie wegen $\emptyset \in \mathcal{A}$ und $\mu(\emptyset) = 0$ folgt die Additivität sofort aus der σ-Additivität, wenn wir $A_k = \emptyset$ für $k = n+1, n+2, \ldots$ setzen, denn dann gilt

$$\mu\Big(\bigcup_{k=1}^{n} A_k\Big) = \mu\Big(\bigcup_{k\in\mathbb{N}} A_k\Big) = \sum_{k\in\mathbb{N}} \mu(A_k) = \sum_{k=1}^{n} \mu(A_k) + \sum_{k=n+1}^{\infty} \underbrace{\mu(A_k)}_{=0} = \sum_{k=1}^{n} \mu(A_k)$$

  (ii) Mit $A, B \in \mathcal{A}$ sind stets auch $A_1 := B \cap A$ und $A_2 := B \setminus A$ in $\mathcal{A}$. Desweiteren gelten offenbar $A_1 \cap A_2 = \emptyset$ sowie $A_1 \cup A_2 = B$. Also folgt mit der Nichtnegativität und der nach (i) gültigen Additivität eines Maßes sofort auch

$$\mu(B) = \mu(A_1 \dot{\cup} A_2) = \mu(A_1) + \underbrace{\mu(A_2)}_{\geq 0} \geq \mu(A_1)\ ,$$

  was im Fall $A \subset B$ wegen $B \cap A = A$ mit der Behauptung übereinstimmt.

  (iii) Im Fall eines endliches Maßraumes $(\Omega, \mathcal{A}, \mu)$ gilt $\mu(\Omega) < \infty$. Da für ein beliebiges $A \in \mathcal{A}$ offensichtlich $A \subset \Omega$ gilt, folgt aus der in (ii) bewiesenen Isotonie von Maßen demnach $\mu(A) \leq \mu(\Omega) < \infty$.

(b) Wählen wir $\mu : \mathcal{A} \to [0, \infty],\ A \mapsto \mu(A) := 0$, dann ist jede $\mathcal{A}$-messbare Menge eine $\mathcal{A}$-Nullmenge, somit auch Ω selbst. Also gilt einerseits $\mu(\Omega) = 0 < \infty$, andererseits müsste im Fall der Vollständigkeit jede Teilmenge von Ω in der σ-Algebra $\mathcal{A}$, also $\mathcal{A} \supset \mathcal{P}(\Omega)$ gelten. Da jedoch stets $\mathcal{A} \subset \mathcal{P}(\Omega)$ gilt, folgte dann $\mathcal{A} = \mathcal{P}(\Omega)$, im Widerspruch zur Voraussetzung.

(c) Wegen $\lambda(\emptyset) = a\mu(\emptyset) + b\nu(\emptyset) = a \cdot 0 + b \cdot 0$ ergibt sich zunächst $\lambda(\emptyset) = 0$. Da a, b nichtnegative reelle (und damit insbesondere endliche) Zahlen sind, folgt mit der Endlichkeit und Isotonie von μ und ν ebenso die Nichtnegativität und Endlichkeit von λ, also $\lambda : \mathcal{A} \to [0, \infty]$, denn für beliebiges $A \in \mathcal{A}$ gilt nach (a)(iii)

$$\lambda(A) = a \underbrace{\mu(A)}_{<\infty} + b \underbrace{\nu(A)}_{<\infty} < \infty$$

Für eine Folge $(A_k)_{k\in\mathbb{N}}$ paarweise disjunkter $\mathcal{A}$-messbarer Mengen gilt – aufgrund der Endlichkeit des Maßes und der Nichtnegativität aller beteiligten Größen – offensichtlich auch

$$\infty > \lambda\Big(\bigcup_{k\in\mathbb{N}} A_k\Big) = a\mu\Big(\bigcup_{k\in\mathbb{N}} A_k\Big) + b\nu\Big(\bigcup_{k\in\mathbb{N}} A_k\Big) = a\sum_{k\in\mathbb{N}} \mu(A_k) + b\sum_{k\in\mathbb{N}} \nu(A_k)$$
$$= \sum_{k\in\mathbb{N}} (a\mu(A_k) + b\nu(A_k)) = \sum_{k\in\mathbb{N}} \lambda(A_k)\ ,$$ womit die σ-Additivität gezeigt ist.

(d) Stimmen sowohl μ als auch ν mit dem Nullmaß überein, so haben wir keine Chance. Ist jedoch $\mu(\Omega) > 0$, so können wir $a = \frac{1}{\mu(\Omega)}$ und $b = 0$ wählen. Gilt zusätzlich $\nu(\Omega) > 0$, so ist jede Wahl $a = \frac{p}{\mu(\Omega)}, b = \frac{q}{\nu(\Omega)}$ mit $p + q = 1$ und $p, q \geq 0$ möglich.

Zusatzaufgabe 11.4: (Konstruktion einer Menge $A \in \mathcal{P}(\mathbb{R}) \setminus \mathcal{M}(\mathbb{R})$)

Auf $\mathbb{R}$ sei die Äquivalenzrelation $a \sim b \;:\Longleftrightarrow\; a - b \in \mathbb{Q}$ gegeben. Nach dem Auswahlaxiom ist es möglich, eine Menge $A \subset [0,1]$ so zu konstruieren, dass sie genau ein Element aus jeder Äquivalenzklasse aus jeder Äquivalenzklasse $[r]_\sim = \{s \in \mathbb{R} \mid r \sim s\}$ enthält. Zeigen Sie:

(a) Mit $t + A = \{t + s \mid s \in A\}$ folgt $[0,1] \subset \bigcup_{t \in \mathbb{Q} \cap [-1,1]} (t + A) \subset [-1, 2]$ (b) $A \notin \mathcal{M}(\mathbb{R})$

Lösung zu Zusatzaufgabe 11.4:

(a) (i) Sei $C = [r]_\sim$ eine beliebige Äquivalenzklasse und $c \in C$. Dann finden wir mindestens ein $q \in [c-1, c]$, also $c - 1 \leq q \leq c$. Dies ist äquivalent zu $0 \leq c - q \leq 1$. Setzen wir nun $s := c - q$, so haben wir ein $s \in C$ mit $s \in [0,1]$ gefunden.

(ii) Seien $s, t \in \mathbb{Q}$ mit $s \neq t$.

- Angenommen, es existiert ein $v \in (s + A) \cap (t + A)$. Dann existieren $a_s, a_t \in A$, so dass $v = s + a_s = t + a_t$. Dies ist aber äquivalent zu $a_s - a_t = t - s \in \mathbb{Q}$, also $a_s \sim a_t$ und somit $[a_s]_\sim = [a_t]_\sim$. Nach Konstruktion enthält jedoch A aus jeder Äquivalenzklasse bezüglich $\sim$ nur genau ein Element. Daher musste $a_s = a_t$ und somit auch $s = t$ gegolten haben. Dies ist jedoch ein Widerspruch.

Also haben wir gezeigt, dass $s, t \in \mathbb{Q} \wedge s \neq t \implies (s + A) \cap (t + A) = \emptyset$ gilt.

(iii) Sei $r \in [0,1]$ beliebig, jedoch fest gewählt.

- Nach Konstruktion von A existiert dann genau ein $a_r \in A$ mit $r \sim a_r$. Folglich existiert ein $t \in \mathbb{Q}$ mit $r = a_r + t$. Wegen $r \in [0,1]$ und $a_r \in A \subset [0,1]$ muss dann $t = r - a_r \in [-1,1]$ gelten. Damit ist $r \in t + A$ für ein $t \in [-1,1] \cap \mathbb{Q}$.

Da $r \in [0,1]$ beliebig war, haben wir gezeigt, dass $[0,1] \subset \bigcup_{t \in \mathbb{Q} \cap [-1,1]} (t + A)$.

(iv) Sei $t \in \mathbb{Q} \cap [-1,1]$ beliebig. Wegen $A \subset [0,1]$ gilt offenbar $-1 \leq t + a \leq 2$ für jedes $a \in A$, also auch $t + A \subset [-1, 2]$. Da die Vereinigung gleichmäßig beschränkter Mengen selbst wieder gleichmäßig beschränkt ist, folgt auch $\bigcup_{t \in \mathbb{Q} \cap [-1,1]} (t + A) \subset [-1, 2]$.

(b) Angenommen, es gelte $A \in \mathcal{M}(\mathbb{R})$. Aufgrund der Translationsinvarianz wäre dann auch für beliebiges $t \in \mathbb{Q}$ die Menge $t + A \in \mathcal{M}(\mathbb{R})$ mit $\lambda_1(A) = \lambda_1(t + A)$.

- Angenommen, es gelte $\lambda_1(A) = 0$. Mit der σ-Additivität (denn in (a) haben wir die paarweise Disjunktheit der Mengen $t + A$ gezeigt) sowie aufgrund der Isotonie eines Maßes (vgl. ZA 11.3 (a)) ergäbe sich – da $\mathbb{Q}$ abzählbar, und somit auch jede Teilmenge von $\mathbb{Q}$ abzählbar – in Zusammenhang mit der Translationsinvarianz nach (a) dann der Widerspruch

$$1 = \lambda_1([0,1]) \leq \lambda_1\Bigg(\bigcup_{t \in \mathbb{Q} \cap [-1,1]} (t + A) \Bigg) \overset{\text{(ii)}}{=} \sum_{t \in \mathbb{Q} \cap [-1,1]} \lambda_1(t + A) = \sum_{t \in \mathbb{Q} \cap [-1,1]} \lambda_1(A) = 0 .$$

- Angenommen, es gelte $\lambda_1(A) > 0$. Wiederum mit der σ-Additivität und der Isotonie eines Maßes (vgl. ZA 11.3 (a)) ergäbe sich in Zusammenhang mit der Translationsinvarianz nach (a) dann der Widerspruch

$$3 = \lambda_1([-1,2]) \geq \lambda_1\Bigg(\bigcup_{t \in \mathbb{Q} \cap [-1,1]} (t + A) \Bigg) \overset{\text{(ii)}}{=} \sum_{t \in \mathbb{Q} \cap [-1,1]} \lambda_1(t + A) = \sum_{t \in \mathbb{Q} \cap [-1,1]} \lambda_1(A) = \infty .$$

Alle Übungsaufgaben sind verfügbar unter: `https://studip.uni-rostock.de/studip/`

Zusatzmaterial zur Analysis II – Serie 12

Kapitel L: Messbare Funktionen, Integrierbarkeit (einfacher Funktionen)

- **Definition 2.5**: Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, $\mathcal{N} = \{A \subset \Omega \mid \exists B \in \mathcal{A} : (\mu(B) = 0 \wedge A \subset B)\}$. Weiter sei $\overline{\mathcal{A}} := \{M \cup N \mid M \in \mathcal{A} \wedge N \in \mathcal{N}\}$. Dann heißt die durch
$$\overline{\mu}(M \cup N) = \mu(M) \qquad \text{für } M \in \mathcal{A},\ N \in \overline{\mathcal{A}}$$
definierte Funktion $\overline{\mu} : \overline{\mathcal{A}} \to [0, \infty]$ die **Vervollständigung von** μ.

- **Theorem 2.6 [Vervollständigung von Maßen]**: Mit den in Definition 2.5 eingeführten Bezeichnungen folgt:
  (a) $\overline{\mathcal{A}}$ ist eine σ-Algebra, welche die σ-Algebra $\mathcal{A}$ enthält.
  (b) $(\Omega, \overline{\mathcal{A}}, \overline{\mu})$ ist vollständig.
  (c) Es gilt $\overline{\mu}|_{\mathcal{A}} = \mu$, d.h., die Maße $\overline{\mu}$ und μ stimmen auf $\mathcal{A}$ überein.

- **Definition 3.1**: Sei $(\Omega, \mathcal{A})$ ein Messraum und $D \in \mathcal{A}$. Eine Funktion $f : D \to \mathbb{R} \cup \{-\infty, \infty\}$ heißt **$\mathcal{A}$-messbar auf** D, falls $\forall \alpha \in \mathbb{R} : f^{-1}(]\alpha, \infty]) \in \mathcal{A}$ gilt.[1]

- **Theorem 3.2**: Sei $(\Omega, \mathcal{A})$ ein Messraum und seien f, g beliebige $\mathcal{A}$-messbare Funktionen auf Ω. Dann gelten:
  (a) $\{\omega \in \Omega : f(\omega) < g(\omega)\} \in \mathcal{A}$. (ii) $f^{-1}(\{\infty\}) \in \mathcal{A}$ und $f^{-1}(\{-\infty\}) \in \mathcal{A}$.

- **Lemma 3.3 [Eigenschaften von $\mathcal{A}$-messbaren Funktionen]**: Sei $(\Omega, \mathcal{A})$ ein Messraum und seien $(f_n)_{n\in\mathbb{N}}$ und f, g eine Folge von bzw. beliebige $\mathcal{A}$-messbare Funktionen (mit möglicherweise unterschiedlichem Definitionsbereich). Desweiteren sei $\lambda \in \mathbb{R}$ und φ eine stetige Funktion auf einer offenen Menge $G \subset \mathbb{R}$. Dann sind die Funktionen
  (a) λf, $f + g$, $\max\{f, g\}$, $\min\{f, g\}$, $|f|$, fg, $\frac{f}{g}$
  (b) $\sup_{n\in\mathbb{N}} f_n$, $\inf_{n\in\mathbb{N}} f_n$, $\limsup_{n\to\infty} f_n$, $\liminf_{n\to\infty} f_n$, $\lim_{n\to\infty} f_n$
  (c) $\varphi \circ f$

  – falls sie definiert sind – $\mathcal{A}$-messbar (und ihr Definitionsbereich ist jeweils in $\mathcal{A}$).

- **Definition/Lemma 3.4 [Einfache Funktionen]**: Sei $(\Omega, \mathcal{A})$ ein Messraum und e eine $\mathcal{A}$-messbare (reellwertige) Funktion mit endlichem Bild $f(\Omega)$, d.h., es existieren reelle paarweise verschiedene Zahlen $c_k,\ k = 1, \dots, n$, so dass
$$f(\Omega) = \bigcup_{k\in\{1,\dots,n\}} \{c_k\} ,$$
dann heißt e **einfache Funktion auf** Ω.
Sind $\alpha_j,\ j = \{1, \dots, k\}$, die aufsteigend geordneten paarweise verschiedenen Elemente von $e(\Omega)$ und $A_j = e^{-1}(\{\alpha_j\}),\ j = \{1, \dots, k\}$, die entsprechenden (offenbar paarweise disjunkten und nach Voraussetzung messbaren) Urbildmengen, so gilt
$$e = \sum_{j=1}^{k} \alpha_j \mathbf{1}_{A_j} ,$$
welche die **kanonische Darstellung von** f genannt wird.[2]

[1] Hierbei bezeichnet $f^{-1}(B) := \{x \in D : f(x) \in B\}$ das Urbild von B unter f.
[2] Alle endlichen Linearkombinationen von Indikatorfunktionen von messbaren Mengen sind einfache Funktionen.

- **Theorem 3.5**: Sei $(\Omega, \mathcal{A})$ ein Messraum und f eine nichtnegative numerische $\mathcal{A}$-messbare Funktion. Dann existiert eine Folge $(e_n)_{n\in\mathbb{N}}$ nichtnegativer einfacher Funktionen auf Ω, so dass $e_k \leq e_{k+1}$ für alle $k \in \mathbb{N}$ und $\lim\limits_{n\to\infty} e_n = f$ (punktweise).

- **Definition 3.6** [**μ-fast überall (μ-f.ü.)**]: Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Wir sagen, eine Eigenschaft gelte **μ-fast überall**, falls eine μ-Nullmenge N existiert, so dass die Eigenschaft für alle $\omega \in \Omega \setminus N$ erfüllt ist.

- **Definition 3.7** [**μ-messbar**]: Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Wir sagen, eine Funktion f definiert auf $D \in \mathcal{A}$ ist **μ-messbar auf** Ω, falls $\mu(\Omega \setminus D) = 0$ und f eine $\mathcal{B}$-messbare Funktion auf D ist, wobei $\mathcal{B} = \{A \in \mathcal{A} \mid A \subset D\}$.

- **Lemma 3.8**: Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum.

  (a) Die Relation
  $$f \sim g \;:\Longleftrightarrow\; f = g \;\; \mu\text{-f.ü.}$$
  ist eine Äquivalenzrelation auf der Menge aller (numerischen) Funktionen auf Ω.

  (b) Ist f eine μ-messbare Funktion definiert auf $D \in \mathcal{A}$. Dann existiert eine $\mathcal{A}$-messbare Funktion g auf Ω, so dass $f = g$ auf D, beispielsweise
  $$g(\omega) \;=\; \begin{cases} f(\omega)\,, & \text{falls } \omega \in D\,, \\ 0\,, & \text{sonst.} \end{cases}$$

  (c) Ist $(\Omega, \mathcal{A}, \mu)$ vollständig, f eine μ-messbare Funktion und $g = f$ μ-f.ü., dann ist auch g eine μ-messbare Funktion.

- **Definition** [**Abstraktes μ-Integral**]: Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $D \in \mathcal{A}$.

  (a) Für eine nichtnegative einfache Funktion $e = \sum\limits_{j=1}^{k} \alpha_j \mathbf{1}_{A_j}$ definieren wir das μ-Integral als
  $$\int_D e(\omega) d\mu(\omega) \;:=\; \sum_{j=1}^{k} \alpha_j \mu(A_j \cap D)\,.$$

  (b) Für eine nichtnegative μ-messbare Funktion f auf D definieren wir das μ-Integral als
  $$\int_D f(\omega) d\mu(\omega) \;:=\; \sup_{\substack{0\leq e\leq f \\ e \text{ einfach}}} \int_D e(\omega) d\mu(\omega)$$
  (Hinweis: Aufgrund der μ-Messbarkeit ist die Ungleichung $0 \leq e \leq f$ im Sinne von μ-f.ü. interpretierbar.)

Zusatzaufgabe 12.1:

(a) Seien $(\Omega, \mathcal{A}, \mu)$, $(\Omega, \mathcal{B}, \nu)$ Maßräume mit $\mathcal{A} \subset \mathcal{B}$ und $\nu|_{\mathcal{A}} = \mu$.
Zeigen Sie, dass jede μ-Nullmenge auch eine ν-Nullmenge ist.

(b) Angenommen, $(\Omega, \mathcal{B}, \nu)$ sei die Vervollständigung von $(\Omega, \mathcal{A}, \mu)$. Weiter sei $(\Omega, \mathcal{A}_1, \mu_1)$ vollständig mit $\mathcal{A} \subset \mathcal{A}_1$ und $\mu = \mu_1|_{\mathcal{A}}$. Zeigen Sie, dass $\mathcal{B} \subset \mathcal{A}_1$ und $\nu = \mu_1|_{\mathcal{B}}$.

Lösung zu Zusatzaufgabe 12.1:

(a) Sei $A \subset \Omega$ eine μ-Nullmenge. Dann gilt $A \in \mathcal{A}$ und $\mu(A) = 0$. Wegen $\mathcal{A} \subset \mathcal{B}$ und $\nu|_{\mathcal{A}} = \mu$ folgt damit auch $A \in \mathcal{B}$ und $\nu(A) = \nu|_{\mathcal{A}}(A) = \mu(A) = 0$, also dass A auch eine ν-Menge ist.

(b) Nach Definition 2.5 der Vervollständigung folgt

$$\mathcal{N} = \{A \subset \Omega \mid \exists B \in \mathcal{A} : (\mu(B) = 0 \wedge A \subset B)\} \subset \mathcal{B} = \{M \cup N \mid M \in \mathcal{A} \wedge N \in \mathcal{N}\} ,$$

denn es ist $\emptyset \in \mathcal{A}$. Ebenso ist daher $\mathcal{A} \subset \mathcal{B}$. Somit gilt $\sigma(\mathcal{N} \cup \mathcal{A}) \subset \mathcal{B}$, denn $\mathcal{B}$ ist nach Satz 2.6 eine σ-Algebra. Andererseits gilt aber auch $\sigma(\mathcal{N} \cup \mathcal{A}) \supset \mathcal{B}$, denn andernfalls könnten wir einen Widerspruch zur Vollständigkeit konstruieren. Also ist $\sigma(\mathcal{N} \cup \mathcal{A}) = \mathcal{B}$.
Sei nun $(\Omega, \mathcal{A}_1, \mu_1)$ vollständig mit $\mathcal{A} \subset \mathcal{A}_1$ und $\mu = \mu_1|_{\mathcal{A}}$. Dann ist nach (a) jede μ-Nullmenge auch eine μ_1-Nullmenge, also insbesondere $\mathcal{N} \subset \mathcal{A}_1$. Wegen $\mathcal{A} \subset \mathcal{A}_1$ folgt dann auch

$$\mathcal{B} = \sigma(\mathcal{N} \cup \mathcal{A}) \subset \mathcal{A}_1 .$$

Da sich jedes Element aus $\mathcal{B}$ als (sogar disjunkte) Vereinigung einer Menge aus $\mathcal{A}$ und einer Menge aus $\mathcal{N}$ darstellen lässt, folgt für jedes $B \in \mathcal{B}$ nach Konstruktion der Vervollständigung der Isotonie von Maßen

$$\begin{aligned} B \in \mathcal{B} &\implies \exists A \in \mathcal{A}\ \exists N \in \mathcal{N} : B = N \cup A \wedge N \cap A = \emptyset \\ &\implies \nu(B) = \nu(A \cup N) = \mu(A) = \mu_1(A) = \mu_1(A) + \mu_1(N) = \mu_1(A \cup N) = \mu_1(B) \end{aligned}$$

also wie behauptet $\nu = \mu_1|_{\mathcal{B}}$.

Zusatzaufgabe 12.2:

(a) Seien $(\Omega, \mathcal{A}), (\tilde{\Omega}, \tilde{\mathcal{A}})$ zwei Messräume, so dass $\Omega = \{u, v, w, x, y, z\}$ und $\tilde{\Omega} = \{a, b, c\}$. Durch $f(u) = f(v) = f(x) = f(z) = b$ und $f(w) = f(y) = a$ sei weiter eine Funktion $f : \Omega \to \tilde{\Omega}$ gegeben.

(i) Bestimmen Sie unter f die Urbilder der Elemente von $\tilde{\mathcal{A}}$, falls $\tilde{\mathcal{A}} = \mathcal{P}(\tilde{\Omega})$.

(ii) Welches ist die kleinste σ-Algebra $\mathcal{A}$ auf Ω, so dass $\forall B \in \tilde{\mathcal{A}} : f^{-1}(B) \in \mathcal{A}$ gilt.[3]

(iii) Auf der Potenzmenge von Ω sei durch

$$\nu(\emptyset) = \nu(\{x\}) = \nu(\{w\}) = \nu(\{z\}) = 0 , \qquad \nu(\{y\}) = 1 ,\ \nu(\{u\}) = 6 ,\ \nu(\{v\}) = 35 ,$$

eine Funktion definiert. Zeigen Sie, dass $(\Omega, \mathcal{P}(\Omega), \nu)$ ein endlicher Maßraum ist.

(iv) Geben Sie die kleinste σ-Algebra $\mathcal{A} \subset \mathcal{P}(\Omega)$ an, für welche $(\Omega, \mathcal{A}, \nu|_{\mathcal{A}})$ mit ν aus (iii) vollständig wird.

(v) Geben Sie die Vervollständigung von $(\Omega, \mathcal{A}, \mu)$ mit $\mathcal{A} = \{\emptyset, \{x, w\}, \{y, z, u, v\}, \Omega\}$ und $\mu = \nu|_{\mathcal{A}}$ für ν aus (iii) an.

(b) Sei $(\Omega, \mathcal{A})$ ein beliebiger Messraum und $A \subset \Omega$ beliebig. Unter welcher Bedingung ist die **charakteristische Funktion/Indikatorfunktion** $\mathbf{1}_A : \Omega \to \mathbb{R}$ von A, definiert durch

$$\mathbf{1}_A(\omega) := \begin{cases} 1 , \text{falls } \omega \in A , \\ 0 , \text{sonst} , \end{cases}$$

messbar ?

Lösung zu Zusatzaufgabe 12.2:

(a) (i) Für die Elemente von $\mathcal{P}(\tilde{\Omega}) = \{\emptyset, \tilde{\Omega}, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}\}$ ergeben sich die Urbilder

$$\begin{aligned} f^{-1}(\emptyset) &= f^{-1}(\{c\}) = \emptyset , & f^{-1}(\{a\}) &= f^{-1}(\{a, c\}) = \{w, y\} , \\ f^{-1}(\{b\}) &= f^{-1}(\{b, c\}) = \{u, v, x, z\} , & f^{-1}(\{a, b\}) &= f^{-1}(\tilde{\Omega}) = \Omega . \end{aligned}$$

[3] Dies ist die allgemeine Definition einer $(\mathcal{A}, \tilde{\mathcal{A}})$-messbaren Funktion.

(ii) Die kleinste σ-Algebra $\mathcal{A}$ auf Ω, für die $\forall B \in \tilde{\mathcal{A}} \ : \ f^{-1}(B) \in \mathcal{A}$ gilt, ist

$$\mathcal{A} \ = \ \sigma\left(\{\emptyset, \{w,y\}, \{u,v,x,z\}, \Omega\}\right) \ = \ \{\emptyset, \{w,y\}, \{u,v,x,z\}, \Omega\}$$

(iii) Aufgrund der Endlichkeit von Ω überprüft man leicht, dass die σ-Additivität von ν, welche hier mit der Additivität von ν übereinstimmt, offenbar erfüllt ist. Insbesondere ist

$$\begin{aligned} \nu(\Omega) &= \nu\Big(\{u\}\dot{\cup}\{v\}\dot{\cup}\{w\}\dot{\cup}\{x\}\dot{\cup}\{y\}\dot{\cup}\{z\}\Big) \\ &= \nu\big(\{u\}\big) + \nu\big(\{v\}\big) + \nu\big(\{w\}\big) + \nu\big(\{x\}\big) + \nu\big(\{y\}\big) + \nu\big(\{z\}\big) \\ &= 0+1+0+6+35+0 \ = \ 42 \ < \ \infty \ . \end{aligned}$$

(iv) Die kleinste σ-Algebra $\mathcal{A} \subset \mathcal{P}(\Omega)$, für welche $(\Omega, \mathcal{A}, \nu|_{\mathcal{A}})$ vollständig wird, erhalten wir nach ZA 12.1 (b) beispielsweise als Vervollständigung des Maßraumes $(\Omega, \mathcal{B}, \mu)$ mit $\mathcal{B} = \{\emptyset, \Omega\}$ und $\mu = \nu|_{\mathcal{B}}$. Da dieser bereits vollständig ist (denn die leere Menge ist die einzige μ-Nullmenge und enthält außer sich selbst keine weiteren Nullmengen), stimmt er nach ZA 12.1 (b) mit seiner Vervollständigung überein.

(v) In diesem Fall ergibt sich die Menge $\mathcal{N} = \{\emptyset, \{x\}, \{w\}, \{x,w\}\}$, denn es ist $\{x,w\} \in \mathcal{A}$ mit $\nu|_{\mathcal{A}}(\{x,w\}) = \nu(\{x\}\dot{\cup}\{w\}) = \nu(\{x\}) + \nu(\{w\}) \ = \ 0$ die einzige nichttriviale $\nu|_{\mathcal{A}}$-Nullmenge. Somit erhalten wir

$$\begin{aligned} \overline{\mathcal{A}} &= \{M \cup N \mid M \in \mathcal{A} \wedge N \in \mathcal{N}\} \\ &= \{\emptyset, \{x\}, \{w\}, \{x,w\}, \{y,z,u,v\}, \{x,y,z,u,v\}, \{w,y,z,u,v\}, \Omega\} \end{aligned}$$

und offenbar $\overline{\mu} = \nu|_{\overline{\mathcal{A}}}$, also den nach Theorem 2.6 vollständigen Maßraum $(\Omega, \overline{\mathcal{A}}, \nu_{\overline{\mathcal{A}}})$ als Vervollständigung von $(\Omega, \mathcal{A}, \mu)$.

(b) Nach Definition 3.1 genügt es, die Frage zu beantworten, wann für beliebiges $\alpha \in \mathbb{R}$ das Urbild $\mathbf{1}_A^{-1}(]\alpha,\infty]) \ = \ \{\omega \in \Omega \mid \mathbf{1}_A(\omega) > \alpha\}$ Element der σ-Algebra ist.

- Sei $\alpha \geq 1$, dann gilt $0 \notin]\alpha,\infty]$ und ebenso $1 \notin]\alpha,\infty]$, d.h., kein $\omega \in \Omega$ wird unter $\mathbf{1}_A$ in die Menge $]\alpha,\infty] := \{x \in \mathbb{R} \cup \{-\infty,\infty\} \mid x > \alpha\}$ abgebildet, wonach das Urbild $\mathbf{1}_A^{-1}(]\alpha,\infty])$ identisch der leeren Menge ist. Die leere Menge ist jedoch Element jeder σ-Algebra, also ergibt sich in diesem Fall $\mathbf{1}_A^{-1}(]\alpha,\infty]) \in \mathcal{A}$.
- Sei $\alpha < 0$, dann gilt $\{0,1\} \subset]\alpha,\infty]$. Da jedes $\omega \in \Omega$ unter $\mathbf{1}_A$ in die Menge $\{0,1\}$ abgebildet wird, folgt in diesem Fall für das Urbild $\mathbf{1}_A^{-1}(]\alpha,\infty]) = \Omega$. Da die leere Menge Element jeder σ-Algebra ist und das Komplement einer messbaren Menge wieder messbar, folgt auch in diesem Fall $\mathbf{1}_A^{-1}(]\alpha,\infty]) \in \mathcal{A}$.
- Sei nun $\alpha \in [0,1[$, dann gilt $0 \notin]\alpha,\infty]$ und $1 \in]\alpha,\infty]$, also

  $$\mathbf{1}_A^{-1}(]\alpha,\infty]) \ = \ \mathbf{1}_A^{-1}(\{1\}) \ = \ A \ .$$

  Daher folgt die Äquivalenz

  $$\mathbf{1}_A^{-1}(]\alpha,\infty]) \in \mathcal{A} \quad \Longleftrightarrow \quad A \in \mathcal{A} \ .$$

Daher gilt nun offenbar $\forall \alpha \in \mathbb{R} \ : \ \mathbf{1}_A^{-1}(]\alpha,\infty]) \in \mathcal{A}$ genau dann, wenn $A \in \mathcal{A}$, d.h., die Indikatorfunktion $\mathbf{1}_A$ ist genau messbar, wenn es die Menge A ist.

Zusatzaufgabe 12.3:

Skizzieren Sie die Mengen $C_k \ := \ \left\{ \sum_{j\in\mathbb{N}} a_j 3^{-j} \mid a_j \in \{0,1,2\} \wedge \forall j \leq k \ : \ a_j \neq 1 \right\}$ für $k = 0,1,2$. Sind die Mengen C_k Lebesgue-messbar ? Falls ja, geben Sie $\lambda_1\left(\bigcap_{k\in\mathbb{N}_0} C_k\right)$ an.

Lösung zu Zusatzaufgabe 12.3:

Wurde in allen drei Übungsgruppen ausführlich behandelt.

Alle Übungsaufgaben sind verfügbar unter: `https://studip.uni-rostock.de/studip/`

Zusatzmaterial zur Analysis II – Serie 13

Das abstrakte μ-Integral

- **Definition 5.1 [Einfache Funktion]**: Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Eine **einfache Funktion** ist eine reellwertige $\mathcal{A}$-messbare Funktion auf Ω, deren Wertebereich endlich ist. Jede beliebige einfache Funktion lässt sich in der (nicht eindeutigen) Gestalt

$$e \;=\; \sum_{j=1}^{m} \alpha_j \mathbf{1}_{A_j} \tag{13.1}$$

mit $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ und paarweise disjunkten Mengen $A_1, \ldots, A_m \in \mathcal{A}$ schreiben.

- **Lemma 5.2**: Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Seien $A_1, \ldots, A_m \in \mathcal{A}$ und $B_1, \ldots, B_n \in \mathcal{A}$ paarweise disjunkte Mengen und $\alpha_j,\ j = 1, \ldots, m,$ und $\beta_k,\ k = 1, \ldots, n$, nichtnegative reelle Zahlen, dann gilt[1]

$$\sum_{j=1}^{m} \alpha_j \mathbf{1}_{A_j} \;\le\; \sum_{k=1}^{n} \beta_k \mathbf{1}_{B_k} \qquad \Longrightarrow \qquad \sum_{j=1}^{m} \alpha_j \mu(A_j) \;\le\; \sum_{k=1}^{n} \beta_k \mu(B_k)\ .$$

- **Definition 5.3 [Abstraktes Lebesgue-Integral]**: Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $D \in \mathcal{A}$.

  (a) Ist e eine einfache Funktion (13.1) mit paarweise disjunkten Mengen $A_1, \ldots, A_m \in \mathcal{A}$ und nichtnegativen $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$, dann definieren wir durch

$$\int_D e \, d\mu \;:=\; \sum_{j=1}^{m} \alpha_j \mu(A_j \cap D) \tag{13.2}$$

  das **μ-Integral von e**. (Das vorige Lemma garantiert uns die Wohldefiniertheit.)

  (b) Ist f eine nichtnegative numerische μ-messbare Funktion auf D, dann definieren wir durch

$$\int_D f \, d\mu \;:=\; \sup_{\substack{0 \le e \le f \\ e \text{ einfach}}} \int_D e \, d\mu \tag{13.3}$$

  das **μ-Integral von f**.

  (c) Für numerische Funktionen bezeichnet $f^+ := \max\{f, 0\}$ den **Positivteil von f** und $f^- := \max\{-f, 0\}$ den **Negativteil von f**. Ist f eine numerische $\mathcal{A}$-messbare Funktion auf D, dann definieren wir das **μ-Integral von f** durch

$$\int_D f \, d\mu \;:=\; \int_D f^+ \, d\mu \;-\; \int_D f^- \, d\mu \tag{13.4}$$

  – vorausgesetzt, mindestens eines der Integrale ist endlich.[2]

Bemerkungen:

(a) Ist f auf Ω definiert, gilt offenbar $\int_D f \, d\mu \;=\; \int_\Omega f \mathbf{1}_D \, d\mu$.

(b) Offenbar gilt $\int_D f \, d\mu \;=\; \int_D f \, d\nu$, falls $\nu = \mu|_D$ die Einschränkung von μ auf die σ-Algebra $\mathcal{A}|_D := \{A \in \mathcal{A} \;:\; A \subset D\}$ von Teilmengen von D ist.

[1] Dabei vereinbaren wir die in der Maßtheorie gebräuchliche Konvention $0 \cdot \infty = \infty \cdot 0 = 0$.
[2] Dabei vereinbaren wir die Rechenregeln $\infty - a = \infty$ und $a - \infty = -\infty$ für reelle Zahlen $a \in \mathbb{R}$.

(c) Es ist sinnvoll, das μ-Integral **μ-fast überall** einzuführen, d.h., ist $\Omega \setminus D$ eine μ-Nullmenge und f auf D definiert, dann setzen wir

$$\int_\Omega f \; d\mu \; := \; \int_D f \; d\mu \; , \tag{13.5}$$

falls die rechte Seite definiert ist. (Offenbar ist die Definition unabhängig von D).

Wir bezeichnen weiter mit $\mathcal{L}^* := \mathcal{L}^*(\mu)$ die Menge aller μ-messbaren, μ-fast überall definierten numerischen Funktionen auf Ω, für welche das μ-Integral existiert. Insbesondere bezeichnet[3]

$$\mathcal{L}^1 \; := \; \mathcal{L}^1(\mu) \; := \; \left\{ f \in \mathcal{L}^*(\mu) \;\middle|\; \int_\Omega f \; d\mu \in \mathbb{R} \right\} \; . \tag{13.6}$$

Falls $f \in \mathcal{L}^1(\mu)$, bezeichnen wir f als **μ-integrierbar**. $\mathcal{L}^1(\mu)$ ist ein linearer Raum über $\mathbb{R}$. Aus Lemma 5.2 und der Konstruktion des μ-Integrals ergibt sich, dass das μ-Integral ein lineares monotones Funktional auf $\mathcal{L}^1(\mu)$ ist.

- **Theorem 5.4 [Beppo-Levi]**: Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Sei $(f_n)_{n\in\mathbb{N}}$ eine aufsteigende Folge nichtnegativer μ-messbarer numerischer Funktionen auf Ω, welche μ-fast überall punktweise konvergiert. Für das μ-Integral des μ-fast überall definierten punktweisen Grenzwertes f gilt dann

$$\lim_{n\to\infty} \int_\Omega f_n d\mu \; = \; \int_\Omega f d\mu \; . \tag{13.7}$$

Satz von Fubini (auf $\mathbb{R}^n$), Normalbereiche

- **Satz von Fubini**: Gegeben seien die σ-endlichen Lebesgueschen Maßräume $(\mathbb{R}^k, \mathcal{M}(\mathbb{R}^k), \lambda_k)$ und $(\mathbb{R}^m, \mathcal{M}(\mathbb{R}^m), \lambda_m)$ für natürliche Zahlen $k, m \in \mathbb{N}$ und $h \in \mathcal{L}^*(\lambda_{k+m})$, dann gilt

$$\int_{\mathbb{R}^k\times\mathbb{R}^m} h \; d\lambda_{k+m} \; = \; \int_{\mathbb{R}^k} \left(\int_{\mathbb{R}^m} h(\omega, \tilde\omega) d\lambda_m(\tilde\omega) \right) d\lambda_k(\omega) \; = \; \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^k} h(\omega, \tilde\omega) d\lambda_k(\omega) \right) d\lambda_m(\tilde\omega)$$

- Eine Teilmenge $D \subset \mathbb{R}^2$ wird als **Normalbereich im $\mathbb{R}^2$** bezeichnet, falls eine der folgenden Bedingungen erfüllt ist:

  (a) $\exists \; \varphi_1, \varphi_2 : [a,b] \to \mathbb{R}$ stetig mit $D = \{(x,y) \in \mathbb{R}^2 \; : \; a \le x \le b, \; \varphi_1(x) \le y \le \varphi_2(x)\}$.

  (b) $\exists \; \psi_1, \psi_2 : [c,d] \to \mathbb{R}$ stetig mit $D = \{(x,y) \in \mathbb{R}^2 \; : \; \psi_1(y) \le x \le \psi_2(y), \; c \le y \le d\}$.

  Der **Flächeninhalt** eines Normalbereiches $D \in \mathbb{R}^2$ ist dann beispielsweise (nach Fubini)

$$\lambda_2(D) \; := \; \int_D 1 \; d\mu(x,y) \; = \; \int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} 1 \; d\lambda_1(y) \right) d\lambda_1(x) \; = \; \int_c^d \left(\int_{\psi_1(y)}^{\psi_2(y)} 1 \; d\lambda_1(x) \right) d\lambda_1(y)$$

- Eine Teilmenge $D \subset \mathbb{R}^3$ wird als **Normalbereich im $\mathbb{R}^3$** bezeichnet, falls stetige Funktionen $\varphi_1, \varphi_2 : [a,b] \to \mathbb{R}$ und $\xi_1, \xi_2 : [a,b] \times \mathbb{R} \to \mathbb{R}$ mit

$$D \; = \; \left\{ (x,y,z) \in \mathbb{R}^3 \; : \; a \le x \le b \; , \quad \varphi_1(x) \le y \le \varphi_2(x) \; , \quad \xi_1(x,y) \le z \le \xi_2(x,y) \right\}$$

  existieren (ggf. (x,y,z) permutiert). Das **Volumen** eines Normalbereiches $D \in \mathbb{R}^3$ ist

$$\lambda_3(D) \; := \; \int_D 1 \; d\lambda_3(x,y,z) \; = \; \int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} \left(\int_{\xi_1(x,y)}^{\xi_2(x,y)} 1 \; d\lambda_1(z) \right) d\lambda_1(y) \right) d\lambda_1(x) \; .$$

Zusatzaufgabe 13.1:

Ist die Funktion $f = \mathbf{1}_\mathbb{Q}$ eine Lebesgue-messbare Funktion ? Falls ja, bestimmen Sie das Lebesgue-Integral $\int_\mathbb{R} \mathbf{1}_\mathbb{Q} d\lambda_1$.

---

[3] **Anmerkung:** Oftmals finden wir anstelle von (13.6) als Definition äquivalenterweise, dass das μ-Integral der Funktion $|f|$ endlich sein muss.

Lösung zu Zusatzaufgabe 13.1:

Nach Zusatzaufgabe 12.2 (b) ist die Funktion $\mathbf{1}_{\mathbb{Q}}$ genau dann Lebesgue-messbar, wenn $\mathbb{Q}$ Lebesgue-messbar ist, also $\mathbb{Q} \in \mathcal{M}(\mathbb{R})$ gilt. Letzteres ist jedoch der Fall (denn abzählbare Mengen sind sogar Lebesguesche Nullmengen). Da es sich insbesondere um eine einfache Funktion mit der kanonischen Darstellung

$$f \; = \; 0 \cdot \mathbf{1}_{f^{-1}(\{0\})} \; + \; 1 \cdot \mathbf{1}_{f^{-1}(\{1\})} \; = \; 0\mathbf{1}_{\mathbb{R}\setminus\mathbb{Q}} \; + \; 1 \cdot \mathbf{1}_{\mathbb{Q}}$$

handelt, erhalten wir für ihr Integral

$$\int_{\mathbb{R}} \mathbf{1}_{\mathbb{Q}} d\lambda_1 \; = \; 0 \cdot \lambda_1(\mathbb{R} \setminus \mathbb{Q}) \; + \; 1 \cdot \lambda_1(\mathbb{Q}) \; = \; 0 \cdot \infty \; + \; 1 \cdot 0 \; = \; 0 \,,$$

wobei wir die in der Maßtheorie übliche Vereinbarung „ $0 \cdot \infty = 0$ “ angewandt haben.

Zusatzaufgabe 13.2: Zeigen Sie die folgenden Eigenschaften von $\mathcal{L}^1(\mu)$:

(a) $f \in \mathcal{L}^1(\mu) \quad \Longrightarrow \quad |f| \in \mathcal{L}^1(\mu)$ und es gilt $\left|\int_\Omega f \; d\mu\right| \; \leq \; \int_\Omega |f| \; d\mu$

(b) f μ-messbar $\wedge \; g \in \mathcal{L}^1(\mu) \; \wedge \; |f| \leq g \quad \Longrightarrow \quad f \in \mathcal{L}^1(\mu)$

(c) $f \in \mathcal{L}^1(\mu) \quad \Longrightarrow \quad f$ ist μ-f.ü. endlich.

Lösung zu Zusatzaufgabe 13.2:

(a) Mit $f = f^+ - f^-$ und der μ-Integrierbarkeit von f folgt die Endlichkeit der μ-Integrale von f^+ und f^-, und mit $|f| = f^+ + f^-$ sowie der Linearität somit auch die Endlichkeit des μ-Integrals $\int_\Omega |f| d\mu \; = \; \int_\Omega f^+ d\mu \; + \; \int_\Omega f^- d\mu$. Da $f \leq |f|$ und $-f \leq |f|$ gilt, folgen mit der Monotonie des μ-Integrals $\int_\Omega f \, d\mu \; \leq \; \int_\Omega |f| \, d\mu$ und $-\int_\Omega f \, d\mu \; \leq \; \int_\Omega |f| \, d\mu$, also $\left|\int_\Omega f \, d\mu\right| \; \leq \; \int_\Omega |f| \, d\mu$.

(b) Da $|f| \leq g$, also $0 \leq f^+ + f^- \leq g$ gilt, besitzen f^+ und f^- wegen $0 \leq f^+ \leq g$ bzw. $0 \leq f^- \leq g$ und der Monotonie des μ-Integrals ein endliches Integral, sind also μ-integrierbar. Aufgrund der Linearität folgt auch die μ-Integrierbarkeit von $f = f^+ - f^-$.

(c) Sei $f \in \mathcal{L}^1(\mu)$. Dann existiert nach (a) ein $c \in \mathbb{R}$ mit $\int_\Omega |f| d\mu \; = \; c \; < \; \infty$. Aufgrund der Messbarkeit von f und somit auch $|f|$ ist das Urbild $A := f^{-1}(\{-\infty, \infty\}) = |f|^{-1}(\{\infty\}) = \{\omega \in \Omega \mid |f(\omega)| = \infty\}$ messbar, also gilt $A \in \mathcal{A}$. Mit der Monotonie und (b) erhalten wir somit

$$0 \leq n \cdot \mu(A) \leq \int_\Omega n \cdot \mathbf{1}_A \; d\mu \leq \int_\Omega |f| \cdot \mathbf{1}_A \; d\mu \leq \int_\Omega |f| \; d\mu = c < \infty$$

für ein beliebiges $n \in \mathbb{N}$. Also $\mu(A) \leq \frac{c}{n}$ für jedes $n \in \mathbb{N}$. Dies ist aber nur möglich, falls $\mu(A) = 0$. Somit ist A eine μ-Nullmenge. Nach Konstruktion von A bedeutet dies aber genau, dass $|f|$ und damit auch f μ-fast überall endlich ist, wie behauptet.

Zusatzaufgabe 13.3:

(a) Bestimmen Sie mittels des Satzes von Fubini das Integral der Funktion

(i) $f(x, y) \; := \; x^2 + 2xy$ über dem in $\mathbb{R}^2$ liegende Dreieck mit Ecken $(0, 0)$, $(1, 0)$ und $(0, 2)$;

(ii) $f(x, y, z) \; := \; xy$ über der Menge

$$K \; := \; \left\{(x, y, z) \in \mathbb{R}^3 \;\middle|\; 0 \; \leq \; x \; \leq \; 2 \,, \quad |y| \; \leq \; x \,, \quad 0 \; \leq \; z \; \leq \; 8 - x^2 - y^2\right\} \,.$$

(b) Berechnen Sie das Integral $\displaystyle\int_{[0,\infty[} \left(\int_{[0,\infty[} y\mathrm{e}^{-(1+x^2)y^2} \; d\lambda_1(y)\right) \; d\lambda_1(x)$.

Zeigen Sie dann mittels des Satzes von Fubini $\displaystyle\int_{[0,\infty[} \mathrm{e}^{-x^2} \; d\lambda_1(x) \; = \; \frac{\sqrt{\pi}}{2}$.

Lösung zu Zusatzaufgabe 13.3:

(a) (i) Die Funktion f ist als stetige beschränkte Funktion über dem vorgegebenen Dreieck offensichtlich Lebesgue-integrierbar. Der Wert kann nach dem Satz von FUBINI über die iterierten Integrale berechnet werden. Dazu beachte man, dass das Dreieck als Teilmenge von $\mathbb{R}^2$ explizit die Darstellung

$$D := \Big\{(x,y) \in \mathbb{R}^2 \;\Big|\; 0 \le x \le 1,\ 0 \le y \le 2-2x\Big\}$$

hat, so dass wir insgesamt

$$\begin{aligned}\int_D f(x,y)\, d\lambda_2(x,y) &= \int_0^1 \left(\int_0^{2-2x} x^2 + 2xy\, dy\right) dx \;=\; \int_0^1 \left[x^2y + xy^2\,\right]_{y=0}^{y=2-2x}\, dy \\ &= \int_0^1 \left(2x^2 - 2x^3 + 4x + 4x^3 - 8x^2\right) dx \\ &= \int_0^1 \left(2x^3 - 6x^2 + 4x\right) dx \;=\; \left[\frac{1}{2}x^4 - 2x^3 + 2x^2\right]_{x=0}^{x=1} \;=\; \frac{1}{2}\end{aligned}$$

erhalten.

(ii) Die Lebesgue-Integrierbarkeit der Funktion f über K folgt unmittelbar aus der Tatsache, dass f eine stetige beschränkte Funktion ist und K offensichtlich endliches Maß besitzt. Mit dem Satz von FUBINI folgt

$$\begin{aligned}\int_K f(x,y,z)\, d\lambda_3(x,y,z) &= \int_0^2 \int_{-x}^{x} \int_0^{8-x^2-y^2} xy\, dz\, dy\, dz \;=\; \int_0^2 \int_{-x}^{x} xyz\Big|_{z=0}^{z=8-x^2-y^2} dy\, dx \\ &= \int_0^2 \int_{-x}^{x} 8xy - x^3y - xy^3\, dy\, dx \;=\; \int_0^2 \left[4xy^2 - \tfrac{1}{2}x^3y^2 - \tfrac{1}{4}xy^4\right]_{y=-x}^{y=x} dx \;=\; 0.\end{aligned}$$

(b) Da sowohl für beliebiges $x \ge 0$ das uneigentliche Riemann-Integral

$$\int_0^\infty y\mathrm{e}^{-(1+x^2)y^2}\, dy \;=\; \lim_{K\to\infty} \int_0^K y\mathrm{e}^{-(1+x^2)y^2}\, dy \;=\; \lim_{K\to\infty}\left(-\frac{1}{2}\frac{1}{1+x^2}\mathrm{e}^{-(1+x^2)y^2}\Big|_{y=0}^{K}\right) \;=\; \frac{1}{2}\frac{1}{1+x^2}$$

als auch das dann resultierende uneigentliche Riemann-Integral

$$\frac{1}{2}\int_0^\infty \frac{1}{1+x^2}\, dx \;=\; \frac{1}{2}\lim_{K\nearrow\infty}\int_0^K \frac{1}{1+x^2}\, dx \;=\; \frac{1}{2}\lim_{K\nearrow\infty} \arctan(x)\Big|_{x=0}^{K} \;=\; \frac{\pi}{4}$$

existiert und dabei die Integranden jeweils positiv sind, sind die einzelnen Integrale auch Lebesgue-integrierbar. Es gilt $\int_{[0,\infty)}\left(\int_{[0,\infty)} y\mathrm{e}^{-(1+x^2)y^2}\, d\lambda_1(y)\right)\, d\lambda_1(x) = \frac{\pi}{4}$. Nun ist aber nach dem Satz von TONELLI die Funktion $f : [0,\infty) \times [0,\infty) \ni (x,y) \longmapsto y\mathrm{e}^{-(1+x^2)y^2} \in \mathbb{R}$ ebenso λ_2-integrierbar und wir können die Integrationsreihenfolge nach dem Satz von FUBINI vertauschen. Mit Hilfe der Potenzgesetze und der Substitution $t = xy,\ dt = ydx$, ergibt sich

$$\begin{aligned}\frac{\pi}{4} &= \int_{[0,\infty)}\left(\int_{[0,\infty)} f(x,y)\, d\lambda_1(y)\right) d\lambda_1(x) &&= \int_{[0,\infty)}\left(\int_{[0,\infty)} f(x,y)\, d\lambda_1(x)\right) d\lambda_1(y) \\ &= \int_{[0,\infty)}\left(\int_{[0,\infty)} y\mathrm{e}^{-(1+x^2)y^2}\, d\lambda_1(x)\right) d\lambda_1(y) &&= \int_{[0,\infty)}\left(\int_{[0,\infty)} \mathrm{e}^{-x^2y^2}\, y\, dx\right)\mathrm{e}^{-y^2}\, dy \\ &= \int_{[0,\infty)}\left(\int_{[0,\infty)} \mathrm{e}^{-t^2}\, dt\right)\mathrm{e}^{-y^2}\, dy &&= \left(\int_{[0,\infty)} \mathrm{e}^{-y^2}\, dy\right)^2\end{aligned}$$

und somit nach Wurzelziehen die Behauptung.

Zusatzmaterial zur Analysis II – Serie 14

Satz von Lebesgue, Lemma von Fatou, $\mathcal{L}^p$-Räume

- **Theorem [Lebesgue – Satz von der majorisierten Konvergenz]**: Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge μ-messbarer numerischer Funktionen auf Ω, welche μ-f.ü. konvergiert. Desweiteren existiere eine Funktion $g \in \mathcal{L}^1(\mu)$, so dass
$$\forall n \in \mathbb{N} \;:\; |f_n| \;\leq\; g \quad \mu - \text{fast überall}, \tag{14.1}$$
dann liegt der μ-f.ü. definierte Grenzwert $f := \lim\limits_{n\to\infty} f_n$ bereits in $\mathcal{L}^1(\mu)$ und es gilt
$$\lim_{n\to\infty} \int_\Omega f_n d\mu \;=\; \int_\Omega f d\mu \;. \tag{14.2}$$

- **Theorem [Lebesgue für Reihen]**: Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge μ-messbarer numerischer Funktionen auf Ω. Weiter existiere eine Funktion $g \in \mathcal{L}^1(\mu)$, so dass
$$\forall m \in \mathbb{N} \;:\; \left|\sum_{n=1}^{m} f_n\right| \;\leq\; g \quad \mu\text{-fast überall}. \tag{14.3}$$
Angenommen, die Reihe $\sum\limits_{n\in\mathbb{N}} f_n$ konvergiere μ-fast überall. Dann gilt
$$\int_\Omega \sum_{n\in\mathbb{N}} f_n d\mu \;=\; \sum_{n\in\mathbb{N}} \int_\Omega f_n d\mu \;. \tag{14.4}$$

- **Theorem [Lemma von Fatou]**: Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge μ-messbarer numerischer Funktionen auf Ω und $g \in \mathcal{L}^1(\mu)$. Im Fall
$$\forall n \in \mathbb{N} \;:\; f_n \;\geq\; g \quad \mu\text{-fast überall} \tag{14.5}$$
gilt dann
$$\int_\Omega \liminf_{n\to\infty} f_n d\mu \;\leq\; \liminf_{n\to\infty} \int_\Omega f_n d\mu \;. \tag{14.6}$$

- **Theorem**: Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und f eine μ-messbare nichtnegative numerische Funktion auf Ω. Dann gilt
$$\int_\Omega f d\mu \;=\; 0 \qquad \Longrightarrow \qquad f = 0 \quad \mu\text{-fast überall}. \tag{14.7}$$

- **Theorem**: Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $f \in \mathcal{L}^1(\mu)$. Gilt $\int_E f \; d\mu = 0$ für jedes $E \in \mathcal{A}$, dann gilt μ-fast überall $f = 0$.

- **Folgerung**: Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Sind $f, g \in \mathcal{L}^1(\mu)$ und gilt $\int_E f \; d\mu \;\leq\; \int_E g \; d\mu$ für jedes $E \in \mathcal{A}$, dann gilt μ-fast überall $f \leq g$.

- **Definition [$\mathcal{L}^p$-Räume]**: Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $p \in [1, \infty]$. Wir bezeichnen mit $\mathcal{L}^p := \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$
  (a) im Fall $p < \infty$ die Menge $\left\{ f \;:\; \Omega \to [-\infty, \infty] \;\middle|\; f \;\mu\text{-messbar und } \int_\Omega |f|^p d\mu < \infty \right\}$
  (b) im Fall $p = \infty$ die Menge $\left\{ f \;:\; \Omega \to [-\infty, \infty] \;\middle|\; f \;\mu\text{-messbar und } \mu\text{-f. ü. beschränkt} \right\}$.
  Dabei heißt f eine μ-fast überall beschränkte Funktion, wenn ein $M < \infty$ existiert, so dass μ-fast überall $|f| \leq M$ erfüllt ist.

- In Verallgemeinerung bezeichnen wir für Lebesgue-messbare Mengen $M \subset \mathbb{R}^n$ mit $\mathcal{L}^*(M)$ die Menge aller λ_n-messbaren numerischen Funktionen, dessen Integral
$$\int_M f d\lambda_n \ := \ \int_{\mathbb{R}^n} f \cdot \mathbf{1}_M d\lambda_n \tag{14.8}$$
existiert (aber nicht unbedingt endlich sein muss).

- **Theorem [Satz von Fubini für Teilmengen des $\mathbb{R}^n$]**: Sei $M \subset \mathbb{R}^{m+k}$ eine Lebesgue-messbare Menge und $f \in \mathcal{L}^*(M)$. Desweiteren sei $M_x := \{y \in \mathbb{R}^k \mid (x,y) \in \mathbb{R}^{m+k}\}$. Dann ist die Funktion
$$x \mapsto g(x) \ := \ \int_{M_x} f(x,\cdot) d\lambda_k \tag{14.9}$$
für λ_m-fast alle $x \in \mathbb{R}^m$ definiert[1] und $g \in \mathcal{L}^*(\mathbb{R}^m)$ mit
$$\int_{\mathbb{R}^m} g \ d\lambda_m \ = \ \int_M f \ d\lambda_{m+k} \ . \tag{14.10}$$
Mit anderen Worten, es gilt
$$\int_M f(x,y) \ d\lambda_{m+k}(x,y) \ = \ \int_{\mathbb{R}^m} \left(\int_{M_x} f(x,y) \ d\lambda_k(y) \right) d\lambda_m(x) \ . \tag{14.11}$$
Bemerkungen:

  (a) Die Annahme $\mathcal{L}^*(M)$ ist beispielsweise erfüllt, falls $f \in \mathcal{L}^1(M)$ gilt (Satz von Tonelli) oder falls f eine messbare nichtnegative numerische Funktion ist (Satz von Fubini im engeren Sinn). Bei der Anwendung vom Satz von Fubini beginnen wir gewöhnlich mit dem Satz von Fubini im engeren Sinn, um die Integrierbarkeit von $|f|$ zu zeigen, und wenden anschließend den Satz von Tonelli an.

  (b) Ist $\pi(M)$ die Projektion von M auf den $\mathbb{R}^m$ (d.h., $\pi(M) := \{x \in \mathbb{R}^m \mid M_x \neq \emptyset\}$), dann gilt
$$\int_M f(x,y) \ d\lambda_{m+k}(x,y) \ = \ \int_{\pi(M)} \left(\int_{M_x} f(x,y) \ d\lambda_k(y) \right) d\lambda_m(x) \ , \tag{14.12}$$
  vorausgesetzt, die Menge $\pi(M)$ ist messbar.

- **Theorem [Transformationssatz für Teilmengen des $\mathbb{R}^n$]**: Seien U und V offene Teilmengen des $\mathbb{R}^n$ und $\Phi : U \to V$ ein C^1-Diffeomorphismus.[2] Weiter sei $f : V \to \mathbb{R}$ eine λ_n-messbare Funktion und $E \subset V$ eine Lebesgue-messbare Teilmenge. Mit
$$g \ := \ (f \circ \Phi)(\cdot) \left| \det d\Phi(\cdot) \right| \tag{14.13}$$
ist dann
$$f \in \mathcal{L}^*(E) \qquad \Longleftrightarrow \qquad g \in \mathcal{L}^*(\Phi^{-1}(E)) \tag{14.14}$$
und – falls eine der beiden Seiten erfüllt ist – gilt die Gleichung
$$\int_E f(y) \ d\lambda_n(y) \ = \ \int_{\Phi^{-1}(E)} (f \circ \Phi)(x) \cdot |\det d\Phi(x)| \ d\lambda_n(x) \ . \tag{14.15}$$

[1] Beachte, dass das Integral über die leere Menge stets Null ist.

[2] d.h., Φ ist bijektiv sowie Φ und Φ^{-1} sind stetig differenzierbar. Insbesondere wissen wir aus Korollar 2.58 (Skript Sommersemester 2009 bzw. 2010, Abschnitt 2.5 zu Diffeomorphismen), dass eine injektive C^1-Abbildung $\Phi : U \to \mathbb{R}^n$ mit $\det d\Phi(x) \neq 0$ für alle $x \in U$ bereits ein Diffeomorphismus von U auf $\Phi(U)$ ist.

Zusatzaufgabe 14.1: (Substitutionsregel/Transformationssatz)

(a) Seien $[a,b], [\alpha,\beta]$ beschränkte nichtdegenerierte Intervalle in $\mathbb{R}$. Zeigen Sie: Ist $t : [a,b] \to [\alpha,\beta]$ eine bijektive stetig differenzierbare Abbildung, und $f : [\alpha,\beta] \to \mathbb{R}$ eine Riemann-integrierbare Funktion, dann sind f und $f \circ t$ auch λ_1-integrierbar und es gilt

$$\int_{[a,b]} f(t(x)) \cdot |t'(x)| \, d\lambda_1(x) = \int_{[\alpha,\beta]} f(y) \, d\lambda_1(y) .$$

(b) Berechnen Sie die Fläche des Kreissektors $\Omega := \{(x,y) \in \mathbb{R}^2 \,|\, 1 < x^2 + y^2 < 4 \,,\, x,y > 0\}$, indem Sie auf die Polarkoordinaten-Abbildung $f(r,\varphi) := (r\cos(\varphi), r\sin(\varphi))$ den Transformationssatz und anschließend Satz von Fubini anwenden.

(c) Zeigen Sie, dass die Abbildung $F(x,y) := (x(1-y), xy)$ in der Nähe jedes Punktes $(x,y) \in \mathbb{R}^2$ mit $x \neq 0$ eine differenzierbare Umkehrabbildung besitzt. Beweisen Sie, dass F sogar ein Diffeomorphismus von $]0,\infty[\;\times\;]0,1[$ auf $]0,\infty[\;\times\;]0,\infty[$ ist.

(d) Berechnen Sie das Integral $\int\limits_{]0,\infty[\times]0,\infty[} \mathrm{e}^{-(u+v)^2} \, d\lambda_2(u,v)$, indem Sie den Transformationssatz auf F aus (c) anwenden.

Lösung zu Zusatzaufgabe 14.1:

(a) Da f R-integrierbar, ist f auch λ_1-integrierbar auf $[\alpha,\beta]$ mit $\int\limits_{[\alpha,\beta]} f(y) \, d\lambda_1(y) = \int\limits_{\alpha}^{\beta} f(y) \, dy$.

Da t bijektiv, muss $t([a,b]) = [\alpha,\beta]$ gelten, so dass wegen der (aus der Injektivität und Stetigkeit folgenden) strengen Monotonie die folgenden Fälle unterschieden werden müssen:

- Es gilt $t(a) = \beta$ und $t(b) = \alpha$ und $t' \leq 0$ (streng monoton fallend).
- Es gilt $t(a) = \alpha$ und $t(b) = \beta$ und $t' \geq 0$ (streng monoton wachsend).

Nach der Substitutionsregel aus der Analysis I (vgl. Forster I, Satz 4, §19) folgt im ersten Fall

$$\int_{[a,b]} (f \circ t)(x) \cdot |t'(x)| \, d\lambda_1(x) = -\int_a^b f(t(x)) \cdot t'(x) dx = -\int_\beta^\alpha f(y) \, dy = \int_\alpha^\beta f(y) \, dy = \int_{[\alpha,\beta]} f(y) \, d\lambda_1(y)$$

und im zweiten Fall

$$\int_{[a,b]} (f \circ t)(x) \cdot |t'(x)| \, d\lambda_1(x) = \int_a^b f(t(x)) \cdot t'(x) dx = \int_\alpha^\beta f(y) \, dy = \int_{[\alpha,\beta]} f(y) \, d\lambda_1(y) .$$

(b) Zunächst gilt $df(r,\varphi) = \begin{pmatrix} \cos(\varphi) & -r\sin(\varphi) \\ \sin(\varphi) & r\cos(\varphi) \end{pmatrix}$, also $\det(df(r,\varphi)) = r$. Mit dem Transformationssatz und Polarkoordinaten sowie anschließender Anwendung vom Satz von Fubini gilt

$$\lambda_2(\Omega) = \int_\Omega 1 \, d\lambda_2(x,y) = \int_{]1,2[\times]0,\frac{\pi}{2}[} r d\lambda_2(r,\varphi) = \int_1^2 \left(\int_0^{\frac{\pi}{2}} r \, d\varphi \right) dr = \frac{\pi}{2} \int_1^2 r \, dr = \frac{\pi}{4} r^2 \Big|_{r=1}^{r=2} = \frac{3\pi}{4} .$$

(c) Es gilt $dF(x,y) = \begin{pmatrix} (1-y) & -x \\ y & x \end{pmatrix}$, und daher ist wegen $\det \begin{pmatrix} (1-y) & -x \\ y & x \end{pmatrix} = x$ die Ableitung $dF(x,y)$ in jedem Punkt $(x,y) \in \mathbb{R}^2$ mit $x \neq 0$ invertierbar. Nach dem Satz über lokale Umkehrbarkeit besitzt F in der Nähe solcher Punkte eine differenzierbare Umkehrabbildung.

Zur globalen Diffeomorphie: Zunächst einmal bildet F den Streifen $(0,\infty) \times (0,1)$ wirklich nach $(0,\infty) \times (0,\infty)$ ab, denn $x(1-y)$ und xy sind positiv für $x \in (0,\infty)$, $y \in (0,1)$.

Für $F(x,y) = (u,v)$ gilt $x = u + v$ sowie $y = \frac{v}{u+v}$, wobei y bei $u,v > 0$ auch wirklich wohldefiniert ist und in $(0,1)$ liegt. Also ist $F : (0,\infty) \times (0,1) \to (0,\infty) \times (0,\infty)$ bijektiv, die Umkehrabbildung ist explizit durch $F^{-1}(u,v) = (u+v, \frac{v}{u+v})$ gegeben. Bijektivität und lokale Diffeomorphie implizieren aber globale Diffeomorphie.

(d) Mit $g(u,v) := \mathrm{e}^{-(u+v)^2}$ und $(g \circ F)(x,y) = \mathrm{e}^{-x^2}$ gilt wegen $\det(dF(x,y)) = x$ nach dem Transformationssatz (aufgrund der Nichtnegativität existiert das (ggf. unendliche) λ_2-Integral in jedem Fall) und Satz von Fubini

$$\int\limits_{]0,\infty[\,\times\,]0,\infty[} \mathrm{e}^{-(u+v)^2}\, d\lambda_2(u,v) \;=\; \int\limits_{]0,\infty[\,\times\,]0,1[} \mathrm{e}^{-x^2} x\, d\lambda_2(x,y) \;=\; \int\limits_{]0,1[}\Big(\int\limits_{]0,\infty[} \mathrm{e}^{-x^2} x\, d\lambda_1(x)\Big)\, d\lambda_1(y)$$

Mit $z = x^2,\ x\,dx = \frac{1}{2}\,dz$ erhalten wir $\displaystyle\int \mathrm{e}^{-x^2} x\,dx \;=\; \frac{1}{2}\int \mathrm{e}^{-z}\,dz \;=\; -\frac{1}{2}\mathrm{e}^{-z}$ und somit

$$\int\limits_{]0,\infty[\,\times\,]0,\infty[} \mathrm{e}^{-(u+v)^2}\, d\lambda_2(u,v) \;=\; \int\limits_{]0,\infty[} \mathrm{e}^{-x^2} x\, d\lambda_1(x) \;=\; \lim_{R\to\infty}\int_0^R \frac{1}{2}\mathrm{e}^{-z}\,dz \;=\; \frac{1}{2}\,.$$

Zusatzaufgabe 14.2:

(a) Zeigen Sie, dass die Abbildung $\Psi(\xi,r) := \Big(r\xi, r\sqrt{1-\xi^2}\Big)$ in der Nähe jedes Punktes $(\xi,r) \in \mathbb{R}^2$ mit $|\xi| < 1$ und $r \neq 0$ eine differenzierbare Umkehrabbildung besitzt. Beweisen Sie, dass Ψ ein Diffeomorphismus vom Halbstreifen $(-1,1)\times(0,\infty)$ auf $\mathbb{R}\times(0,\infty)$ ist.

(b) Berechnen Sie das Integral
$$\int_{(-1,1)\times(0,\infty)} \frac{1}{(1+r^2\xi^2)\cdot(1+r\sqrt{1-\xi^2})^2}\cdot\frac{r}{\sqrt{1-\xi^2}}\,d\lambda_2(\xi,r)\ ,$$
indem Sie den Transformationssatz auf Ψ und anschließend den Satz von Fubini anwenden.

(c) Zeigen Sie, dass $\mathcal{M}(\mathbb{R}^2) \neq \mathcal{M}(\mathbb{R})\times\mathcal{M}(\mathbb{R})$ gilt.

Lösung zu Zusatzaufgabe 14.2:

(a) Die Jacobi-Matrix $J\Psi(\xi,r) = \begin{pmatrix} r & \xi \\ -\frac{r\xi}{\sqrt{1-\xi^2}} & \sqrt{1-\xi^2}\end{pmatrix}$ besitzt die Determinante

$$\det(J\Psi(\xi,r)) \;=\; r\sqrt{1-\xi^2} + r\frac{\xi^2}{\sqrt{1-\xi^2}} \;=\; \frac{r}{\sqrt{1-\xi^2}}\big(1-\xi^2+\xi^2\big) \;=\; \frac{r}{\sqrt{1-\xi^2}}\ ,$$

also ist Ψ in der Nähe jedes Punktes (ξ,r) mit $|\xi|<1$ und $r\neq 0$ ein lokaler Diffeomorphismus. Tatsächlich ist Ψ sogar ein Diffeomorphismus von $(-1,1)\times(0,\infty)$ auf $\mathbb{R}\times(0,\infty)$, denn $r\xi = u$ und $r\sqrt{1-\xi^2} = v$ können wir bei $v>0$, $|\xi|<1$ und $r>0$ wegen $r^2 = u^2+v^2$ auflösen durch

$$r = \sqrt{u^2+v^2} \qquad\text{und}\qquad \xi = \frac{u}{\sqrt{u^2+v^2}}\,.$$

(b) Aus dem Transformationssatz folgt mit $f(u,v) = \frac{1}{1+u^2}\frac{1}{(1+v)^2}$

$$\begin{aligned}
&\int_{(-1,1)\times(0,\infty)} \frac{1}{(1+r^2\xi^2)\cdot(1+r\sqrt{1-\xi^2})^2}\cdot\frac{r}{\sqrt{1-\xi^2}}\,d\lambda_2(\xi,r)\\
&= \int_{(-1,1)\times(0,\infty)} (f\circ\Psi)(\xi,r)\cdot|\det(J\Psi)(\xi,r)|\,d\lambda_2(\xi,r) \;=\; \int_{\mathbb{R}\times(0,\infty)} f(u,v)d\lambda_2(u,v)\\
&= \left(\int_{\mathbb{R}}\frac{1}{1+u^2}d\lambda_1(u)\right)\left(\int_{(0,\infty)}\frac{1}{(1+v)^2}d\lambda_1(v)\right) \;=\; \left(\arctan(u)\Big|_{-\infty}^{\infty}\right)\left(-\frac{1}{1+v}\Big|_0^{\infty}\right) = \pi\,.
\end{aligned}$$

(c) Offenbar ist $\mathcal{M}(\mathbb{R}^2) \supset \mathcal{M}(\mathbb{R})\times\mathcal{M}(\mathbb{R})$. Aufgrund der Vollständigkeit von $(\mathbb{R}^n,\mathcal{M}(\mathbb{R}^n),\lambda_n)$ muss jede Teilmenge der λ_2-Nullmenge $N := \{0\}\times\mathbb{R}$ ebenfalls in $\mathcal{M}(\mathbb{R}^2)$ liegen, somit auch $M := \{0\}\times A$ mit der nach ZA 11.4 existenten Teilmenge $A \in \mathcal{P}(\mathbb{R})\setminus\mathcal{M}(\mathbb{R})$. Jedoch ist dann $M \notin \mathcal{M}(\mathbb{R})\times\mathcal{M}(\mathbb{R})$.