Lösungen der AmV-Aufgaben 2007/08

1. (a) Es ist

$$f'(x) = x^{q-1} - a\,,$$

und da $q > 1$ nach Konstruktion, ist $q-1 > 0$, d. h. f' hat eine eindeutige Nullstelle bei $x_0 := a^{1/(q-1)}$. Da $f''(x) = (q-1)x^{q-2} > 0$ für alle $x \in (0,\infty)$, ist x_0 tatsächlich ein Minimum von f. Es gilt

$$f(x_0) = \frac{a^p}{p} + \frac{a^{q/(q-1)}}{q} - aa^{1/(q-1)} = \frac{a^p}{p} + \frac{a^p}{q} - a^p = 0\,.$$

(b) Für jedes $b > 0$ gilt also mit (a)

$$0 = f(x_0) \le f(b) = \frac{a^p}{p} + \frac{b^q}{q} - ab \quad\Longrightarrow\quad ab \le \frac{a^p}{p} + \frac{b^q}{q}\,.$$

Für $a = 0$ oder $b = 0$ ist die Ungleichung trivialerweise erfüllt.

2. (a) Aus $|x|_p = 1$ folgt $|x_k| \le 1$ für alle $k \in \{1,\dots,n\}$, also $|x_k|^q \le |x_k|^p$, da $q > p$. Damit folgt $\sum_{k=1}^n |x_k|^q \le \sum_{k=1}^n |x_k|^p = 1$ und daraus $|x|_q \le 1$. Für einen beliebigen Vektor $x \in \mathbb{R}^n \setminus \{0\}$ (Für den Nullvektor ist nichts zu zeigen.) folgt

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

Die Hölder-Ungleichung besagt für $1/\alpha + 1/\beta = 1$

$$\sum_{k=1}^n \big||x_k|^p \cdot 1\big| \le \left(\sum_{k=1}^n \big||x_k|^p\big|^\alpha \right)^{1/\alpha} \left(\sum_{k=1}^n |1|^\beta \right)^{1/\beta}.$$

Ziehen der p-ten Wurzel ergibt

$$|x|_p \le n^{1/p\beta} |x|_{p\alpha}\,.$$

Nun wählen wir $\alpha = q/p > 1$, woraus sich $\beta = 1/(1-p/q) = q/(q-p)$ und damit die Behauptung ergibt.

(b) Wegen $|x|_\infty = 1$ gilt $|x_i| \le 1$ für alle $i \in \{1,\dots,n\}$, und es gibt (mindestens) einen Index $k \in \{1,\dots,n\}$ mit $|x_k| = 1$. Dann ist

$$|x|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p} = \left(1 + \sum_{i=1,\, i\neq k}^n |x_i|^p \right)^{1/p} \begin{cases} \ge 1\,, \\ \le n^{1/p}\,. \end{cases}$$

Wegen $n^{1/p} \to 1$ für $p \to \infty$ folgt mit dem Quetschlemma $|x|_p \to 1$ für $p \to \infty$. Für einen beliebigen Vektor $x \in \mathbb{R}^n \setminus \{0\}$ (Für den Nullvektor ist nichts zu zeigen.) folgt wie in (a)

$$|x|_p = \left| |x|_\infty \frac{x}{|x|_\infty} \right|_p = |x|_\infty \left| \frac{x}{|x|_\infty} \right|_p \xrightarrow{p\to\infty} |x|_\infty \left| \frac{x}{|x|_\infty} \right|_\infty = |x|_\infty\,.$$

3. Seien grundsätzlich $f, g \in \mathcal{C}(I)$ und $1 < p, q < \infty$ mit $1/p + 1/q = 1$.

   (a) Sei zunächst $\|f\|_p = \|g\|_q = 1$. Dann folgt mit der Youngschen Ungleichung aus Aufgabe 1. (b)

$$\int_a^b f(t)g(t)\,\mathrm{d}t \le \int_a^b \big|f(t)\big|\,\big|g(t)\big|\,\mathrm{d}t \le \int_a^b \Big(\frac{1}{p}\big|f(t)\big|^p + \frac{1}{q}\big|g(t)\big|^q\Big)\,\mathrm{d}t$$
$$= \frac{1}{p}\|f\|_p^p + \frac{1}{q}\|g\|_q^q = \frac{1}{p} + \frac{1}{q} = 1\,.$$

   Seien nun f, g beliebig mit $\|f\|_p, \|g\|_q > 0$. (Für $f \equiv 0$ oder $g \equiv 0$ ist nichts zu zeigen.) Dann folgt

$$\int_a^b f(t)g(t)\,\mathrm{d}t = \|f\|_p\|g\|_q \int_a^b \frac{f(t)}{\|f\|_p}\,\frac{g(t)}{\|g\|_q}\,\mathrm{d}t \le \|f\|_p\|g\|_q\,.$$

   (b) Mit $h(t) := |f(t) + g(t)|^{p-1}$ folgt mit der Hölderschen Ungleichung

$$\|f+g\|_p^p = \int_a^b \big|f(t)+g(t)\big|^p\,\mathrm{d}t = \int_a^b \big|f(t)+g(t)\big|h(t)\,\mathrm{d}t$$
$$\le \int_a^b \big|f(t)\big|h(t)\,\mathrm{d}t + \int_a^b \big|g(t)\big|h(t)\,\mathrm{d}t \le \|f\|_p\|h\|_q + \|g\|_p\|h\|_q\,.$$

   Weiterhin gilt

$$\|h\|_q^q = \int_a^b \big|h(t)\big|^q\,\mathrm{d}t = \int_a^b \big|f(t)+g(t)\big|^{(p-1)q}\,\mathrm{d}t = \int_a^b \big|f(t)+g(t)\big|^p\,\mathrm{d}t = \|f+g\|_p^p\,,$$

   woraus $\|h\|_q = \|f+g\|_p^{p/q} = \|f+g\|_p^{p-1}$ und schließlich

$$\|f+g\|_p^p \le \|f\|_p\|f+g\|_p^{p-1} + \|g\|_p\|f+g\|_p^{p-1}$$

   folgt. Division durch $\|f+g\|_p^{p-1} \ne 0$ liefert die Behauptung. (Für $f + g \equiv 0$ ist nichts zu zeigen.)

4. (a) Positivität, Definitheit und Symmetrie sind klar. Für die Dreiecksungleichung unterscheiden wir vier Fälle:

   i. $x, y, z \in \mathbb{R}^2$ paarweise linear abhängig:

$$d(x,z) = |x-z|_2 = |x-y+y-z|_2 \le |x-y|_2 + |y-z|_2 = d(x,y) + d(y,z)\,.$$

   ii. $x, y \in \mathbb{R}^2$ linear abhängig, $z \in \mathbb{R}^2$ zu diesen linear unabhängig:

$$d(x,z) = |x|_2+|z|_2 = |x-y+y|_2+|z|_2 = |x-y|_2+|y|_2+|z|_2 = d(x,y)+d(y,z)\,.$$

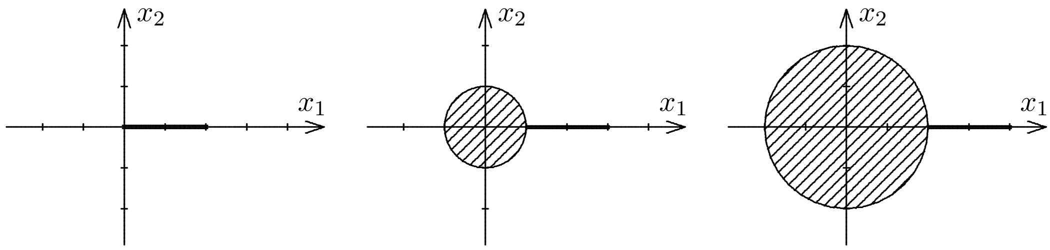
   iii. $x, z \in \mathbb{R}^2$ linear abhängig, $y \in \mathbb{R}^2$ zu diesen linear unabhängig:

$$d(x,z) = |x-z|_2 \le |x|_2 + |z|_2 + |y|_2 + |y|_2 = d(x,y) + d(y,z)\,.$$

   iv. $x, y, z \in \mathbb{R}^2$ paarweise linear unabhängig:

$$d(x,z) = |x|_2 + |z|_2 \le |x|_2 + |y|_2 + |y|_2 + |z|_2 = d(x,y) + d(y,z)\,.$$

(b) Diagramme für $B_1((1,0)^{\mathrm{T}})$, $B_2((1,0)^{\mathrm{T}})$ und $B_3((1,0)^{\mathrm{T}})$:



5. (a) Seien $x, y \in B_r(a)$. Dann ist

$$d(x,y) \leq d(x,a) + d(a,y) \leq 2r\,.$$

Übergang zum Supremum liefert $\operatorname{diam}(B_r(a)) \leq 2r$.
Die Ungleichung ist strikt für $d(x,y) = 1$ für $x \neq y$, denn es ist $\operatorname{diam}(B_0(a)) = 0$ und $\operatorname{diam}(B_r(a)) = 1$ für $r > 0$, falls $|X| > 1$. Die Ungleichung ist also strikt, falls $r \neq 0$ und $r \neq 1/2$. (Der pathologische Fall $|X| = 1$ tut es natürlich auch immer.)
Für einen normierten Raum $(X, \|\cdot\|)$ ist

$$\begin{aligned}\operatorname{diam}(B_r(a)) = \sup\bigl\{\|x-y\| \bigm| x,y \in B_r(a)\bigr\} &\overset{\text{i.}}{=} \sup\bigl\{\|x-y\| \bigm| x,y \in B_r(0)\bigr\} \\ &\overset{\text{ii.}}{\geq} \sup\bigl\{\|x+x\| \bigm| x \in B_r(0)\bigr\} = 2\sup\bigl\{\|x\| \bigm| x \in B_r(0)\bigr\} = 2r\,.\end{aligned}$$

Zusammen mit „$\leq$“ ergibt sich Gleichheit.
Bei i. wurden $\|x-y\| = \|(x-a)-(y-a)\|$ und $x \in B_r(a) \iff x - a \in B_r(0)$ benutzt.
Bei ii. wurde benutzt, dass mit $x \in B_r(0)$ auch $-x \in B_r(0)$ gilt und dass das Supremum höchstens kleiner wird, wenn man die Menge verkleinert.

(b) Wenn A beschränkt ist, dann existiert ein $R > 0$ mit $\operatorname{diam} A \leq R$. Damit gilt nach (a) $A \subset B_{2R}(a)$ für jedes $a \in A$.
Es gelte nun $A \subset B_r(a)$ für ein $r > 0$ und $a \in X$. Wegen $\operatorname{diam} A \leq \operatorname{diam}(B_r(a)) \leq 2r$ nach (a) ist dann A beschränkt.

(c) Dass A offen ist, ist äquivalent dazu, dass es für jedes $x \in A$ ein $\varepsilon_x > 0$ gibt, so dass $U_{\varepsilon_x}(x) \subset A$. Dies ist gleichbedeutend zu $U_{\varepsilon_x}(x) \cap (X \setminus A) = \emptyset$, d. h. zu $\operatorname{dist}(x, X \setminus A) \geq \varepsilon_x$. Diese Äquivalenz beweist die Behauptung.

(d) Zu $\bar{A}$: „$\subset$“: Sei $x \in X$ mit $\operatorname{dist}(x,A) = \varepsilon > 0$. Dann ist $U_\varepsilon(x) \subset X \setminus A$, und für jede abgeschlossene Menge $B \subset X$ mit $A \subset B$ ist $B' = B \setminus U_\varepsilon(x)$ abgeschlossen mit $A \subset B'$. Also folgt $x \notin \bar{A}$.
„$\supset$“: Sei $x \notin \bar{A}$. Dann gibt es eine abgeschlossene Menge $B \subset X$ mit $B \subset A$ und $x \notin B$. Dann ist aber $X \setminus B$ offen und $x \in X \setminus B$, d. h. es gibt ein $\varepsilon > 0$, so dass $U_\varepsilon(x) \subset X \setminus B$. Also folgt $\operatorname{dist}(x,A) \geq \varepsilon$.
Zu ∂A: „$\subset$“: Sei $x \in \partial A$, d. h. $x \in \bar{A}$ und $x \notin A^\circ$. Dann folgt nach der ersten Aussage von (d) $\operatorname{dist}(x,A) = 0$. Würde jedoch $\operatorname{dist}(x, X \setminus A) = \varepsilon > 0$ gelten, so wäre $U_\varepsilon(x) \subset A$, also $x \in A^\circ$. Daher muss auch $\operatorname{dist}(x, X \setminus A) = 0$ gelten.
„$\supset$“: Sei nun $x \in X$ mit $\operatorname{dist}(x,A) = \operatorname{dist}(x, X \setminus A) = 0$. Dann folgt nach der ersten Aussage von (d) $x \in \bar{A}$. Würde jedoch $x \in A^\circ$ gelten, so gäbe es ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subset A$, also $\operatorname{dist}(x, X \setminus A) \geq \varepsilon$. Daher muss $x \notin A^\circ$ und damit $x \in \partial A$ gelten.

6. (a) A ist nicht offen, denn für den Punkt $a := (2,0)^{\mathrm{T}} \in A$ gibt es kein $\varepsilon > 0$, so dass $U_\varepsilon(a) \subset A$, denn in $U_\varepsilon(a)$ liegt z. B. $(2, \varepsilon/2)^{\mathrm{T}} \notin A$.
   A ist nicht abgeschlossen, denn für den Punkt $a := (1,0)^{\mathrm{T}} \in \mathbb{R}^2 \setminus A$ gibt es kein $\varepsilon > 0$, so dass $U_\varepsilon(a) \subset \mathbb{R}^2 \setminus A$, denn in $U_\varepsilon(a)$ liegt z. B. $(1+\varepsilon/2, 0)^{\mathrm{T}} \notin \mathbb{R}^2 \setminus A$.
   Es ist $A^\circ = \emptyset$, denn $\emptyset$ ist die einzige offene Menge, die in A liegt.
   Es ist $\bar{A} = [1,3] \times \{0\}$, denn diese Menge ist abgeschlossen und gleichzeitig die kleinste, die A enthält.
   Damit folgt $\partial A = \bar{A} \setminus A^\circ = [1,3] \times \{0\}$.
   B ist offen, dann für den Punkt $b := (x,y)^{\mathrm{T}} \in B$ erfüllt $U_\varepsilon(b) \subset B$, falls man $\varepsilon = \min\{1-|x|, 1-|y|, |x|\}/2$ für $x \neq 0$ bzw. $\varepsilon = \min\{1-|x|, 1-|y|, |y|\}/2$ für $x = 0$ wählt.
   B ist nicht abgeschlossen, denn für den Punkt $b := (1,0)^{\mathrm{T}} \in \mathbb{R}^2 \setminus B$ gibt es kein $\varepsilon > 0$, so dass $U_\varepsilon(b) \subset \mathbb{R}^2 \setminus B$, denn in $U_\varepsilon(b)$ liegt z. B. $(1-\varepsilon/2, 0)^{\mathrm{T}} \notin \mathbb{R}^2 \setminus B$.
   Es ist $B^\circ = B$, da B selbst offen ist.
   Es ist $\bar{B} = [-1,1]^2$, denn diese Menge ist abgeschlossen und gleichzeitig die kleinste, die B enthält.
   Damit folgt $\partial B = \bar{B} \setminus B^\circ = ([-1,1] \times \{-1,1\}) \cup (\{-1,1\} \times [-1,1]) \cup (\{0\} \times [0,1])$.
   Bemerkung: Alle Ergebnisse bleiben unverändert, wenn d_∞ durch d_1 oder d_2 ersetzt wird. Das liegt daran, dass d_1, d_2 und d_∞ zueinander äquivalente Metriken sind, was durch die beiden folgenden gleichwertigen Eigenschaften definiert ist: (a) Für alle $p, q \in \{1, 2, \infty\}$ gibt es Konstanten $C_{pq}, D_{pq} > 0$ mit $C_{pq} d_p(x,y) \leq d_q(x,y) \leq D_{pq} d_p(x,y)$ für alle $x, y \in X$. (b) Für alle $p, q \in \{1, 2, \infty\}$ gilt $d_p(x_n, a) \to 0 \iff d_q(x_n, a) \to 0$ für alle Folgen $(x_n)_n \subset X$ und $a \in X$.

   (b) Für $a = (a_1, 0)^{\mathrm{T}} \in A$ definieren wir $\varepsilon = (1 - |a_1 - 2|)/2$ mit $0 < \varepsilon \leq 1/2$. Dann ist bezüglich der FEB-Metrik $U_\varepsilon(a) = (a_1 - \varepsilon, a_1 + \varepsilon) \times \{0\}$, denn für jeden Punkt $b \in \mathbb{R}^2$ mit $b_2 \neq 0$ ist $d(a,b) > 1$. Nach Konstruktion von ε gilt aber $U_\varepsilon(a) \subset A$, d. h. A ist offen.
   Bemerkung: Offenbar sind d_∞ und d_{FEB} keine äquivalenten Metriken, denn A ist offen bezüglich d_{FEB}, aber nicht bezüglich d_∞. Auch die Grenzwerte sind verschieden: Für $a_n = (1, 1/n)^{\mathrm{T}}$ und $a = (1,0)^{\mathrm{T}}$ gilt offenbar $d_\infty(a_n, a) \to 0$ für $n \to \infty$, aber $d_{FEB}(a_n, a) \to 2$.

7. Für jedes $n \in \mathbb{N}$ definiere $f_n \in \mathcal{C}(I)$ durch

$$f_n(x) := \begin{cases} 1 - nx & \text{für } x \leq 1/n\,, \\ 0 & \text{für } x > 1/n\,. \end{cases}$$

Für $1 \leq p < \infty$ ist dann

$$\|f_n\|_p^p = \int_0^1 \big|f_n(x)\big|^p \,\mathrm{d}x = \int_0^{1/n} (1-nx)^p \,\mathrm{d}x = -\frac{(1-nx)^{p+1}}{n(p+1)}\Bigg|_{x=0}^{x=1/n} = \frac{1}{n(p+1)}\,,$$

d. h. $\|f_n\|_p = (n(p+1))^{-1/p}$. Ferner ist

$$\|f_n\|_\infty = \sup\Big\{ \big|f_n(x)\big| \;\Big|\; x \in I \Big\} = 1\,.$$

Bemerkung: Man beachte, dass tatsächlich $\|f_n\|_p \to \|f_n\|_\infty$ für $p \to \infty$.

Wären nun $\|\cdot\|_p$ und $\|\cdot\|_q$ für $1 \leq p < q \leq \infty$ äquivalent, so gäbe es Konstanten $C_{pq} > 0$, die nicht von n abhängen, so dass

$$\|f_n\|_q \leq C_{pq} \|f_n\|_p\,.$$

Für $q < \infty$ wächst jedoch der Quotient

$$\frac{\|f_n\|_q}{\|f_n\|_p} = \frac{(n(q+1))^{-1/q}}{(n(p+1))^{-1/p}} = \frac{\sqrt[p]{p+1}}{\sqrt[q]{q+1}}\, n^{1/p-1/q}$$

wegen $1/p > 1/q$ für $n \to \infty$ über alle Schranken, d. h. es kann kein solches C_{pq} geben. Für $q = \infty$ ergibt sich analog

$$\frac{\|f_n\|_q}{\|f_n\|_p} = \frac{1}{(n(p+1))^{-1/p}} = \sqrt[p]{p+1}\, n^{1/p}$$

mit demselben Ergebnis, da $1/p > 0$.

Also ist kein Paar $(\|\cdot\|_p, \|\cdot\|_q)$ von Normen äquivalent.

8. (a) Für alle $x, y \in X$ gilt

$$\begin{aligned} d(x,y) + \operatorname{dist}(y,A) &= d(x,y) + \inf\bigl\{d(y,a) \bigm| a \in A\bigr\} \\ &= \inf\bigl\{d(x,y) + d(y,a) \bigm| a \in A\bigr\} \geq \inf\bigl\{d(x,a) \bigm| a \in A\bigr\} = \operatorname{dist}(x,A)\,, \end{aligned}$$

und damit $\operatorname{dist}(x,A) - \operatorname{dist}(y,A) \leq d(x,y)$. Durch Vertauschen von x und y erhält man die Ungleichung $\operatorname{dist}(y,A) - \operatorname{dist}(a,A) \leq d(x,y)$, d. h. es folgt

$$\bigl|\operatorname{dist}(x,A) - \operatorname{dist}(y,A)\bigr| \leq d(x,y)\,.$$

(b) Für jeden normierten Raum $(V, \|\cdot\|)$ ist $d(x,y) := \|x-y\|$ eine Metrik. Wähle $A = \{0\} \subset V$. Dann ist offensichtlich $\operatorname{dist}(x,A) = d(x,0) = \|x\|$ (und $\operatorname{dist}(y,A) = \|y\|$), und mit (a) folgt

$$\bigl|\|x\| - \|y\|\bigr| = \bigl|\operatorname{dist}(x,A) - \operatorname{dist}(y,A)\bigr| \leq d(x,y) = \|x-y\|\,.$$

9. (a) Es ist

$$f(tx) = \frac{tx_1 tx_2}{tx_1 + tx_2} = \frac{x_1 x_2}{x_1 + x_2}\, t \xrightarrow{t \to 0} 0\,,$$

d. h. $f(tx) \to 0$ für $t \to 0$.

(b) Idee: Wähle einen Weg zum Punkt $(0,0)^{\mathrm{T}}$ derart, dass er tangential zur zweiten Winkelhalbierenden „einbiegt“, also z. B. die Folge $x_n = (n^{-1}, n^{-2} - n^{-1})^{\mathrm{T}}$. Wegen $x_{n,1} + x_{n,2} = n^{-2} > 0$ liegt x_n tatsächlich in G; ferner gilt $x_n \to (0,0)^{\mathrm{T}}$ für $n \to \infty$. Aber

$$\bigl|f(x_n) - f(0)\bigr| = \left|\frac{n^{-1}(n^{-2} - n^{-1})}{n^{-1} + n^{-2} - n^{-1}}\right| = \left|\frac{1}{n} - 1\right| \xrightarrow{n \to \infty} 1\,,$$

d. h. f ist in $(0,0)^{\mathrm{T}}$ nicht stetig.

10. (a) „(i) $\Longrightarrow$ (ii)“: Zu zeigen ist, dass für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass $f(U_\delta(a)) \subset U_\varepsilon(f(a))$, d. h. aus $x \in U_\delta(a)$ folgt $f(x) \in U_\varepsilon(f(a))$, oder, noch anders geschrieben, aus $d_X(x,a) < \delta$ folgt $d_Y(f(x), f(a)) < \varepsilon$. Nehmen wir an, dies sei falsch, dann gäbe es ein $\varepsilon > 0$ und eine Folge $(a_n)_n \subset X$ mit $a_n \to a$ für $n \to \infty$ bezüglich d_X, so dass: Zu jedem $k \in \mathbb{N}$ gibt es ein $n_k \in \mathbb{N}$ mit $d_X(a_{n_k}, a) < 1/k$ und $d_Y(f(a_{n_k}), f(a)) \geq \varepsilon$. Nach (i) müsste aber $f(a_{n_k}) \to f(a)$ für $k \to \infty$ bezüglich d_Y folgen, also ein Widerspruch zur letzten Ungleichung.

„(ii) $\Longrightarrow$ (iii)": Sei V eine Umgebung von $f(a)$. Dann ist $f(a)$ ein innerer Punkt von V, d. h. es gibt eine offene Kugel um $f(a)$ in V, z. B. $U_\varepsilon(f(a))$ mit $\varepsilon = \operatorname{dist}(f(a), \partial V)/2 > 0$. Nach (ii) gibt es ein $\delta > 0$ mit $f(U_\delta(a)) \subset U_\varepsilon(f(a)) \subset V$. Anwenden von f^{-1} liefert $U_\delta(a) \subset f^{-1}(f(U_\delta(a))) \subset f^{-1}(V)$. Also ist $f^{-1}(V)$ eine Umgebung von a, da die offene Kugel $U_\delta(a)$ ganz drin liegt.

„(iii) $\Longrightarrow$ (i)": Sei $(a_n)_n$ eine Folge in X mit $a_n \to a$ für $n \to \infty$ bezüglich d_X. Wir müssen zeigen, dass dann $f(a_n) \to f(a)$ für $n \to \infty$ bezüglich d_Y. Sei V eine beliebige Umgebung von $f(a)$. Da nach (iii) dann $f^{-1}(V)$ eine Umgebung von a ist, gibt es ein $n_0 \in \mathbb{N}$, so dass $a_n \in f^{-1}(V)$ für alle $n \geq n_0$. Dann aber ist $f(a_n) \in V$ für alle $n \geq n_0$, was die Konvergenz beweist.

(b) Da A_1 und A_2 abgeschlossen und $f|_{A_1}$ und $f|_{A_2}$ stetig sind, ist nur auf dem Schnitt $A_1 \cap A_2$ etwas zu zeigen. Sei also $a \in A_1 \cap A_2$ und $(a_n)_n$ eine Folge in X mit $a_n \to a$, und für jedes $n_0 \in \mathbb{N}$ gebe es Folgenglieder $a_{n_1} \in A_1 \setminus A_2$ und $a_{n_2} \in A_2 \setminus A_1$ für $n_1, n_2 \geq n_0$. (Sonst ist wieder nichts zu zeigen.) Für jedes $\varepsilon > 0$ gibt es aber ein $n_0 \in \mathbb{N}$, so dass $|f(a_{n_1}) - f(a)| < \varepsilon$ und $|f(a_{n_2}) - f(a)| < \varepsilon$, d. h.

$$\begin{aligned}\big|f(a_{n_1}) - f(a_{n_2})\big| &= \big|f(a_{n_1}) - f(a) + f(a) - f(a_{n_2})\big| \\ &\leq \big|f(a_{n_1}) - f(a)\big| + \big|f(a) - f(a_{n_2})\big| < 2\varepsilon\,,\end{aligned}$$

also $f(a_n) \to f(a)$. Damit ist f stetig in a.

(c) Wähle $X = [0,2]$, $A_1 = [0,1]$, $A_2 = (1,2]$, $Y = \mathbb{R}$ und $f = \mathbb{1}_{A_1} : X \to Y$. Dann sind natürlich $f|_{A_1} \equiv 1$ und $f|_{A_2} \equiv 0$ stetig, aber f nicht.

(d) Seien $A_1 = \{(x,y)^{\mathrm{T}} \mid x \leq y\}$ und $A_2 = \{(x,y)^{\mathrm{T}} \mid x \geq y\}$ abgeschlossen mit $X = A_1 \cup A_2$. Ferner sind $f|_{A_1}$ mit $f|_{A_1}(x,y) = x - y$ und $f|_{A_2}$ mit $f|_{A_2}(x,y) = (x-y)^2$ stetig, also nach (b) auch f.

11. (a) Für $p = 1$ ergibt sich

$$\begin{aligned}\|T\| &= \sup\Big\{|Tx| \;\Big|\; x \in \mathbb{R}^n\,,\ |x|_1 \leq 1\Big\} = \sup\Big\{|\langle x, a\rangle| \;\Big|\; x \in \mathbb{R}^n\,,\ |x|_1 \leq 1\Big\} \\ &\overset{\text{i.}}{=} \sup\Big\{|x|_2|a|_2 \;\Big|\; x \in \mathbb{R}^n\,,\ |x|_1 \leq 1\Big\} = |a|_2 \sup\Big\{|x|_2 \;\Big|\; x \in \mathbb{R}^n\,,\ |x|_1 \leq 1\Big\} \\ &\overset{\text{ii.}}{=} |a|_2\,.\end{aligned}$$

Bei i. wurde die Cauchy-Schwarz-Ungleichung $|\langle x, a\rangle| \leq |x|_2|a|_2$ benutzt. Durch Bilden des Supremums gilt allerdings Gleichheit, da es immer ein x gibt, das zu a linear abhängig ist.

Bei ii. wurde die Normabschätzung $|x|_2 \leq |x|_1$ nach Aufgabe 2. (a) benutzt, und auch hier gilt Gleichheit, da das Supremum auch über $x = (1, 0, \ldots, 0)^{\mathrm{T}}$ gebildet wird.

Für $p = \infty$ ergibt sich

$$\|T\| = \ldots = |a|_2 \sup\Big\{|x|_2 \;\Big|\; x \in \mathbb{R}^n\,,\ |x|_\infty \leq 1\Big\} \overset{\text{iii.}}{=} \sqrt{n}\,|a|_2\,.$$

Bei iii. wurde die Normabschätzung $|x|_2 \leq \sqrt{n}\,|x|_\infty$ nach Aufgabe 2. (a,b) benutzt, und auch hier gilt Gleichheit, da das Supremum auch über $x = (1, \ldots, 1)^{\mathrm{T}}$ gebildet wird.

(b) Für $p = \infty$ seien $f, g \in \mathcal{C}(I)$. Dann ist $|Tf - Tg| = |f(1) - g(1)|$ und

$$\|f - g\|_\infty = \sup\Big\{\big|f(x) - g(x)\big| \;\Big|\; x \in I\Big\} \geq \big|f(1) - g(1)\big|$$

d. h. $|Tf - Tg| \leq \|f - g\|_\infty$. Also ist T sogar Lipschitz-stetig mit Konstante 1.
Für $p < \infty$ definieren wir für jedes $n \in \mathbb{N}$ eine Funktion $f_n \in \mathcal{C}(I)$ durch $f_n(x) = x^n$; ferner sei $g \in \mathcal{C}(I)$ die Nullfunktion. Dann ist $|Tf_n - Tg| = 1$ für alle $n \in \mathbb{N}$, aber

$$\|f_n - g\|_p^p = \int_0^1 \big|f_n(x)\big|^p \,\mathrm{d}x = \int_0^1 x^{np} \,\mathrm{d}x = \frac{x^{np+1}}{np+1}\bigg|_{x=0}^{x=1} = \frac{1}{np+1}\,.$$

Das bedeutet, dass $f_n \to g$ in der p-Norm für $n \to \infty$, aber $Tf_n \not\to Tg$ im Betrag für $n \to \infty$. Also ist T nicht stetig.

12. (a) i. Sei $(x_k)_k$ konvergent gegen $a \in X$, d. h. für jedes $\varepsilon > 0$ gibt es ein $k_0 \in \mathbb{N}$ mit $d(x_k, a) < \varepsilon/2$ für alle $k \geq k_0$. Dann gilt aber für alle $k, k' \geq k_0$

$$d(x_k, x_{k'}) \leq d(x_k, a) + d(a, x_{k'}) < \varepsilon\,,$$

d. h. $(x_k)_k$ ist eine Cauchyfolge.

ii. Gelte die Cauchyfolgeneigenschaft für ein beliebiges $\varepsilon > 0$, d. h. es gebe ein $k_0 \in \mathbb{N}$ mit $d(x_k, x_{k'}) < \varepsilon$ für alle $k, k' \geq k_0$. Für die Menge $M := \{x_k \mid k \in \mathbb{N}\}$ aller Folgenglieder gilt dann

$$\begin{aligned}\operatorname{diam} M &= \sup\{d(x_k, x_{k'}) \mid k, k' \in \mathbb{N}\} \\ &= \max\Big\{\sup \underbrace{\{d(x_k, x_{k_0}) \mid k < k_0\}}_{=:M_1},\ \sup \underbrace{\{d(x_{k_0}, x_{k'}) \mid k_0 < k'\}}_{=:M_2}, \\ &\qquad \sup \underbrace{\{d(x_k, x_{k'}) \mid k < k_0 < k'\}}_{=:M_3}\Big\}\,.\end{aligned}$$

M_1 enthält nur endlich viele Elemente, d. h. $\sup M_1 < \infty$. Weiter ist $\sup M_2 \leq \varepsilon$, weil $d(x_{k_0}, x_{k'}) < \varepsilon$ für alle $k' > k_0$ gilt. Für ein Element von M_3 gilt $d(x_k, x_{k'}) \leq d(x_k, x_{k_0}) + d(x_{k_0}, x_{k'})$, wobei der der erste Summand vom Typ aus M_1 und der zweite vom Typ aus M_2 ist. Damit ist $\operatorname{diam} M < \infty$, also die Folge beschränkt.

iii. Für alle $\ell \in \mathbb{N}$ ist

$$d(x_\ell, a) \leq d(x_\ell, x_{k_j}) + d(x_{k_j}, a)\,.$$

Sei nun $\varepsilon > 0$. Dann gibt es ein $\ell_0 \in \mathbb{N}$ mit $d(x_\ell, x_{k_j}) < \varepsilon/2$ für alle $\ell, k_j \geq \ell_0$, weil $(x_\ell)_\ell$ eine Cauchyfolge ist. Außerdem gibt es ein $j_0 \in \mathbb{N}$ mit $d(x_{k_j}, a) < \varepsilon/2$ für alle $j \geq j_0$ wegen der Konvergenz der Teilfolge.
Wählt man in der obigen Ungleichung also $\ell \geq \ell_0$ beliebig, so erhält man $d(x_\ell, a) \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$, sofern $j \geq \max\{j_0, \ell_0\}$ gewählt wird. Dies zeigt $x_\ell \to a$ für $\ell \to \infty$.

(b) Sei $(x_k)_k$ eine Cauchyfolge in (A, d). Wegen der Kompaktheit von A besitzt sie eine in (A, d) konvergente Teilfolge, aber nach (a) iii. ist dann bereits die Folge selbst konvergent. Also ist (A, d) vollständig.

(c) Seien $a \in X$ ein Häufungspunkt von A und $(x_k)_k$ eine Folge in $A \setminus \{a\}$ mit $x_k \to a$. (Dass so eine Folge existiert, ist Definition des Häufungspunktes.) Ist (A, d) vollständig, dann liegt der Grenzwert a in A, d. h. A enthält alle ihre Häufungspunkte, also ist A abgeschlossen.

Sei nun (A, d) nicht vollständig, d. h. es gebe eine Cauchyfolge $(x_k)_k$ in (A, d), die nicht konvergiert. Wegen der Vollständigkeit von (X, d) konvergiert $(x_k)_k$ jedoch gegen ein $a \in X$. Da $(x_k)_k$ in (A, d) nicht konvergiert, muss $a \notin A$ gelten, aber a ist ein Häufungspunkt von A, und damit ist A nicht abgeschlossen in X.

13. (a) Wir beweisen diese Aussage absichtlich sehr ausführlich mit der Überdeckungseigenschaft. Wir wählen eine beliebge offene Überdeckung von B, d. h. eine Familie $(U_i)_{i \in I}$ offener Mengen $U_i \subset X$ für alle $i \in I$ mit einer Indexmenge I, so dass $B \subset \bigcup_{i \in I} U_i$. Nun stellen wir fest, dass

$$A \subset (X \setminus B) \cup \bigcup_{i \in I} U_i \,,$$

denn für jedes $a \in A$ gilt: Ist $a \in B$, so ist $a \in \bigcup_{i \in I} U_i$; ist $a \notin B$, so ist $a \in X \setminus B$. Ferner ist B nach Voraussetzung abgeschlossen und daher $X \setminus B$ offen, d. h. $(X \setminus B, U_i)_{i \in I}$ ist eine offene Überdeckung von A. Da A kompakt ist, existiert eine endliche Teilüberdeckung, d. h. Indizes $i_1, \ldots, i_n \in I$ mit

$$A \subset (X \setminus B) \cup \bigcup_{\ell=1}^{n} U_{i_\ell} \,.$$

($X \setminus B$ lassen wir o. B. d. A. dabei.) Nun ist aber $(U_{i_1}, \ldots, U_{i_n})$ eine offene Überdeckung von B, denn ist $b \in B$, so gilt $b \notin X \setminus B$. Damit besitzt jede offene Überdeckung $(U_i)_{i \in I}$ von B eine endliche Teilüberdeckung, d. h. B ist kompakt.

Bemerkung: In diesem Beweis wurde an keiner Stelle eine Eigenschaft der Metrik benutzt. Daher ist er ohne Veränderung sofort für topologische Hausdorff-Räume verwendbar.

(b) Wir nehmen an, f sei nicht gleichmäßig stetig. Dann existiert ein $\varepsilon > 0$, so dass es für jedes $\delta = 1/n > 0$ Punkte $a_n, b_n \in A$ gibt mit

$$d_X(a_n, b_n) < \frac{1}{n} \quad \text{und} \quad d_Y\big(f(a_n), f(b_n)\big) \geq \varepsilon \,.$$

Wegen der Kompaktheit von A haben die Folgen $(a_n)_n$ und $(b_n)_n$ konvergente Teilfolgen $(a_{m_k})_k$ und $(b_{n_k})_k$. Da aber $d_X(a_{m_k}, b_{n_k}) \to 0$ für $k \to \infty$, haben beide Teilfolgen denselben Grenzwert $x \in A$. Es ist jedoch $d_Y(f(a_{m_k}), f(b_{n_k})) \geq \varepsilon$, also ein Widerspruch zur Stetigkeit von f in x.

Zusatz: Motiviert durch das Überdeckungsbeispiel wollen wir ein paar Grundlagen zu topologischen Räumen betrachten, die in der Vorlesung keinen Platz gefunden haben, aber für spätere Veranstaltungen durchaus wesentlich sind. Leider ist dieser Zusatz mal wieder länger geworden als erhofft/erwartet. Im Folgenden sei $X \neq \emptyset$ eine beliebige Menge.

Eine Teilmenge $\mathcal{T} \subset \mathcal{P}(X)$ der Potenzmenge von X heißt eine *Topologie* auf X, falls:

- Es gilt $\emptyset, X \in \mathcal{T}$.
- Sind $U_i \in \mathcal{T}$ für alle $i \in I$ mit einer beliebigen Indexmenge I, so ist auch $\bigcup_{i \in I} U_i \in \mathcal{T}$.
- Sind $U, V \in \mathcal{T}$, so ist auch $U \cap V \in \mathcal{T}$.

Ist $\mathcal{T}$ eine Topologie auf X, so heißt das Paar $(X, \mathcal{T})$ ein *topologischer Raum*. Eine Menge $M \subset X$ heißt genau dann *offen (in X)*, wenn $M \in \mathcal{T}$, und *abgeschlossen (in X)*, wenn $X \setminus M$ offen ist. Für $X' \subset X$ heißt $(X', \mathcal{T}')$ ein *topologischer Unterraum* von $(X, \mathcal{T})$, falls $\mathcal{T}'$ die von $\mathcal{T}$ *induzierte Teilraumtopologie (Relativtopologie)* auf X' ist, d. h. es gilt $\mathcal{T}' = \{X' \cap U \mid U \in \mathcal{T}\}$. Eine Menge in $(X', \mathcal{T}')$ ist also genau dann offen, wenn sie der Schnitt von X' mit einer offenen Menge in aus $(X, \mathcal{T})$ ist. Eine solche Menge heißt *relativ offen* in X bezüglich X'. Für eine Menge $M \subset X$ definieren wir das *Innere* durch $M^\circ = \bigcup_{M \supset U \text{ offen}} U$ und den *Abschluss* durch $\bar{M} = \bigcap_{M \subset A \text{ abg.}} A$.

Eine Menge $U \in \mathcal{T}$ heißt eine *offene Umgebung* von $x \in X$, falls $x \in U$. Eine Menge $M \subset X$ heißt eine *Umgebung* von $x \in X$, falls es eine Menge $U \in \mathcal{T}$ mit $x \in U \subset M$ gibt. Ein topologischer Raum $(X, \mathcal{T})$ heißt *hausdorffsch* oder ein *(topologischer) Hausdorff-Raum*, falls es zu je zwei Punkten $x, y \in X$ mit $x \neq y$ offene Umgebungen U von x und V von y mit $U \cap V = \emptyset$ gibt. In Hausdorff-Räumen kann man einen sinnvollen Konvergenzbegriff definieren: Eine Folge $(a_n)_n$ in X konvergiert genau dann gegen ein $a \in X$, wenn in jeder Umgebung von a fast alle Folgenglieder liegen. Zu jeder Umgebung M von a gibt es also ein $n_0 \in \mathbb{N}$, so dass $a_n \in M$ für alle $n \geq n_0$.

Das Paar $(X, \mathcal{T})$ mit $X = \mathbb{R}$ und $\mathcal{T} = \{\emptyset, \mathbb{R}\}$ ist ein topologischer Raum, der die gröbstmögliche Topologie, die sog. *Klumpentopologie*, trägt. Er ist nicht hausdorffsch, denn die einzige Umgebung eines jeden Punktes ist ganz $\mathbb{R}$. Nach der obigen Definition der Folgenkonvergenz würde also jede Folge gegen jeden Grenzwert konvergieren. Das Paar $(X, \mathcal{T})$ mit $X = \mathbb{R}$ und $\mathcal{T} = \mathcal{P}(\mathbb{R})$ ist ein topologischer Raum, der die feinstmögliche Topologie, die sog. *diskrete Topologie*, trägt. Zu jedem Punkt $x \in \mathbb{R}$ ist $\{x\}$ eine (sogar offene) Umgebung. Eine Folge $(a_n)_n$ konvergiert also gegen $a \in X$, falls fast alle Folgenglieder in $\{a\}$ liegen, also wenn die Folge stationär wird. Jeder metrische Raum ist hausdorffsch bezüglich der durch die Metrik d induzierten Topologie; sie ist die gröbste Topologie, die alle offenen Kugeln $U_r(x) = \{y \in X \mid d(x, y) < r\}$ enthält. Beispielsweise wird durch die *diskrete Metrik* $d(x, y) = 1$ für alle $x \neq y$ die diskrete Topologie induziert.

Ein topologischer Raum $(X, \mathcal{T})$ bzw. eine Teilmenge $M \subset X$ heißt *quasikompakt*, wenn jede offene Überdeckung eine endliche Teilüberdeckung besitzt. Ist der Raum zusätzlich hausdorffsch, so spricht man von *kompakt*. In Aufgabe 13. (a) haben wir gezeigt, dass abgeschlossene Teilmengen kompakter Mengen kompakt sind. In Aufgabe 19. (a) wird gezeigt, dass der Schnitt einer kompakten mit einer abgeschlossenen Menge kompakt ist. Wir zeigen hier noch, dass kompakte Mengen abgeschlossen sind: Seien dazu $(X, \mathcal{T})$ hausdorffsch, $K \subset X$ kompakt und $a \in X \setminus K$ ein fester Punkt. Zu jedem $x \in K$ gibt es nun eine offene Umgebung $U(x)$, so dass $a \notin \bar{U}(x)$ (Hausdorff-Eigenschaft). Dann gilt offenbar $K \subset \bigcup_{x \in K} U(x)$, und da K kompakt ist, reichen endlich viele $\{x_1, \ldots, x_n\} \subset K$ aus, so dass $K \subset \bigcup_{\ell=1}^n U(x_\ell)$. Daraus folgt durch Abschlussbildung $\bar{K} \subset \bigcup_{\ell=1}^n \bar{U}(x_\ell)$ und damit $a \notin \bar{K}$. Da a beliebig war, folgt $\bar{K} = K$, d. h. K ist abgeschlossen.

Aus Aufgabe 19. (a) folgt insbesondere, dass der Schnitt zweier kompakter Mengen kompakt ist. Diese Aussage kann nicht ohne weiteres auf quasikompakte Mengen verallgemeinert werden, wie das folgende Gegenbeispiel zeigt: Seien $X = \mathbb{N} \,\dot{\cup}\, \{a, b\}$ mit $a \neq b$ und $\mathcal{T} = \mathcal{P}(\mathbb{N}) \cup \{A, B, X\}$ mit $A = \mathbb{N} \cup \{a\}$ und $B = \mathbb{N} \cup \{b\}$. Dann ist $(X, \mathcal{T})$ ein topologischer Raum, der nicht hausdorffsch ist, denn z. B. lassen sich die Punkte $1 \in X$ und $a \in X$ nicht durch Umgebungen trennen. Die beiden Mengen A und B sind quasikompakt, denn jede offene Überdeckung von A muss eine offene Umgebung von a enthalten, d. h. A oder X. Damit ist eine mögliche endliche Teilüberdeckung einer beliebigen Überdeckung von A z. B. lediglich $\{A\}$ oder $\{X\}$; ebenso funktioniert der Beweis für B. Der Schnitt $A \cap B = \mathbb{N}$ ist natürlich nicht quasikompakt, wie die offene Überdeckung $\mathbb{N} \subset \bigcup_{n \in \mathbb{N}} \{n\}$ zeigt.

Eine Funktion $f \colon X \to X'$ zwischen zwei topologischen Räumen $(X, \mathcal{T})$ und $(X', \mathcal{T}')$ heißt *stetig*, wenn Urbilder offener Mengen wieder offen sind, d. h. wenn $f^{-1}(U') \in \mathcal{T}$ für alle $U' \in \mathcal{T}'$. (Statt der offenen können auch die abgeschlossenen Mengen benutzt werden.) Sind $K \subset X$ kompakt, X' hausdorffsch und $f \colon X \to X'$ stetig, so ist $f(K)$ kompakt. Ist zusätzlich $X' \subset \mathbb{R}$, dann bedeutet dies, dass $f(K)$ beschränkt und bezüglich der von $d(x, y) = |x - y|$ induzierten Topologie abgeschlossen ist, d. h. f nimmt ein globales Maximum und Minimum auf K an. Durch Verfeinern von $\mathcal{T}$ bzw. Vergröbern von $\mathcal{T}'$ kann jede Funktion $f \colon X \to X'$ stetig gemacht werden. Im Extremfall $\mathcal{T} = \mathcal{P}(X)$ oder $\mathcal{T}' = \{\emptyset, X'\}$ ist jede Funktion stetig.

Ist $f \colon X \to X'$ bijektiv und stetig und $f^{-1} \colon X' \to X$ ebenfalls stetig, so heißt f ein *Homöomorphismus*. Die Funktion $f \colon (-\pi, \pi] \to S^1$ mit der Einheitskreislinie $S^1 = \{(x, y)^{\mathrm{T}} \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ und $f(t) = (\cos t, \sin t)^{\mathrm{T}}$ ist bijektiv und stetig, f^{-1} jedoch ist unstetig in $(-1, 0)^{\mathrm{T}}$, d. h. f ist kein Homöomorphismus. Das rührt daher, dass zwar $(\pi - \varepsilon, \pi]$ relativ offen in $(-\pi, \pi]$, aber $f((\pi - \varepsilon, \pi])$ nicht relativ offen in S^1 ist, d. h. offene Mengen müssen keine offenen Urbilder unter f^{-1} haben. Homöomorphismen jedoch haben gute Eigenschaften: Sie bilden kompakte Mengen/Räume auf kompakte Mengen/Räume ab, und sie bilden (weg)zusammenhängende Mengen/Räume auf (weg)zusammenhängende Mengen/Räume ab. Allerdings ist die Abbildung $\arctan\colon \mathbb{R} \to (-\pi/2, \pi/2)$ ein Homöomorphismus, obwohl der Wertebereich beschränkt ist und der Definitionsbereich nicht. Das liegt daran, dass Beschränktheit keine topologische, sondern eine metrische Eigenschaft ist.

Die Menge S^1 ist kompakt und zusammenhängend. Gäbe es einen Homöomorphismus $f\colon I \to S^1$ mit $I \subset \mathbb{R}$, so müsste also auch I kompakt und zusammenhängend sein, also ein beschränktes Intervall. Entfernt man nun aus S^1 irgendeinen Punkt, so bleibt die Menge zusammenhängend, das Intervall jedoch wird unzusammenhängend, wenn man einen inneren Punkt herausnimmt. Somit kann es keinen Homöomorphismus geben. Wir betrachten statt dessen $f\colon \widehat{\mathbb{R}} \to S^1$ mit $\widehat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$, $f(t) = (\cos(2\arctan t), \sin(2\arctan t))^{\mathrm{T}} = (t^2-1, 2t)^{\mathrm{T}}/(t^2+1)$ für $t \in \mathbb{R}$ und $f(\infty) = (-1,0)^{\mathrm{T}}$. Dann ist zunächst f bijektiv und stetig in jedem $t \in \mathbb{R}$. Wir definieren nun: Eine Teilmenge $U \subset \widehat{\mathbb{R}}$ mit $\infty \in U$ heißt offen, wenn $\widehat{\mathbb{R}} \setminus U$ kompakt ist. Damit wird f stetig in ∞, und nun ist auch f^{-1} stetig, also f ein Homöomorphismus. $\widehat{\mathbb{R}}$ ist also kompakt; man nennt $\widehat{\mathbb{R}}$ die *Ein-Punkt-Kompaktifizierung* von $\mathbb{R}$. Insbesondere bleibt $\widehat{\mathbb{R}}$ – so wie S^1 – zusammenhängend, wenn man einen Punkt entfernt.

Auf der kompakten Menge $\widehat{\mathbb{R}}$ gelten dieselben Sätze aus EHM, die auf kompakten Teilmengen von $\mathbb{R}$ gelten. Zum Beispiel sagt Bolzano-Weierstraß: Jede beschränkte Folge besitzt eine konvergente Teilfolge. Anders gesprochen: Jede Folge mit Werten in einer kompakten Menge besitzt eine konvergente Teilfolge. Dies gilt auch für $\widehat{\mathbb{R}}$: Entweder ist die Folge in $\mathbb{R}$ beschränkt, dann gilt der Satz wie bekannt. Ist $(a_n)_n$ jedoch unbeschränkt, so existieren Folgenglieder a_n mit $|a_n| > R$ für jedes $R > 0$. Damit gibt es sogar unendlich viele Folgenglieder dieser Art für jedes R, d. h. in jeder Umgebung um ∞ liegen unendlich viele Folgenglieder. Damit gibt es eine konvergente Teilfolge, nämlich eine, die gegen ∞ konvergiert.

14. Alle f_n sind punktsymmetrisch zum Ursprung. Für $m, n \in \mathbb{N}$ mit $n > m$ gilt $f_n(x) \geq f_m(x)$ für alle $x \in [0,1]$ und damit

$$\begin{aligned}
\|f_n - f_m\|_p^p &= \int_{-1}^{1} \bigl|f_n(x) - f_m(x)\bigr|^p \,\mathrm{d}x = 2\int_0^1 \bigl(f_n(x) - f_m(x)\bigr)^p \,\mathrm{d}x \\
&= 2\int_0^{1/n} (nx - mx)^p \,\mathrm{d}x + 2\int_{1/n}^{1/m} (1 - mx)^p \,\mathrm{d}x \\
&= \frac{2(n-m)^p x^{p+1}}{p+1}\bigg|_{x=0}^{x=1/n} - \frac{2(1-mx)^{p+1}}{m(p+1)}\bigg|_{x=1/n}^{x=1/m} \\
&= \frac{2}{p+1}\left(\frac{(n-m)^p}{n^{p+1}} + \frac{(1-m/n)^{p+1}}{m}\right) \\
&= \frac{2}{p+1}\left(\frac{1}{n}\Bigl(1 - \frac{m}{n}\Bigr)^p + \frac{1}{m}\Bigl(1 - \frac{m}{n}\Bigr)^{p+1}\right) \leq \frac{4}{m(p+1)}\,,
\end{aligned}$$

also ist $(f_n)_n$ eine Cauchyfolge. Die punktweise Grenzfunktion $f\colon I \to \mathbb{R}$ ist gegeben durch $f(x) = -1$ für $x < 0$, $f(x) = 1$ für $x > 0$ und $f(0) = 0$, und es gilt

$$\begin{aligned}
\|f_n - f\|_p^p &= \int_{-1}^{1} \bigl|f_n(x) - f(x)\bigr|^p \,\mathrm{d}x = 2\int_0^1 \bigl(1 - f_n(x)\bigr)^p \,\mathrm{d}x \\
&= 2\int_0^{1/n} (1 - nx)^p \,\mathrm{d}x = -\frac{2(1-nx)^{p+1}}{n(p+1)}\bigg|_{x=0}^{x=1/n} = \frac{2}{n(p+1)}\,.
\end{aligned}$$

$(f_n)_n$ konvergiert also tatsächlich in der p-Norm gegen f. Da dieses f aber nach der Eindeutigkeit des Grenzwerts die einzige derartige Funktion ist und da aber f nicht stetig ist, ist $(f_n)_n$ in $(\mathcal{C}(I), \|\cdot\|_p)$ *nicht* konvergent.

Zusatz: Die Stetigkeit der Grenzfunktion kann nur dann sichergestellt werden, wenn eine Folge stetiger Funktionen gleichmäßig konvergiert, was äquivalent zur Konvergenz bezüglich der Supremumsnorm $\|\cdot\|_\infty$

ist. Der Raum $(\mathcal{C}(I), \|\cdot\|_\infty)$ ist also tatsächlich vollständig und damit ein Banachraum. Der Beweis ist ebenso einfach wie typisch: Für jedes $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass $|f_n(x) - f_n(y)| < \varepsilon/3$ (gleichmäßige Stetigkeit von f auf dem Kompaktum I) und $|f_n(x) - f(x)| < \varepsilon/3$ (gleichmäßige Konvergenz) für alle $x, y \in I$ mit $|x - y| < \delta$. Dann ist

$$\begin{aligned}\big|f(x) - f(y)\big| &= \big|f(x) - f_n(x) + f_n(x) - f_n(y) + f_n(y) - f(y)\big| \\ &\le \big|f(x) - f_n(x)\big| + \big|f_n(x) - f_n(y)\big| + \big|f_n(y) - f(y)\big| < \varepsilon\,.\end{aligned}$$

Rechnet man für die Funktionenfolge $(f_n)_n$ aus der Aufgabe

$$\begin{aligned}\|f_n - f_m\|_\infty &= \sup\Big\{\big|f_n(x) - f_m(x)\big| \;\Big|\; x \in I\Big\} = \sup\Big\{f_n(x) - f_m(x) \;\Big|\; x \in [0,1]\Big\} \\ &= \max\Big\{\sup\Big\{(n-m)x \;\Big|\; x \in \Big[0, \frac{1}{n}\Big]\Big\}\,,\ \sup\Big\{1 - mx \;\Big|\; x \in \Big[\frac{1}{n}, \frac{1}{m}\Big]\Big\}\Big\} \\ &= \max\Big\{\frac{n-m}{n}, 1 - \frac{m}{n}\Big\} = 1 - \frac{m}{n}\,,\end{aligned}$$

so sieht man, dass dieser Ausdruck keineswegs gegen 0 konvergiert, nur weil m und n groß werden. $(f_n)_n$ ist also keine Cauchyfolge in $(\mathcal{C}(I), \|\cdot\|_\infty)$.

15. (a) Gefragt ist nach den Lösungen der Gleichung

$$f(x) = \mathrm{e}^{-x/10} = x =: g(x)\,.$$

Da f streng monoton fallend und g streng monoton wachsend auf ganz $\mathbb{R}$ ist, was man z. B. mit der Ableitung beweisen kann, kann es höchstens einen Fixpunkt geben. Wegen $\lim_{x\to\pm\infty} g(x) = \pm\infty$ existiert dieser nach dem Zwischenwertsatz, da f und g stetig sind, d. h. es gibt einen eindeutigen Fixpunkt $x_0 \in \mathbb{R}$.

(b) Wir setzen $K := [0,1]$ und stellen fest, dass $f(0) = 1 > g(0) = 0$ und $f(1) < 1 = g(1)$. Daher muss der Fixpunkt x_0 im Innern des Kompaktums K liegen. Ferner ist $f|_K$ eine Selbstabbildung, denn es gilt $f(x) \in K$ für alle $x \in K$. Da f stetig differenzierbar ist, können wir den Mittelwertsatz der Differentialrechung anwenden und erhalten für $x, y \in K$

$$\big|f(x) - f(y)\big| = \big|f'(\xi)(x-y)\big| = \left|-\frac{1}{10}\,\mathrm{e}^{-\xi/10}\right| |x-y| \le \frac{1}{10}|x-y|\,.$$

Damit ist $f|_K$ eine Kontraktion mit Kontraktionsfaktor $1/10$. Also sind alle Voraussetzungen für den Banachschen Fixpunktsatz erfüllt, und für jedes $x \in K$ konvergiert die Iterationsfolge $(f^n(x))_n$ gegen x_0. Also ist x_0 ein attraktiver Fixpunkt.

(c) Wir wissen aus (b), dass das gesuchte Intervall mindestens $[0,1]$ ist. Wählen wir $x > 1$, so folgt $0 < f(x) < 1$, weswegen wir nach einer Iteration in $[0,1]$ landen und damit gegen x_0 konvergieren. Starten wir mit einem $x < 0$, so gilt $f(x) > 0$, d. h. wir landen wieder in einem Bereich, von dem wir schon wissen, dass die Folge gegen x_0 konvergiert. Also gilt $f^n(x) \to x_0$ für $n \to \infty$ für alle $x \in \mathbb{R}$.

16. (a) Es ist

$$\begin{aligned}\partial_k |x|_p = \partial_k \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} &= \frac{1}{p}\left(\sum_{i=1}^n |x_i|^p\right)^{1/p-1} \partial_k \sum_{i=1}^n |x_i|^p \\ = \frac{1}{p}\left(\sum_{i=1}^n |x_i|^p\right)^{-(p-1)/p} &\sum_{i=1}^n \partial_k (x_i^2)^{p/2} = \frac{1}{p}|x|_p^{-(p-1)} \sum_{i=1}^n \frac{p}{2}(x_i^2)^{p/2-1}\cdot 2x_i\delta_{ik}\end{aligned}$$

$$= |x|_p^{-(p-1)} |x_k|^{p-2} x_k$$

und damit

$$\nabla |x|_p = \sum_{k=1}^{n} \partial_k |x|_p \, e_k = \frac{1}{|x|_p^{p-1}} \begin{pmatrix} |x_1|^{p-2} x_1 \\ \vdots \\ |x_n|^{p-2} x_n \end{pmatrix}.$$

Bemerkung: Für $p = 2$ ergibt sich daraus $\nabla |x|_2 = x/|x|_2$. (Merken!)

Weiter ist

$$\nabla \frac{1}{|x|_2} = \nabla |x|_2^{-1} = -|x|_2^{-2} \nabla |x|_2 \overset{\text{(a)}}{=} -\frac{x}{|x|_2^3}.$$

(b) Für $v \in \mathbb{R}^n \setminus \{0\}$ und $t \neq 0$ ist

$$\frac{f(0+tv) - f(0)}{t} = \frac{|tv|_p}{t} = \frac{|t|\,|v|_p}{t} = \operatorname{sgn}(t)\,|v|_p.$$

Für $t > 0$ ist dieser Ausdruck konstant $|v_p| > 0$, für $t < 0$ konstant $-|v_p| < 0$. Somit kann der Grenzwert für $t \to 0$ nicht existieren.

17. (a) Für $v \in \mathbb{R}^2 \setminus \{0\}$ und $t \neq 0$ hängt der Ausdruck

$$\frac{f(0+tv) - f(0)}{t} = \frac{t^2 v_1^2 t v_2}{(t^2 v_1^2 + t^2 v_2^2) t} = \frac{v_1^2 v_2}{v_1^2 + v_2^2}$$

nicht mehr von t ab, so dass der Grenzwert $t \to 0$ und damit die Richtungsableitung $\partial_v f(0)$ existiert.

(b) Es ist

$$\big|f(x) - f(0)\big| = \left| \frac{x_1^2 x_2}{x_1^2 + x_2^2} \right| = \frac{x_1^2}{x_1^2 + x_2^2} |x_2| \leq |x_2| \xrightarrow{x \to 0} 0,$$

d. h. f ist stetig in 0.

(c) Mit (a) ist

$$\partial_{e_1} f(0) = \partial_{e_2} f(0) = 0 \quad \text{und} \quad \partial_{e_1+e_2} f(0) = \frac{1}{2},$$

d. h. $\partial_{e_1+e_2} f(0) \neq \partial_{e_1} f(0) + \partial_{e_2} f(0)$. Damit ist die Abbildung $v \mapsto \partial_v f(0)$ nicht linear. Wäre f differenzierbar, so würde $\partial_v f(0) = \mathrm{d}f(0)(v)$ gelten, was linear in v wäre. Somit kann es die lineare Abbildung $\mathrm{d}f(0)\colon \mathbb{R}^2 \to \mathbb{R}$ nicht geben, d. h. f ist nicht differenzierbar in 0.

Bemerkung: Ist $f\colon U \to \mathbb{R}$, $U \subset \mathbb{R}^n$ offen, in $a \in U$ differenzierbar, so wissen wir, dass erstens alle Richtungsableitungen in a existieren und zweitens f stetig in a ist. Beide Folgerungen sind hier erfüllt (nämlich (a) und (b)), aber es folgt trotzdem nicht die Differenzierbarkeit!

18. (a) Für $v \in \mathbb{R}^2 \setminus \{0\}$ und $t \neq 0$ betrachten wir

$$\frac{f(0+tv) - f(0)}{t} = \frac{f(tv)}{t}.$$

Der Zähler ist genau dann ungleich 0, wenn $t^2 v_1^2 = t v_2$ ist, d. h. $t v_1^2 = v_2$, da $t \neq 0$. Diese Gleichung hat aber höchstens eine Lösung $t_0 \in \mathbb{R}$, d. h. der Zähler ist 0, falls $|t| < |t_0|$. Damit ist der Grenzwert für $t \to 0$ ebenfalls 0, also $\partial_v f(0) = 0$.

(b) In natürlicher Weise benutzen wir die Folge $x_n = (n^{-1}, n^{-2})^{\mathrm{T}}$. Dann ist $x_{n,1}^2 = x_{n,2}$ für alle $n \in \mathbb{N}$ und $x_n \to (0,0)^{\mathrm{T}}$ für $n \to \infty$. Es folgt

$$\big|f(x_n) - f(0)\big| = 1 \xrightarrow{n\to\infty} 1\,,$$

d. h. f ist nicht stetig in 0.

19. (a) Wir definieren $r := \operatorname{dist}(x, A)$, $K := B_{r+1}(x) \cap A$ und $f\colon K \to \mathbb{R}$ durch $f(a) = d_2(x, a)$. Dann gilt:

- K ist nicht leer wegen der Infimumseigenschaft in der Definition von dist.
- $B_{r+1}(x)$ ist als abgeschlossene und beschränkte Menge im $\mathbb{R}^n$ kompakt.
- K ist als Schnitt einer kompakten und abgeschlossenen Menge kompakt. (Wenn das nicht bekannt ist, siehe den allgemeinen Beweis im Zusatz. Da alles im $\mathbb{R}^n$ ist, kann man aber auch einfach zeigen, dass der Schnitt abgeschlossen und beschränkt ist.)
- f ist stetig (sogar Lipschitz-stetig mit Konstante 1), denn

$$\big|f(a_1) - f(a_2)\big| = \big|d(x, a_1) - d(x, a_2)\big| \le d(a_1, a_2)$$

  für alle $a_1, a_2 \in A$ wegen der umgekehrten Dreiecksungleichung.
- Da f stetig und K kompakt ist, nimmt f das Minimum an.

Damit gibt es ein $z \in K$ (und damit $z \in A$) mit $d_2(x, z) = f(z) = \inf\{f(a) \mid a \in A\} = \operatorname{dist}(x, A)$.

Zusatz: Wir zeigen, dass $K \cap A$ kompakt ist, wenn $K \subset X$ kompakt und $A \subset X$ abgeschlossen sind. Wir wählen eine beliebge offene Überdeckung von $K \cap A$, d. h. eine Familie $(U_i)_{i\in I}$ offener Mengen $U_i \subset X$ für alle $i \in I$ mit einer Indexmenge I, so dass $K \cap A \subset \bigcup_{i\in I} U_i$. Nun stellen wir fest, dass

$$K \subset (X \setminus A) \cup \bigcup_{i\in I} U_i\,,$$

denn für jedes $x \in K$ gilt: Ist $x \in A$, so ist $x \in \bigcup_{i\in I} U_i$; ist $x \notin A$, so ist $x \in X \setminus A$. Ferner ist A nach Voraussetzung abgeschlossen und daher $X \setminus A$ offen, d. h. $(X \setminus A, U_i)_{i\in I}$ ist eine offene Überdeckung von K. Da K kompakt ist, existiert eine endliche Teilüberdeckung, d. h. Indizes $i_1, \ldots, i_n \in I$ mit

$$K \subset (X \setminus A) \cup \bigcup_{\ell=1}^{n} U_{i_\ell}\,.$$

($X \setminus A$ lassen wir o. B. d. A. dabei.) Nun ist aber $(U_{i_1}, \ldots, U_{i_n})$ eine offene Überdeckung von $K \cap A$, denn ist $x \in K \cap A$, so gilt $x \notin X \setminus A$. Damit besitzt jede offene Überdeckung $(U_i)_{i\in I}$ von $K \cap A$ eine endliche Teilüberdeckung, d. h. $K \cap A$ ist kompakt.

(Wem es nicht aufgefallen sein sollte: Dieser Beweis war Copy & Paste von Aufgabe 13. (a)!)

(b) i. Wir definieren die Funktion $f\colon B \to \mathbb{R}$ durch $f(b) := \operatorname{dist}(A, b)$. Dann gilt:

- f ist stetig (sogar Lipschitz-stetig mit Konstante 1), denn

$$\big|f(b_1) - f(b_2)\big| = \big|\operatorname{dist}(A, b_1) - \operatorname{dist}(A, b_2)\big| \le d_2(b_1, b_2)$$

  für alle $b_1, b_2 \in B$ nach Aufgabe 8. (b).
- f ist nicht-negativ; das ist klar.
- Wegen $A \cap B = \emptyset$ gilt $B \subset X \setminus A$, und weil A abgeschlossen ist, ist $X \setminus A$ offen. Also ist jedes $b \in B$ ein äußerer Punkt von A, d. h. $0 \notin \operatorname{Bild} f$.

- Da f stetig und B kompakt ist, nimmt f das Minimum an. Es gibt also ein $b_0 \in B$ mit
$$f(b_0) = \min\{f(b) \mid b \in B\} = \inf\{\operatorname{dist}(A,b) \mid b \in B\} = d_m(A,B)\,.$$

Wegen $f(b_0) > 0$ folgt die Behauptung.

ii. A ist abgeschlossen als kartesisches Produkt abgeschlossener Mengen. Oder: Ist $(x,y)^{\mathrm{T}} \in \mathbb{R}^2 \setminus A$, d. h. $y \neq 0$, so ist $B_{|y|/2}((x,y)^{\mathrm{T}})$ eine Umgebung von $(x,y)^{\mathrm{T}}$ in $\mathbb{R}^2 \setminus A$. Damit ist $\mathbb{R}^2 \setminus A$ offen und A abgeschlossen.
B ist abgeschlossen als Schnitt des abgeschlossenen Nullstellengebildes von $xy-1$ und der abgeschlossenen Viertelebene $[0,\infty)^2$. Oder: Klar ist, dass jeder Punkt aus $\mathbb{R}^2 \setminus (0,\infty)^2$ ein äußerer Punkt von B ist. Sei $(x,y)^{\mathrm{T}} \in (0,\infty)^2 \setminus B$. Dann gibt es genau ein $c > 0$, $c \neq 1$, mit $xy - c = 0$. Daraus folgt aber, dass es ein $\varepsilon > 0$ gibt, so dass $(x-\varepsilon, x+\varepsilon) \times (y-\varepsilon, y+\varepsilon) \subset X \setminus B$, d. h. $X \setminus B$ ist offen und daher B abgeschlossen.
Dass A und B disjunkt sind, ist klar, weil $0 \neq 1/t$ für alle $t > 0$.
Wir zeigen nun, dass es für jedes $\varepsilon > 0$ Punkte $a \in A$ und $b \in B$ gibt, so dass $d(a,b) < \varepsilon$. Damit ist dann $d_m(A,B) = 0$ gezeigt. Wir wählen $a = (2/\varepsilon, 0)^{\mathrm{T}} \in A$ und $b = (2/\varepsilon, \varepsilon/2)^{\mathrm{T}} \in B$. Dann ist $d(a,b) = \varepsilon/2 < \varepsilon$.

Bemerkung: Schon die Wohldefiniertheit macht Schwierigkeiten: Was ist $d_m(A,\emptyset)$? Wegen der Vereinbarung $\inf\emptyset = \infty$ könnte man dies auch für d_m verwenden. Offenbar ist dann d_m nichtnegativ und symmetrisch. Allerdings kann $d_m(A,B) = 0$ für $A \neq B$ gelten, beispielsweise wenn $A \cap B \neq \emptyset$. Auch die Dreiecksungleichung gilt nicht, wie das Beispiel $A = [0,1]$, $B = [1,2]$ und $C = [2,3]$ mit $1 = d_m(A,C) \not\leq d_m(A,B) + d_m(B,C) = 0$ zeigt.

20. (a) Es ist $f_1(x) = |x|_2^3 + 1$ und damit nach Aufgabe 16. (a)
$$\nabla f_1(x) = 3|x|_2^2\, \nabla |x|_2 = 3|x|_2^2\, \frac{x}{|x|_2} = 3|x|_2\, x\,.$$

Ferner ist $f_2(x) = \cos|x|_2$, d. h.
$$\nabla f_2(x) = -\sin|x|_2\, \nabla|x|_2 = -\frac{\sin|x|_2}{|x|_2}\, x\,.$$

Damit ist die Jacobimatrix gegeben durch
$$J_f(x) = \begin{pmatrix} \nabla f_1(x) & \nabla f_2(x) \end{pmatrix}^{\mathrm{T}} = \begin{pmatrix} 3|x|_2 x_1 & 3|x|_2 x_2 \\ -\dfrac{\sin|x|_2}{|x|_2} x_1 & -\dfrac{\sin|x|_2}{|x|_2} x_2 \end{pmatrix},$$

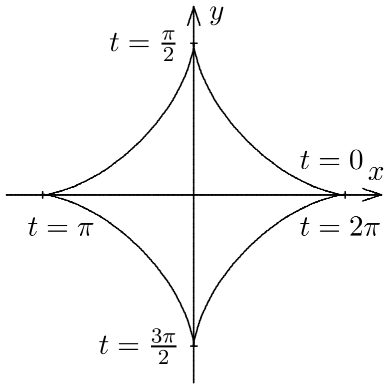
worin man an der ersten Spalte sofort
$$\partial_1 f(x) = \begin{pmatrix} 3|x|_2 x_1 \\ -\dfrac{\sin|x|_2}{|x|_2} x_1 \end{pmatrix}$$

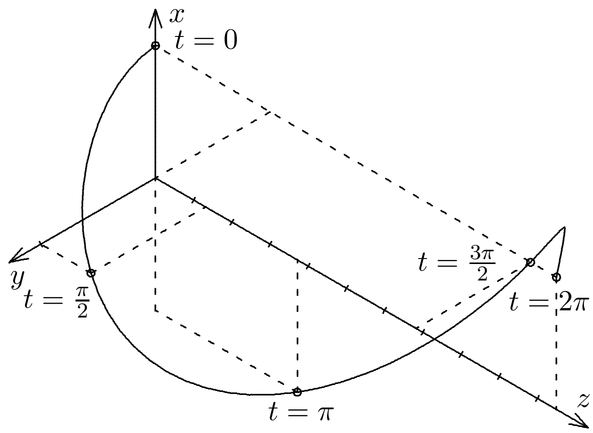
abliest. Die vier partiellen Ableitungen $\partial_i f_k \colon \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}$, $i,k \in \{1,2\}$, sind stetig, d. h. f ist differenzierbar. Damit existiert auch jede Richtungsableitung und lässt sich berechnen durch
$$\partial_v f(x) = J_f(x) \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3|x|_2(2x_1 + x_2) \\ -\dfrac{\sin|x|_2}{|x|_2}(2x_1 + x_2) \end{pmatrix}.$$

(b) Mit der Kettenregel ergibt sich

$$
\begin{aligned}
(f\circ\gamma)'(1) &= J_{f\circ\gamma}(1) = (J_f\circ\gamma)(1)\,J_\gamma(1) = J_f\big(\gamma(1)\big)\,\gamma'(1)\\
&= \begin{pmatrix} 3\sqrt{2} & 3\sqrt{2}\\ -\dfrac{\sin\sqrt{2}}{\sqrt{2}} & -\dfrac{\sin\sqrt{2}}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix} 9\sqrt{2}\\ -\dfrac{3\sin\sqrt{2}}{\sqrt{2}}\end{pmatrix}.
\end{aligned}
$$

21. (a) Diagramme für γ und ϱ:





(b) Mit

$$
\begin{aligned}
\big|\gamma'(t)\big|_2^2 &= (-3\cos^2 t\sin t)^2 + (3\sin^2 t\cos t)^2\\
&= 9\cos^2 t\sin^2 t\cos^2 t + 9\cos^2 t\sin^2 t\sin^2 t = 9\cos^2 t\sin^2 t = \frac{9}{4}\sin^2 2t
\end{aligned}
$$

ergibt sich

$$
\ell(\gamma) = \int_0^{2\pi}\big|\gamma'(t)\big|_2\,\mathrm{d}t = \frac{3}{2}\int_0^{2\pi}|\sin 2t|\,\mathrm{d}t = 6\int_0^{\pi/2}\sin 2t\,\mathrm{d}t = -3\cos 2t\Big|_{t=0}^{t=\pi/2} = 6\,.
$$

Ebenso mit

$$
\big|\varrho'(t)\big|_2^2 = (-\sin t)^2 + (\cos t)^2 + (t^{1/2})^2 = t+1
$$

ergibt sich

$$
\ell(\varrho) = \int_0^{2\pi}\big|\varrho'(t)\big|_2\,\mathrm{d}t = \int_0^{2\pi}(t+1)^{1/2}\,\mathrm{d}t = \frac{2}{3}(t+1)^{3/2}\Big|_{t=0}^{t=2\pi} = \frac{2}{3}\big((2\pi+1)^{3/2}-1\big)\,.
$$

Bemerkung: Um das Ergebnis zu testen, sind elementargeometrische Abschätzungen der Form $4\sqrt{2} < \ell(\gamma) < 8$ sehr hilfreich. Für $\ell(\varrho)$ kann man sich $\ell(\varrho)^2 \approx (2\pi)^2 + (2(2\pi)^{3/2}/3)^2$ überlegen.

(c) Wir setzen $\varphi(x) := |x|_2^2/2$. Dann gilt mit Aufgabe 16. (a) $\nabla\varphi(x) = |x|_2\nabla|x|_2 = x$, d. h. $\nabla\varphi = F$. Damit folgt mit Hilfe von Satz 9.16

$$
\int_\gamma F = \int_\gamma \nabla\varphi = \varphi\big(\gamma(2\pi)\big) - \varphi\big(\gamma(0)\big) = \varphi(1,0) - \varphi(1,0) = 0\,.
$$

Bemerkung: Für diese Rechnung war es nicht einmal notwendig, die Funktion φ zu kennen; ihre Existenz war hinreichend dafür, dass das Integral verschwindet.

Das andere Integral berechnen wir explizit:

$$\begin{aligned}\int_\varrho G &= \int_0^{2\pi} \Big\langle G\big(\varrho(t)\big), \varrho'(t)\Big\rangle\,\mathrm{d}t = \int_0^{2\pi} \left\langle \begin{pmatrix} -\sin t \\ \cos t \\ \frac{2}{3}t^{3/2} \end{pmatrix}, \begin{pmatrix} -\sin t \\ \cos t \\ t^{1/2} \end{pmatrix} \right\rangle \mathrm{d}t \\ &= \int_0^{2\pi} \left(\sin^2 t + \cos^2 t + \frac{2}{3}t^2\right) \mathrm{d}t = \left(t + \frac{2}{9}t^3\right)\Big|_{t=0}^{t=2\pi} = 2\pi\left(1 + \frac{8\pi^2}{9}\right).\end{aligned}$$

22. (a) Für $a, h \in \mathbb{R}^n$ mit $h \neq 0$ berechnen wir

$$\begin{aligned}&\Big|(fg)(a+h) - (fg)(a) - \big(g(a)\mathrm{d}f(a) + f(a)\mathrm{d}g(a)\big)(h)\Big| \\ &\quad= \Big|f(a+h)g(a+h) - f(a)g(a+h) + f(a)g(a+h) - f(a)g(a) \\ &\qquad\quad - g(a+h)\mathrm{d}f(a)(h) + g(a+h)\mathrm{d}f(a)(h) - g(a)\mathrm{d}f(a)(h) - f(a)\mathrm{d}g(a)(h)\Big| \\ &\quad\le \big|f(a+h)g(a+h) - f(a)g(a+h) - g(a+h)\mathrm{d}f(a)(h)\big| \\ &\qquad\quad + \big|f(a)g(a+h) - f(a)g(a) - f(a)\mathrm{d}g(a)(h)\big| \\ &\qquad\quad + \big|g(a+h)\mathrm{d}f(a)(h) - g(a)\mathrm{d}f(a)(h)\big| \\ &\quad= \big|g(a+h)\big|\,\big|f(a+h) - f(a) - \mathrm{d}f(a)(h)\big| \\ &\qquad\quad + \big|f(a)\big|\,\big|g(a+h) - g(a) - \mathrm{d}g(a)(h)\big| + \big|g(a+h) - g(a)\big|\,\big|\mathrm{d}f(a)(h)\big|\end{aligned}$$

und betrachten die drei Summanden nach Division durch $|h|_2$ getrennt: In

$$\big|g(a+h)\big| \frac{|f(a+h) - f(a) - \mathrm{d}f(a)(h)|}{|h|_2}$$

konvergiert der erste Faktor wegen der Stetigkeit von g für $h \to 0$ gegen $|g(a)|$ und der zweite wegen der Differenzierbarkeit von f gegen 0. In

$$\big|f(a)\big| \frac{|g(a+h) - g(a) - \mathrm{d}g(a)(h)|}{|h|_2}$$

ist der erste Faktor konstant und der zweite konvergiert wegen der Differenzierbarkeit von g für $h \to 0$ gegen 0. In

$$\big|g(a+h) - g(a)\big| \frac{|\mathrm{d}f(a)(h)|}{|h|_2} = \big|g(a+h) - g(a)\big| \left|\mathrm{d}f(a)\left(\frac{h}{|h|_2}\right)\right|,$$

wo die Linearität von $\mathrm{d}f(a)$ benutzt wurde, konvergiert der erste Faktor wegen der Stetigkeit von g für $h \to 0$ gegen 0 und der zweite bleibt beschränkt, weil $h/|h|_2$ beschränkt und $\mathrm{d}f(a)$ linear und beschränkt ist.

Damit ist der angesetzte Ausdruck $g(a)\mathrm{d}f(a) + f(a)\mathrm{d}g(a)$ die (eindeutig bestimmte) Ableitung von fg in a, und diese liegt als Linearkombination von $\mathrm{d}f(a)$ und $\mathrm{d}g(a)$ tatsächlich in $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$.

(b) Da f stetig in $a \in \mathbb{R}^n$ mit $f(a) \neq 0$ ist, gibt es ein $r > 0$, so dass $f(x) \neq 0$ für alle $x \in U_r(a)$. Für $h \in \mathbb{R}^n$ mit $0 < |h|_2 < r$ berechnen wir

$$\big|f(a)f(a+h)\big| \left|\frac{1}{f(a+h)} - \frac{1}{f(a)} + \frac{1}{f(a)^2}\mathrm{d}f(a)(h)\right|$$

$$\begin{aligned}
&= \left| f(a+h) - f(a) - \frac{f(a+h)}{f(a)} \mathrm{d}f(a)(h) \right| \\
&\le \left| f(a+h) - f(a) - \mathrm{d}f(a)(h) \right| + \left| \left(1 - \frac{f(a+h)}{f(a)}\right) \mathrm{d}f(a)(h) \right|
\end{aligned}$$

und betrachten die zwei Summanden nach Division durch $|h|_2$ getrennt: Der erste Summand

$$\frac{|f(a+h) - f(a) - \mathrm{d}f(a)(h)|}{|h|_2}$$

konvergiert wegen der Differenzierbarkeit von f für $h \to 0$ gegen 0. In

$$\frac{1}{|h|_2} \left| \left(1 - \frac{f(a+h)}{f(a)}\right) \mathrm{d}f(a)(h) \right| = \left| 1 - \frac{f(a+h)}{f(a)} \right| \left| \mathrm{d}f(a) \left(\frac{h}{|h|_2} \right) \right|,$$

wo die Linearität von $\mathrm{d}f(a)$ benutzt wurde, konvergiert der erste Faktor wegen der Stetigkeit von f für $h \to 0$ gegen 0 und der zweite bleibt beschränkt, weil $h/|h|_2$ beschränkt und $\mathrm{d}f(a)$ linear und beschränkt ist.

Damit ist der angesetzte Ausdruck $-f(a)^{-2}\mathrm{d}f(a)$ die (eindeutig bestimmte) Ableitung von $1/f$ in a, und diese liegt als Vielfaches von $\mathrm{d}f(a)$ tatsächlich in $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$.

Bemerkung: Es genügt nicht, $f \cdot 1/f = 1$ unter Verwendung der Produktregel aus (a) abzuleiten. Das ist nur eine formale Rechnung, die nicht zeigt, dass $1/f$ differenzierbar ist.

(c) *Achtung: Die folgende Rechnung setzt $f \in \mathcal{C}^1$ voraus. Bitte um Rückmeldung, ob es anders geht oder ein Fehler auf dem Aufgabenblatt war!* Seien $x \in \mathbb{R}^n$ mit $x \neq 0$ und $\gamma\colon [0,1] \to \mathbb{R}^n$ mit $\gamma(t) = tx$ ein Weg. Dann ist $\gamma(0) = 0$, $\gamma(1) = x$ und $\gamma'(t) = x$, d. h. mit Satz 9.16 folgt

$$\begin{aligned}
\left| f(x) - f(0) \right| &= \left| \int_\gamma \nabla f \right| = \left| \int_0^1 \langle \nabla f(tx), x \rangle \,\mathrm{d}t \right| \le \int_0^1 \left| \langle \nabla f(tx), x \rangle \right| \mathrm{d}t \\
&\le \int_0^1 \left| \nabla f(tx) \right|_2 |x|_2 \,\mathrm{d}t \le |x|_2 \int_0^1 \mathrm{e}^{-\alpha t |x|_2} \,\mathrm{d}t = -\frac{\mathrm{e}^{-\alpha t |x|_2}}{\alpha} \bigg|_{t=0}^{t=1} \\
&= \frac{1 - \mathrm{e}^{-\alpha |x|_2}}{\alpha} \le \frac{1}{\alpha}.
\end{aligned}$$

Damit ergibt sich

$$\left| f(x) \right| = \left| f(x) - f(0) + f(0) \right| \le \left| f(x) - f(0) \right| + \left| f(0) \right| \le \frac{1}{\alpha} + \left| f(0) \right|,$$

d. h. f ist beschränkt.

23. (a) Für $(x, y) \neq (0,0)$ ist f klarerweise differenzierbar, und es gilt

$$\partial_1 f(x,y) = \frac{(3x^2y - y^3)(x^2 + y^2) - (x^3y - xy^3) \cdot 2x}{(x^2+y^2)^2} = \frac{x^4y + 4x^2y^3 - y^5}{(x^2+y^2)^2}.$$

Wegen $f(y,x) = -f(x,y)$ erhält man $\partial_2 f$ aus $\partial_1 f$ durch Vertauschen von x mit y und Hinzufügen eines Minuszeichens, also

$$\partial_2 f(x,y) = \frac{x^5 - 4x^3y^2 - xy^4}{(x^2+y^2)^2}.$$

Die Ableitungen in $(0,0)^{\mathrm{T}}$ müssen separat berechnet werden. Für $t \neq 0$ ist

$$\frac{f(0+t,0)-f(0,0)}{t} = \frac{0}{t} = 0 \quad \text{und} \quad \frac{f(0,0+t)-f(0,0)}{t} = \frac{0}{t} = 0\,,$$

d. h. für $t \to 0$ folgt $\partial_1 f(0,0) = \partial_2 f(0,0) = 0$. Also ist f in jedem Punkt partiell differenzierbar. Es fehlt noch die Stetigkeit der Ableitungen. Wir betrachten dazu

$$\begin{aligned}\big|\partial_1 f(x,y) - \partial_1 f(0,0)\big| \le \frac{|x^4y| + |4x^2y^3| + |y^5|}{(x^2+y^2)^2} &\le \frac{2x^4+4x^2y^2+2y^4}{(x^2+y^2)^2}|y| \\ &= 2|y| \xrightarrow{y\to 0} 0\,,\end{aligned}$$

d. h. $\partial_1 f$ ist stetig auf ganz $\mathbb{R}^2$. (Außerhalb des Ursprungs ist die Stetigkeit sowieso klar.) Die Stetigkeit von $\partial_2 f$ folgt daraus wegen der oben genannten Symmetrie, d. h. $f \in \mathcal{C}^1(\mathbb{R}^2)$.

(b) Wir rechnen nach, dass

$$\frac{\partial_1 f(0,0+t) - \partial_1 f(0,0)}{t} = \frac{-t^5}{(t^2)^2 t} = -1 \xrightarrow{t\to 0} -1 = \partial_2\partial_1 f(0,0)$$

und

$$\frac{\partial_2 f(0+t,0) - \partial_2 f(0,0)}{t} = \frac{t^5}{(t^2)^2 t} = 1 \xrightarrow{t\to 0} 1 = \partial_1\partial_2 f(0,0)\,.$$

Ablesen ergibt $\partial_2\partial_1 f(0,0) \neq \partial_1\partial_2 f(0,0)$.

24. (a) Es ist

$$\partial_2 F_1(x) = \frac{-(x_1^2+x_2^2) + x_2\cdot 2x_2}{(x_1^2+x_2^2)^2} = \frac{-x_1^2+x_2^2}{(x_1^2+x_2^2)^2}$$

und

$$\partial_1 F_2(x) = \frac{(x_1^2+x_2^2) - x_1\cdot 2x_1}{(x_1^2+x_2^2)^2} = \frac{-x_1^2+x_2^2}{(x_1^2+x_2^2)^2}\,,$$

also $\partial_2 F_1(x) = \partial_1 F_2(x)$.

(b) Es ist

$$\int_\gamma F = \int_0^{2\pi} \Big\langle F\big(\gamma(t)\big), \gamma'(t)\Big\rangle\,\mathrm{d}t = \int_0^{2\pi} \frac{-\sin t(-\sin t) + \cos t\cos t}{\cos^2 t + \sin^2 t}\,\mathrm{d}t = 2\pi\,.$$

Wenn F ein Potential besäße, dann müsste dieses geschlossene Kurvenintegral 0 sein.

(c) Es ist

$$\partial_1 f = \frac{\partial}{\partial x_1}\arctan\frac{x_2}{x_1} = \frac{-x_2/x_1^2}{1+(x_2/x_1)^2} = \frac{-x_2}{x_1^2+x_2^2}$$

und

$$\partial_2 f = \frac{\partial}{\partial x_2}\arctan\frac{x_2}{x_1} = \frac{1/x_1}{1+(x_2/x_1)^2} = \frac{x_1}{x_1^2+x_2^2}\,,$$

d. h. $\nabla f = F|_U$.

(d) Wir zeigen, dass U^* sternförmig zu $a = (1,0)^{\mathrm{T}} \in U^*$ ist. Für einen beliebigen Punkt $x \in U^*$ unterscheiden wir die drei Fälle $x_2 > 0$, $x_2 = 0$ und $x_2 < 0$. $x_2 = 0$ ist klar, denn die Verbindungsstrecke liegt dann auf der x_1-Achse, und der dritte Fall ergibt sich durch Symmetrie aus dem ersten. Sei also nun $x_2 > 0$. Dann beschreibt $\gamma(t) := (1-t)a + tx$, $t \in [0,1]$, die Verbindungsstrecke von a und x. Wegen $a_2 = 0$ und $x_2 > 0$ ist aber $\gamma_2(t) > 0$ für $t > 0$, also $\gamma(t) \in U^*$, d. h. U^* ist sternförmig.

25. (a) $(\bar{x}, \bar{y})^{\mathrm{T}} \in \mathbb{R}^2$ ist genau dann ein kritischer Punkt, wenn

$$\partial_1 f(\bar{x},\bar{y}) = (\bar{x} - 2\bar{y}^2) + (\bar{x} - \bar{y}^2) = 2\bar{x} - 3\bar{y}^2 = 0 \quad \Longrightarrow \quad \bar{x} = \frac{3}{2}\bar{y}^2$$

und

$$\partial_2 f(\bar{x},\bar{y}) = -2\bar{y}(\bar{x} - 2\bar{y}^2) + (\bar{x} - \bar{y}^2)(-4\bar{y}) = -6\bar{x}\bar{y} + 8\bar{y}^3 = -\bar{y}^3 = 0\,,$$

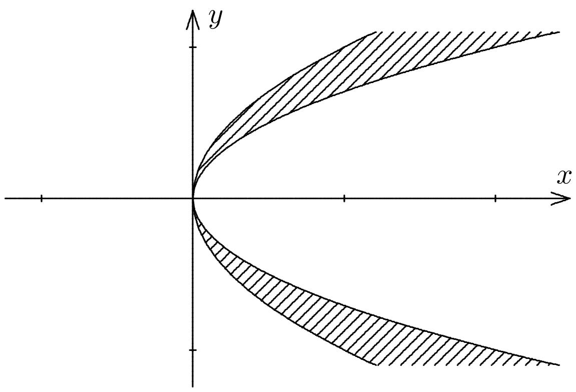
was nur durch $(\bar{x},\bar{y}) = (0,0)$ gelöst wird. Es gilt

$$H_f(0,0) = \begin{pmatrix} 2 & -6y \\ -6y & -6x + 24y^2 \end{pmatrix}\Bigg|_{x=y=0} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

(b) $f(x,y)$ ist genau dann negativ, wenn die beiden Faktoren $x - y^2$ und $x - 2y^2$ verschiedenes Vorzeichen haben. Wir unterscheiden

- $x - y^2 > 0$ und $x - 2y^2 < 0$, d. h. $y^2 < x < 2y^2$;
- $x - y^2 < 0$ und $x - 2y^2 > 0$, d. h. $2y^2 < x < y^2$ (unlösbar).

Also folgt:



(c) Es ist

$$\varphi(t) = \big(ta - (tb)^2\big)\big(ta - 2(tb)^2\big) = 2b^4t^4 - 3ab^2t^3 + a^2t^2\,,$$

d. h. 0 ist eine Nullstelle von $\varphi'(t) = 8b^4t^3 - 9ab^2t^2 + 2a^2t$. Weiter rechnen wir $\varphi''(t) = 24b^4t^2 - 18ab^2t + 2a^2$, also $\varphi''(0) = 2a^2$. Ist $a \neq 0$, so ist 0 tatsächlich eine Minimalstelle. Im Falle $a = 0$ hingegen ergibt sich $\varphi''(t) = 24b^4t^2$, $\varphi'''(t) = 48b^4t$ und $\varphi^{(4)}(t) = 48b^4$. Dann ist $\varphi''(0) = \varphi'''(0) = 0$ und $\varphi^{(4)}(0) > 0$, da $b \neq 0$. Also hat φ auch hier in 0 ein Minimum.

(d) Wir müssen zeigen, dass es in jeder Umgebung um $(0,0)^{\mathrm{T}}$ Punkte gibt, in denen f negativ ist. Betrachten wir einen Punkt $(x,y)^{\mathrm{T}}$, für den $x = 3y^2/2 \neq 0$ gilt, so ist

$$f(x,y) = \Big(\frac{3}{2}y^2 - y^2\Big)\Big(\frac{3}{2}y^2 - 2y^2\Big) = -\frac{1}{4}y^4 < 0\,.$$

Da jede Umgebung solch einen Punkt enthält, hat f kein lokales Minimum in $(0,0)^{\mathrm{T}}$.

26. Für $t \neq 0$ ist

$$\begin{aligned}\frac{f(x+te_i)-f(x)}{t} &= \frac{\langle A(x+te_i), x+te_i\rangle - \langle Ax, x\rangle}{t} \\ &= \frac{\langle Ax, te_i\rangle + \langle Ate_i, x\rangle + \langle Ate_i, te_i\rangle}{t} = (Ax)_i + \langle Ae_i, x\rangle + (Ae_i)_i t\,,\end{aligned}$$

wobei

$$\langle Ae_i, x\rangle = \sum_{k=1}^n (Ae_i)_k x_k = \sum_{k=1}^n \sum_{j=1}^n a_{kj}(e_i)_j x_k = \sum_{k=1}^n a_{ki} x_k = (A^{\mathrm{T}}x)_i\,.$$

Für $t \to 0$ folgt also $\partial_i f(x) = (Ax)_i + (A^{\mathrm{T}}x)_i$ und damit $\nabla f(x) = (A + A^{\mathrm{T}})x$. Für die zweiten Ableitungen betrachten wir

$$\begin{aligned}\frac{\partial_i f(x+te_k) - \partial_i f(x)}{t} &= \frac{((A+A^{\mathrm{T}})(x+te_k))_i - ((A+A^{\mathrm{T}})x)_i}{t} \\ &= \big((A+A^{\mathrm{T}})e_k\big)_i = \sum_{j=1}^n (a_{ij}+a_{ji})(e_k)_j = a_{ik} + a_{ki}\,.\end{aligned}$$

Da dieser Ausdruck nicht mehr von t abhängt, existiert insbesondere der Grenzwert $t \to 0$, und $a_{ik} + a_{ki}$ sind die Einträge der Hessematrix, d. h. $H_f(x) = A + A^{\mathrm{T}}$. Alle Einträge sind konstant und daher stetig, was $f \in \mathcal{C}^2(\mathbb{R}^n)$ beweist. Für den Laplace-Operator gilt schließlich

$$\Delta f(x) = \operatorname{spur}\big(H_f(x)\big) = 2 \operatorname{spur} A\,.$$

27. (a) Für einen kritischen Punkt $(\bar{x}, \bar{y})^{\mathrm{T}} \in \mathbb{R}^2$ muss

$$\partial_1 f(\bar{x},\bar{y}) = \cos\bar{x}\sin\bar{y} = 0 \quad \text{und} \quad \partial_2 f(\bar{x},\bar{y}) = \sin\bar{x}\cos\bar{y} = 0$$

gelten. Wir unterscheiden die beiden Fälle

- $\cos\bar{x} = 0$, damit $\sin\bar{x} = \pm 1$, daher also $\cos\bar{y} = 0$;
- $\sin\bar{y} = 0$, damit $\cos\bar{y} = \pm 1$, daher also $\sin\bar{x} = 0$.

Allgemein gilt für die Hessematrix

$$H_f(x,y) = \begin{pmatrix} -\sin x \sin y & \cos x \cos y \\ \cos x \cos y & -\sin x \sin y \end{pmatrix}.$$

Im ersten Fall ist $\bar{x} = (k+1/2)\pi$ und $\bar{y} = (\ell + 1/2)\pi$ für $k, \ell \in \mathbb{Z}$. Damit folgt

$$H_f(\bar{x},\bar{y}) = \begin{pmatrix} -(-1)^k(-1)^\ell & 0 \\ 0 & -(-1)^k(-1)^\ell \end{pmatrix} = (-1)^{k+\ell+1}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Diese Matrix ist definit, und zwar positiv, wenn $k+\ell$ ungerade, und negativ, wenn $k+\ell$ gerade ist. f besitzt also

- strikte lokale Maxima in $((2k+1/2)\pi,(2\ell+1/2)\pi)^{\mathrm{T}}$ und $((2k-1/2)\pi,(2\ell-1/2)\pi)^{\mathrm{T}}$ für $k,\ell\in\mathbb{Z}$;
- strikte lokale Minima in $((2k+1/2)\pi,(2\ell-1/2)\pi)^{\mathrm{T}}$ und $((2k-1/2)\pi,(2\ell+1/2)\pi)^{\mathrm{T}}$ für $k,\ell\in\mathbb{Z}$.

Im zweiten Fall ist $\bar{x}=k\pi$ und $\bar{y}=\ell\pi$ für $k,\ell\in\mathbb{Z}$. Damit folgt

$$H_f(\bar{x},\bar{y})=\begin{pmatrix}0 & (-1)^k(-1)^\ell\\ (-1)^k(-1)^\ell & 0\end{pmatrix}=(-1)^{k+\ell}\begin{pmatrix}0&1\\1&0\end{pmatrix}.$$

Diese Matrix ist indefinit für alle $k,\ell\in\mathbb{Z}$, denn die quadratische Form

$$\begin{pmatrix}\xi\\ \eta\end{pmatrix}^{\mathrm{T}} H_f(\bar{x},\bar{y})\begin{pmatrix}\xi\\ \eta\end{pmatrix}=2(-1)^{k+\ell}\xi\eta$$

kann jedes Vorzeichen annehmen. f besitzt also

- Sattelpunkte in $(k\pi,\ell\pi)^{\mathrm{T}}$ für alle $k,\ell\in\mathbb{Z}$.

(b) Die Formel für das Taylorpolynom dritter Ordnung ist

$$T_p^3 f(x,y)=\sum_{|\alpha|\le 3}\frac{\partial^\alpha f(p)}{\alpha!}\left(\begin{pmatrix}x\\y\end{pmatrix}-p\right)^\alpha .$$

Ist α_1 gerade, so ist $\partial^\alpha f(x,y)$ ein Vielfaches von $\sin x$, also $\partial^\alpha f(p)=0$. Für die anderen Multiindizes ergibt sich

$$\begin{aligned}
\alpha&=\begin{pmatrix}1\\0\end{pmatrix}: & \frac{\partial^\alpha f(p)}{\alpha!}&=\frac{1}{1!\,0!}\cos 0\sin\frac{\pi}{4}=\frac{1}{\sqrt{2}},\\
\alpha&=\begin{pmatrix}1\\1\end{pmatrix}: & \frac{\partial^\alpha f(p)}{\alpha!}&=\frac{1}{1!\,1!}\cos 0\cos\frac{\pi}{4}=\frac{1}{\sqrt{2}},\\
\alpha&=\begin{pmatrix}3\\0\end{pmatrix}: & \frac{\partial^\alpha f(p)}{\alpha!}&=\frac{1}{3!\,0!}(-\cos 0)\sin\frac{\pi}{4}=-\frac{1}{6\sqrt{2}},\\
\alpha&=\begin{pmatrix}1\\2\end{pmatrix}: & \frac{\partial^\alpha f(p)}{\alpha!}&=\frac{1}{1!\,2!}\cos 0\Big(-\sin\frac{\pi}{4}\Big)=-\frac{1}{2\sqrt{2}}.
\end{aligned}$$

Damit folgt

$$\begin{aligned}
T_p^3 f(x,y)&=\frac{1}{\sqrt{2}}\left(x+x\Big(y-\frac{\pi}{4}\Big)-\frac{1}{6}x^3-\frac{1}{2}x\Big(y-\frac{\pi}{4}\Big)^2\right)\\
&=\frac{1}{\sqrt{2}}\left(\Big(1-\frac{\pi}{4}-\frac{\pi^2}{32}\Big)x+\Big(1+\frac{\pi}{4}\Big)xy-\frac{1}{6}x^3-\frac{1}{2}xy^2\right).
\end{aligned}$$

28. (a) Für jedes $n\in\mathbb{N}$ definieren wir $K_n:=M\cap B_n(0)$ und $g_n:=f|_{K_n}$. Dann ist jedes K_n als Schnitt einer abgeschlossenen mit einer kompakten Menge kompakt (siehe Zusatz von Aufgabe 19. (a)), und es gibt einen Index $n_0\in\mathbb{N}$, so dass $K_n\neq\emptyset$ für alle $n\ge n_0$. Für alle diese $n\ge n_0$ ist g_n stetig auf K_n und nimmt daher ihr Maximum auf K_n an, d. h. es gibt $a_n\in K_n$ mit $f(a_n)=\max\{f(x)\mid x\in K_n\}$. So entsteht eine Folge $(a_n)_{n\ge n_0}$ in M, und wegen $K_{n+1}\supset K_n$ ist $(f(a_n))_{n\ge n_0}$ monoton wachsend.

Wir wollen nun die Aussage in ii. zeigen, d. h. wir nehmen an, dass $\sup\{f(x) \mid x \in M\} < \infty$. Also ist $(f(a_n))_{n \geq n_0}$ nach oben beschränkt und wegen der Monotonie nach unten beschränkt, also beschränkt. Damit folgt erstens, dass $(f(a_n))_{n\geq n_0}$ konvergent ist, und zweitens, dass wegen der allgemeinen Voraussetzung $(a_n)_{n\geq n_0}$ beschränkt ist. Damit besitzt $(a_n)_{n\geq n_0}$ nach dem Satz von Bolzano-Weierstraß eine konvergente Teilfolge $(a_{n_k})_k$ (o. B. d. A. $n_1 \geq n_0$), die für $k \to \infty$ gegen ein $a \in A$ konvergiert. Wegen der Stetigkeit von f konvergiert auch $f(a_{n_k})$ gegen $f(a)$, wobei $f(a) = \lim_{n\to\infty} f(a_n)$. Also nimmt f an der Stelle a ein globales Maximum, nämlich $f(a)$, an.

(b) Es ist

$$\nabla f(x) = \sum_{i=1}^{k} \nabla |x - a^i|_2^2 = \sum_{i=1}^{k} 2(x - a^i) = 2\left(kx - \sum_{i=1}^{k} a^i\right),$$

d. h. $\nabla f(\bar{x}) = 0$ genau dann, wenn

$$\bar{x} = \frac{1}{k}\sum_{i=1}^{k} a^i .$$

Nun betrachten wir für $v \in \mathbb{R}^n$

$$\begin{aligned} f(\bar{x}+v) - f(\bar{x}) &= \sum_{i=1}^{k} \left|\frac{1}{k}\sum_{j=1}^{k} a^j + v - a^i\right|_2^2 - \sum_{i=1}^{k} \left|\frac{1}{k}\sum_{j=1}^{k} a^j - a^i\right|_2^2 \\ &= \sum_{i=1}^{k}\left(|v|_2^2 + 2\left\langle \frac{1}{k}\sum_{j=1}^{k} a^j - a^i, v\right\rangle\right) \\ &= k|v|_2^2 + \frac{2}{k}\sum_{i=1}^{k}\sum_{j=1}^{k}\langle a^j, v\rangle - 2\sum_{i=1}^{k}\langle a^i, v\rangle = k|v|_2^2 , \end{aligned}$$

d. h. $f(\bar{x}+v) > f(\bar{x})$ für alle $v \neq 0$, also hat f bei $\bar{x}$ ein globales Minimum.

29. (a) Wir müssen $f(x) = y$, d. h.

$$x_1^2 - x_2^2 = -1 , \quad 2x_1x_2 = 0 ,$$

nach x auflösen. Durch Fallunterscheidung der zweiten Gleichung erhalten wir

- $x_1 = 0$, also $-x_2^2 = -1$, d. h. $x_2 = 1$ oder $x_2 = -1$;
- $x_2 = 0$, also $x_1^2 = -1$, was unlösbar ist.

y besitzt also genau die beiden Urbilder $\hat{x} = (0,1)^{\mathrm{T}}$ und $\check{x} = (0,-1)^{\mathrm{T}}$ unter f. (Mit der Bemerkung erhält man das noch schneller, da $\hat{x}$ und $\check{x}$ die beiden komplexen Quadratwurzeln von -1 sein müssen.) Die Jacobimatrix von f lautet

$$J_f(x) = \begin{pmatrix} 2x_1 & -2x_2 \\ 2x_2 & 2x_1 \end{pmatrix}$$

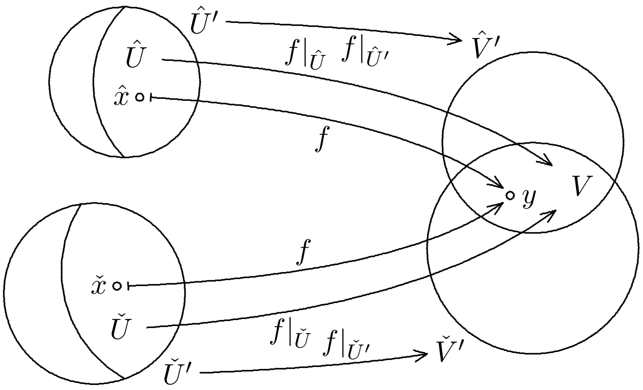
mit der Determinante $\det J_f(x) = 4(x_1^2 + x_2^2) \neq 0$ für $x \in \mathbb{R}^2 \setminus \{0\}$. Nach dem Satz über lokale Diffeomorphismen gibt es dann offene Umgebungen $\hat{U}'$ um $\hat{x}$, $\check{U}'$ um $\check{x}$ und $\hat{V}', \check{V}'$ um y, so dass

$$f|_{\hat{U}'} \in \mathrm{Diff}(\hat{U}', \hat{V}') , \quad f|_{\check{U}'} \in \mathrm{Diff}(\check{U}', \check{V}') .$$

Dann ist aber auch $V := \hat{V}' \cap \check{V}'$ eine offene Umgebung von y, und es gilt

$$f|_{\hat{U}} \in \mathrm{Diff}(\hat{U}, V)\,, \quad f|_{\check{U}} \in \mathrm{Diff}(\check{U}, V)$$

mit $\hat{U} = \hat{U}' \cap f^{-1}(V)$ und $\check{U} = \check{U}' \cap f^{-1}(V)$.



(b) Nach dem Satz über lokale Diffeomorphismen ist die Jacobimatrix der inversen Funktion gleich der Inversen der Jacobimatrix, d. h. wir erhalten

$$J_\varphi(y) = \big(J_{f|_{\hat{U}}}(\hat{x})\big)^{-1} = \begin{pmatrix} 2\hat{x}_1 & -2\hat{x}_2 \\ 2\hat{x}_2 & 2\hat{x}_1 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix}^{-1} = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

und

$$J_\psi(y) = \big(J_{f|_{\check{U}}}(\check{x})\big)^{-1} = \begin{pmatrix} 2\check{x}_1 & -2\check{x}_2 \\ 2\check{x}_2 & 2\check{x}_1 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}^{-1} = \frac{1}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

30. Für jedes $i \in \{1, \dots, k\}$ gilt: Nach dem Satz über lokale Diffeomorphismen gibt es eine offene Umgebung $\bar{U}^i \subset U$ von x^i und eine offene Umgebung $\bar{V}^i \subset \mathbb{R}^n$ von y, so dass $f|_{\bar{U}^i} \in \mathrm{Diff}(\bar{U}^i, \bar{V}^i)$. Nun definieren wir

$$\hat{U}^i := \bar{U}^i \cap \bigcap_{j=1,\, j\neq i}^{k} U_{d_2(x^j, x^i)/2}(x^i)\,.$$

Dann ist $\hat{U}^i$ eine offene Umgebung von x^i mit $\hat{U}^i \subset \bar{U}^i$, d. h. mit $\hat{V}^i := \mathrm{Bild}\, f|_{\hat{U}^i} \subset \bar{V}^i$ ist $f|_{\hat{U}^i} \in \mathrm{Diff}(\hat{U}^i, \hat{V}^i)$.

Die so konstruierten $\hat{U}^i$ sind nun paarweise disjunkt. Die Menge $V := \bigcap_{i=1}^k \hat{V}_i$ ist offen und eine Umgebung von y.

Für jedes $i \in \{1, \dots, k\}$ gilt weiter: Die Menge $U_i := \hat{U}_i \cap f^{-1}(V)$ ist eine offene Umgebung von x^i, und es gilt $f|_{U^i} \in \mathrm{Diff}(U^i, V)$. Da ein Diffeomorphismus bijektiv ist, gibt es zu jedem $\tilde{y} \in V$ genau ein $\tilde{x}^i \in U^i$ mit $f(\tilde{x}^i) = \tilde{y}$.

Da die U^i paarweise disjunkt sind, sind die $\tilde{x}^i$ paarweise verschieden, d. h. jedes $\tilde{y} \in V$ besitzt mindestens k Urbilder in U unter f.

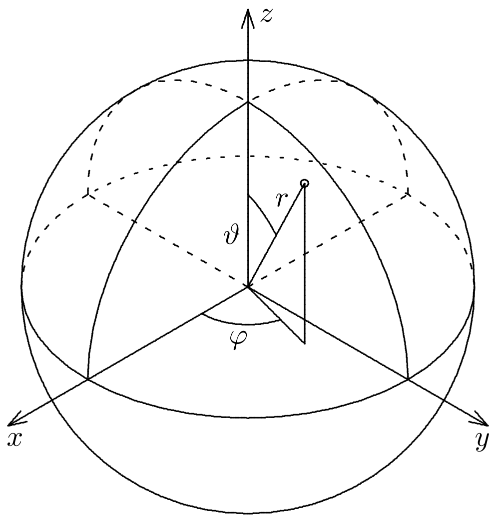
31. (b) Es ist

$$J_{P_3}(r,\varphi,\vartheta) = \begin{pmatrix} \dfrac{\partial(r\cos\varphi\sin\vartheta)}{\partial r} & \dfrac{\partial(r\cos\varphi\sin\vartheta)}{\partial\varphi} & \dfrac{\partial(r\cos\varphi\sin\vartheta)}{\partial\vartheta} \\ \dfrac{\partial(r\sin\varphi\sin\vartheta)}{\partial r} & \dfrac{\partial(r\sin\varphi\sin\vartheta)}{\partial\varphi} & \dfrac{\partial(r\sin\varphi\sin\vartheta)}{\partial\vartheta} \\ \dfrac{\partial(r\cos\vartheta)}{\partial r} & \dfrac{\partial(r\cos\vartheta)}{\partial\varphi} & \dfrac{\partial(r\cos\vartheta)}{\partial\vartheta} \end{pmatrix}$$
$$= \begin{pmatrix} \cos\varphi\sin\vartheta & -r\sin\varphi\sin\vartheta & r\cos\varphi\cos\vartheta \\ \sin\varphi\sin\vartheta & r\cos\varphi\sin\vartheta & r\sin\varphi\cos\vartheta \\ \cos\vartheta & 0 & -r\sin\vartheta \end{pmatrix}$$

und damit

$$\det J_{P_3}(r,\varphi,\vartheta) = r^2\sin\vartheta \begin{vmatrix} \cos\varphi\sin\vartheta & -\sin\varphi & \cos\varphi\cos\vartheta \\ \sin\varphi\sin\vartheta & \cos\varphi & \sin\varphi\cos\vartheta \\ \cos\vartheta & 0 & -\sin\vartheta \end{vmatrix}$$
$$= r^2\sin\vartheta\left(\cos\vartheta\begin{vmatrix} -\sin\varphi & \cos\varphi\cos\vartheta \\ \cos\varphi & \sin\varphi\cos\vartheta \end{vmatrix} - \sin\vartheta\begin{vmatrix} \cos\varphi\sin\vartheta & -\sin\varphi \\ \sin\varphi\sin\vartheta & \cos\varphi \end{vmatrix}\right)$$
$$= r^2\sin\vartheta\big(\cos\vartheta(-\cos\vartheta) - \sin\vartheta\sin\vartheta\big) = -r^2\sin\vartheta\,.$$

(a) Es ist $|P_3(r,\varphi,\vartheta)|_2 = r$, d. h. alle Punkte mit festem r liegen auf der Sphäre $S_r(0)$. Verändert man φ von $-\pi$ bis π, während man $\vartheta \in (0,\pi)$ festhält, so durchläuft der Punkt einen Breitengrad. Verändert man dagegen ϑ von 0 bis π, während man $\varphi \in (-\pi,\pi)$ festhält, so durchläuft der Punkt einen Meridian (halben Längengrad), und zwar vom Nord- zum Südpol.



(c) Nach (a) erhalten wir für festes r einen Teil der Sphäre $S_r(0)$. Wegen $\vartheta \neq 0$ und $\vartheta \neq \pi$ liegen aber der Nord- und Südpol nicht im Bild von $P_3(r,\cdot,\cdot)$. Wegen $\varphi \neq -\pi$ und $\varphi \neq \pi$ liegt aber auch der gesamte 180°-Meridian nicht im Bild, es folgt also

$$\operatorname{Bild} P_3(r,\cdot,\cdot) = S_r(0) \setminus \big\{(x,0,z)^{\mathrm{T}} \;\big|\; x \le 0\,,\ z \in \mathbb{R}\big\}\,.$$

Vereinigt man nun über alle $r \in (0,\infty)$, so erhält man

$$V = \operatorname{Bild} P_3 = \mathbb{R}^3 \setminus \big\{(x,0,z)^{\mathrm{T}} \;\big|\; x \le 0\,,\ z \in \mathbb{R}\big\}\,.$$

(Man beachte, dass zwar $V \neq \mathbb{R}^3$, aber dass $\bar{V} = \mathbb{R}^3$.) Klar ist $P_3 \in \mathcal{C}^1(U, \mathbb{R}^3)$ und $P_3 \colon U \to V$ surjektiv; so wurde gerade V konstruiert. Wir zeigen weiter, dass P_3 injektiv ist: Die oben gebildete Vereinigung über r ist disjunkt, d. h. $\operatorname{Bild} P_3(r, \cdot, \cdot) \cap \operatorname{Bild} P_3(r', \cdot, \cdot) = \emptyset$, falls $r \neq r'$. Es gelte nun $P_3(r, \varphi, \vartheta) = P_3(r, \varphi', \vartheta')$, d. h.

$$\cos\varphi \sin\vartheta = \cos\varphi' \sin\vartheta' , \quad \sin\varphi \sin\vartheta = \sin\varphi' \sin\vartheta' , \quad \cos\vartheta = \cos\vartheta' .$$

Der Cosinus ist auf $(0, \pi)$ injektiv, weswegen $\vartheta = \vartheta'$ folgt. Wegen $\sin\vartheta \neq 0$ erhalten wir daraus $\cos\varphi = \cos\varphi'$ und $\sin\varphi = \sin\varphi'$. Die erste Gleichung hat neben $\varphi = \varphi'$ noch die Lösung $\varphi = -\varphi'$, eingesetzt in die zweite ergibt sich jedoch ein Widerspruch, was die Injektivität von P_3 zeigt. Damit existiert die inverse Abbildung $P_3^{-1} \colon V \to U$, und deren stetige Differenzierbarkeit folgt sofort aus $\det J_{P_3}(r, \varphi, \vartheta) = -r^2 \sin\vartheta \neq 0$ von (b). Damit ist $P_3 \in \operatorname{Diff}(U, V)$ gezeigt.

Bemerkung: Es kann passieren, dass $f \in \mathcal{C}^\infty(U, V)$ bijektiv ist, aber dass f^{-1} nicht differenzierbar ist, z. B. $f \colon \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^3$. Es ist $f^{-1}(y) = (\operatorname{sgn} y) \sqrt[3]{|y|}$, aber diese Funktion ist nicht differenzierbar in $y = 0$. Ist jedoch f^{-1} zumindest einmal stetig differenzierbar, so ist sie automatisch genauso oft differenzierbar wie f selbst.

(d) Aus $b = (2, 0, 0)^{\mathrm{T}} = P_3(r, \varphi, \vartheta)$ ergibt sich $r = 2$, $\vartheta = \pi/2$ und $\varphi = 0$ und damit

$$J_{P_3}\Big(2, 0, \frac{\pi}{2}\Big) = \begin{pmatrix} 1 \cdot 1 & -2 \cdot 0 \cdot 1 & 2 \cdot 1 \cdot 0 \\ 0 \cdot 1 & 2 \cdot 1 \cdot 1 & 2 \cdot 0 \cdot 0 \\ 0 & 0 & -2 \cdot 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix} .$$

Mit dem Satz über lokale Diffeomorphismen folgt dann

$$J_{P_3^{-1}}(b) = \Big(J_{P_3}\big(P_3^{-1}(b)\big)\Big)^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix}^{-1} = \frac{1}{2} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} .$$

32. (a) i. Um die Notation aus dem Satz über implizite Funktionen zu verwenden, definieren wir die Funktion $F \colon \mathbb{R}^4 \to \mathbb{R}^2$ durch

$$F(u, v, x, y) = \begin{pmatrix} 3x^2 - y^2 + u^3 + v^2 - 4 \\ 2x^2 - y^3 + u^2 - v - 1 \end{pmatrix}$$

mit der Jacobimatrix

$$J_F(u, v, x, y) = \begin{pmatrix} 3u^2 & 2v & 6x & -2y \\ 2u & -1 & 4x & -3y^2 \end{pmatrix} .$$

Dies entspricht $k = m = 2$ und $n = 4$ aus dem Satz; $(u, v)^{\mathrm{T}}$ übernimmt die Rolle von x, $(x, y)^{\mathrm{T}}$ die von y. Es ist $F(1, 1, 1, 1) = (0, 0)^{\mathrm{T}}$, und aus

$$J_F(1, 1, 1, 1) = \begin{pmatrix} 3 & 2 & 6 & -2 \\ 2 & -1 & 4 & -3 \end{pmatrix}$$

ergibt sich

$$\det\Big(\frac{\partial F}{\partial(x, y)}(1, 1, 1, 1)\Big) = \begin{vmatrix} 6 & -2 \\ 4 & -3 \end{vmatrix} = -10 \neq 0 .$$

Nach dem Satz existieren dann genau die behaupteten Umgebungen. Für die Jacobimatrix von g in $(1,1)^{\mathrm{T}}$ folgt

$$\begin{aligned} J_g(1,1) &= -\left(\frac{\partial F}{\partial(x,y)}(1,1,1,1)\right)^{-1} \frac{\partial F}{\partial(u,v)}(1,1,1,1) \\ &= -\begin{pmatrix} 6 & -2 \\ 4 & -3 \end{pmatrix}^{-1} \begin{pmatrix} 3 & 2 \\ 2 & -1 \end{pmatrix} = -\frac{1}{10}\begin{pmatrix} 5 & 8 \\ 0 & 14 \end{pmatrix}. \end{aligned}$$

ii. Schreibt man die Variablen von F in der Reihenfolge y, v, x, u, so erhält man

$$J_F(y,v,x,u) = \begin{pmatrix} -2y & 2v & 6x & 3u^2 \\ -3y^2 & -1 & 4x & 2u \end{pmatrix}.$$

Es ist

$$\det\left(\frac{\partial F}{\partial(x,u)}(1,1,1,1)\right) = \begin{vmatrix} 6 & 3 \\ 4 & 2 \end{vmatrix} = 0\,;$$

der Satz über implizite Funktionen kann also nicht angewendet werden.

(b) Mit den Bezeichnungen aus dem Satz über implizite Funktionen ist hier $k = 1$, $m = 2$ und $n = 3$; x übernimmt die Rolle von x, $(y,z)^{\mathrm{T}}$ die von y. Zunächst ist $F(0,0,0) = (0,0)^{\mathrm{T}}$. Mit

$$J_F(x,y,z) = \begin{pmatrix} 2x + 2\cos y & -2x\sin y & \cos z \\ 5x^4 + 2\cos z & \cos y & -2x\sin z \end{pmatrix}$$

folgt

$$\det\left(\frac{\partial F}{\partial(y,z)}(0,0,0)\right) = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1 \neq 0\,,$$

d. h. nach dem Satz gibt es ein $\varepsilon > 0$ und eine $\mathcal{C}^1$-Abbildung $g\colon U_\varepsilon(0) \to \mathbb{R}^2$, wobei $U_\varepsilon(0) = (-\varepsilon, \varepsilon) \subset \mathbb{R}$, so dass $F(x, g(x)) = 0$ für alle $x \in U_\varepsilon(0)$. Für die Jacobimatrix von g ergibt sich

$$J_g(0) = -\left(\frac{\partial F}{\partial(y,z)}(0,0,0)\right)^{-1} \frac{\partial F}{\partial x}(0,0,0) = -\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} -2 \\ -2 \end{pmatrix}.$$

Da g nur von einer Variablen abhängt, kann man dafür auch $g'(0)$ schreiben.

Zusatz: Wegen der Wichtigkeit dieses Satzes sei hier noch einmal das einfachste Beispiel zusammengefasst: Wir betrachten den Fall $m = k = 1$, d. h. $n = 2$, und die Funktion $F\colon \mathbb{R}^n \to \mathbb{R}^m$, gegeben durch $F(x,y) = x^2 + y^2 - 1$. Sei nun $p = (\bar{x}, \bar{y})^{\mathrm{T}} \in \mathbb{R}^2$ mit $F(p) = 0$, d. h. p liegt auf der Einheitskreislinie. Sofort rechnet man $J_F(x,y) = (2x, 2y) \in \mathbb{R}^{m \times n}$, also $\partial F(p)/\partial y = 2\bar{y} \in \mathbb{R}^{m \times m}$. Ist diese „Matrix“ invertierbar, d. h. ist $\bar{y} \neq 0$, so lassen sich die y-Koordinaten der Punkte der Menge $M = \{(x,y)^{\mathrm{T}} \mid F(x,y) = 0\}$ lokal als $\mathcal{C}^1$-Funktion g der x-Koordinaten schreiben. Je nach Vorzeichen von $\bar{y}$ ist dies natürlich $x \mapsto \sqrt{1 - x^2}$ oder $x \mapsto -\sqrt{1 - x^2}$. Anders ausgedrückt: Ist $\bar{y} \neq 0$, so kann die implizit gegebene Gleichung $F(x,y) = 0$ in einer Umgebung von p nach y aufgelöst werden, d. h. es gibt eine Funktion g, so dass $y = g(x)$ in expliziter Form. Im Fall $\bar{y} = 0$ gilt $p = (1,0)^{\mathrm{T}}$ oder $p = (-1,0)^{\mathrm{T}}$, und in diesen beiden Punkten kann die Kreislinie tatsächlich nicht lokal durch eine Funktion in x beschrieben werden. Weiter ist $J_g(\bar{x}) = -(\partial F(p)/\partial y)^{-1} \partial F(p)/\partial x = -\bar{x}/\bar{y}$, was nichts anderes als die Steigung der Tangenten in p ist. Dasselbe Ergebnis erhält man, wenn man $\pm\sqrt{1 - x^2}$ nach x ableitet.

33. (a) Für $f\colon U \to \mathbb{R}$ ist $g := \nabla f = (\partial_1 f, \dots, \partial_n f)^{\mathrm{T}}\colon U \to \mathbb{R}^n$ eine Vektorfunktion. Für deren Jacobimatrix ergibt sich

$$J_g(x) = \begin{pmatrix} \partial_1 g_1(x) & \dots & \partial_n g_1(x) \\ \vdots & \ddots & \vdots \\ \partial_1 g_n(x) & \cdots & \partial_n g_n(x) \end{pmatrix} = \begin{pmatrix} \partial_1\partial_1 f(x) & \dots & \partial_n\partial_1 f(x) \\ \vdots & \ddots & \vdots \\ \partial_1\partial_n f(x) & \cdots & \partial_n\partial_n f(x) \end{pmatrix} = H_f(x)\,,$$

d. h. $J_{\nabla f}(x) = H_f(x)$ für alle $x \in U$.

(b) Wir betrachten die Funktion $g := \nabla f\colon U \to \mathbb{R}^n$. Nach (a) gilt $J_g(x_0) = H_f(x_0)$, und diese Matrix ist nach Voraussetzung invertierbar, d. h. nach dem Satz über lokale Diffeomorphismen existiert eine Umgebung $\tilde{U} \subset U$ von x_0, so dass $g|_{\tilde{U}} \in \operatorname{Diff}(\tilde{U}, \tilde{V})$ mit $\tilde{V} := \operatorname{Bild} g|_{\tilde{U}}$.

Dass $x_0 \in U$ ein kritischer Punkt von f ist, bedeutet, dass $g(x_0) = 0$ gilt. Da $g|_{\tilde{U}}$ injektiv ist, ist $g(x) \neq 0$ für alle $x \in \tilde{U} \setminus \{x_0\}$. Wenn aber $\nabla f(x)$ nicht der Nullvektor ist, dann ist auch $\mathrm{d}f(x)\colon \mathbb{R}^n \to \mathbb{R}$ nicht die Nullabbildung, also gilt $\operatorname{Rang} \mathrm{d}f(x) > 0$. Damit ist $\mathrm{d}f(x)$ surjektiv für alle $x \in \tilde{U} \setminus \{x_0\}$, d. h. alle diese x sind reguläre Punkte von f.

34. Da U offen und $v \mapsto Tv$ ein Diffeomorphismus ist, ist auch $U_0 := \{T^{-1}u \mid u \in U\}$ offen, und da $z \in U$, ist $z_0 := T^{-1}z \in U_0$. Die Funktion $F\colon U_0 \to \mathbb{R}^m$ ist aus $\mathcal{C}^1$, da f und $v \mapsto Tv$ aus $\mathcal{C}^1$ sind und w konstant ist. Ferner ist $F(z_0) = f(Tz_0) - w = f(z) - w = 0$ nach der Definition von w, und mit der Kettenregel folgt

$$J_F(z_0) = J_f(Tz_0)\, J_{v \mapsto Tv}(z_0) = J_f(z)T\,.$$

Wir stellen nun explizit die Matrix $T = (t_{ij})$ auf. Dazu betrachten wir die i-te Komponente der beiden Seiten von $Te_j = e_{\pi(j)}$:

$$(Te_j)_i = \sum_{\ell=1}^{n} t_{i\ell}(e_j)_\ell = \sum_{\ell=1}^{n} t_{i\ell}\delta_{j\ell} = t_{ij}\,, \quad (e_{\pi(j)})_i = \delta_{i,\pi(j)}\,.$$

Damit erhalten wir

$$\big(J_f(z)T\big)_{ij} = \sum_{\ell=1}^{n} \partial_\ell f_i(z)\delta_{\ell,\pi(j)} = \partial_{\pi(j)} f_i(z)\,,$$

also

$$J_F(z_0) = J_f(z)T = \big(\partial_{\pi(1)} f(z), \dots, \partial_{\pi(n)} f(z)\big)\,.$$

Nun sind aber die Vektoren $\partial_{\pi(k+1)} f(z), \dots, \partial_{\pi(n)} f(z)$ nach Voraussetzung linear unabhängig, d. h. die Teilmatrix $(\partial_{\pi(k+1)} f(z), \dots, \partial_{\pi(n)} f(z)) \in \mathbb{R}^{m\times m}$ von $J_F(z_0)$ ist invertierbar. Nach dem Satz über implizite Funktionen existieren damit offene Umgebungen $U_1 \subset \mathbb{R}^k$ von $(z_{0,1}, \dots, z_{0,k})^{\mathrm{T}}$ und $U_2 \subset \mathbb{R}^m$ von $(z_{0,k+1}, \dots, z_{0,n})^{\mathrm{T}}$ und eine $\mathcal{C}^1$-Abbildung $g\colon U_1 \to U_2$, so dass für alle $\xi \in U_1$ und $\eta \in U_2$

$$F(\xi, \eta) = 0 \quad \iff \quad \eta = g(\xi)\,.$$

Mit der Definition von F schreibt sich die linke Gleichung als $f(T(\xi, \eta)) = w$. Die Äquivalenz besagt: Für ein $\zeta \in V := T(U_1 \times U_2)$ gilt $f(\zeta) = w$ genau dann, wenn $\zeta = T(\xi^{\mathrm{T}}, g(\xi)^{\mathrm{T}})^{\mathrm{T}}$ für ein $\xi \in U_1$. In Formeln geschrieben heißt das

$$f^{-1}\big(\{w\}\big) \cap V = \left\{ T \begin{pmatrix} \xi \\ g(\xi) \end{pmatrix} \,\middle|\, \xi \in U_1 \right\}.$$

Bemerkung: Die Aussage dieser Aufgabe verallgemeinert den Satz über implizite Funktionen dahingehend, dass F in dem Punkt, in dem aufgelöst werden soll, nicht 0 sein muss und dass die m Koordinaten, nach denen aufgelöst werden soll, nicht die letzten m Stück sein müssen. Statt dessen gilt jetzt allgemeiner: Seien $m, n \in \mathbb{N}$ mit $m < n$, $U \subset \mathbb{R}^n$ offen, $F \in \mathcal{C}^1(U, \mathbb{R}^m)$, $\bar{z} \in U$ und $w = F(\bar{z})$. Ferner gelte für die m Indizes $1 \le j_1 < j_2 < \ldots < j_m \le n$, dass die Teilmatrix

$$\big(\partial_{j_1} F(\bar{z}), \ldots, \partial_{j_m} F(\bar{z})\big) \in \mathbb{R}^{m \times m}$$

der Jacobimatrix $J_F(\bar{z}) \in \mathbb{R}^{m \times n}$ invertierbar ist. Sind dann $1 \le i_1 < i_2 < \ldots < i_{n-m} \le n$ die fehlenden $n - m$ Indizes, d. h. es gilt $\{i_1, \ldots, i_{n-m}\} \,\dot\cup\, \{j_1, \ldots, j_m\} = \{1, \ldots, n\}$, so gibt es eine offene Umgebung $U' \subset \mathbb{R}^{n-m}$ von $(\bar{z}_{i_1}, \ldots, \bar{z}_{i_{n-m}})^{\mathrm{T}}$, eine offene Umgebung $V' \subset \mathbb{R}^m$ von $(\bar{z}_{j_1}, \ldots, \bar{z}_{j_m})^{\mathrm{T}}$ und eine $\mathcal{C}^1$-Funktion $g \colon U' \to V'$, so dass für alle $z \in \mathbb{R}^n$ mit $(z_{i_1}, \ldots, z_{i_{n-m}})^{\mathrm{T}} \in U'$ und $(z_{j_1}, \ldots, z_{j_m})^{\mathrm{T}} \in V'$

$$F(z) = w \iff (z_{j_1}, \ldots, z_{j_m})^{\mathrm{T}} = g(z_{i_1}, \ldots, z_{i_{n-m}})\,.$$

35. (a) Für die Jacobimatrix gilt

$$J_f(x, y, z) = \begin{pmatrix} 2x & 2y & 0 \\ 0 & 2y & 2z \end{pmatrix}.$$

Das Differential $\mathrm{d}f(x, y, z)$ ist genau dann surjektiv, wenn $J_f(x, y, z)$ vollen Rang, also Rang 2 hat. Wir unterscheiden:

- $x \neq 0$ und $z \neq 0$: Dann gilt automatisch $\operatorname{Rang} J_f(x, y, z) = 2$.
- Entweder $x = 0$ oder $z = 0$: Dann gilt $\operatorname{Rang} J_f(x, y, z) = 2 \iff y \neq 0$.
- $x = 0$ und $z = 0$: Dann gilt $\operatorname{Rang} J_f(x, y, z) \le 1$.

Zusammen ergibt sich: Das Differential $\mathrm{d}f(x, y, z)$ ist genau dann surjektiv, wenn höchstens eine der Koordinaten x, y oder z gleich 0 ist. Die Menge der regulären Punkte ist also $\operatorname{RegP} f = \mathbb{R}^3 \setminus (A_x \cup A_y \cup A_z)$ mit $A_x := \{(x, 0, 0)^{\mathrm{T}} \mid x \in \mathbb{R}\}$, $A_y := \{(0, y, 0)^{\mathrm{T}} \mid y \in \mathbb{R}\}$ und $A_z := \{(0, 0, z)^{\mathrm{T}} \mid z \in \mathbb{R}\}$.

Für die regulären Werte betrachten wir zunächst

$$f(A_x) = \left\{ \begin{pmatrix} x^2 \\ 0 \end{pmatrix} \,\middle|\, x \in \mathbb{R} \right\},\ f(A_y) = \left\{ \begin{pmatrix} y^2 \\ y^2 \end{pmatrix} \,\middle|\, y \in \mathbb{R} \right\},\ f(A_z) = \left\{ \begin{pmatrix} 0 \\ z^2 \end{pmatrix} \,\middle|\, z \in \mathbb{R} \right\}.$$

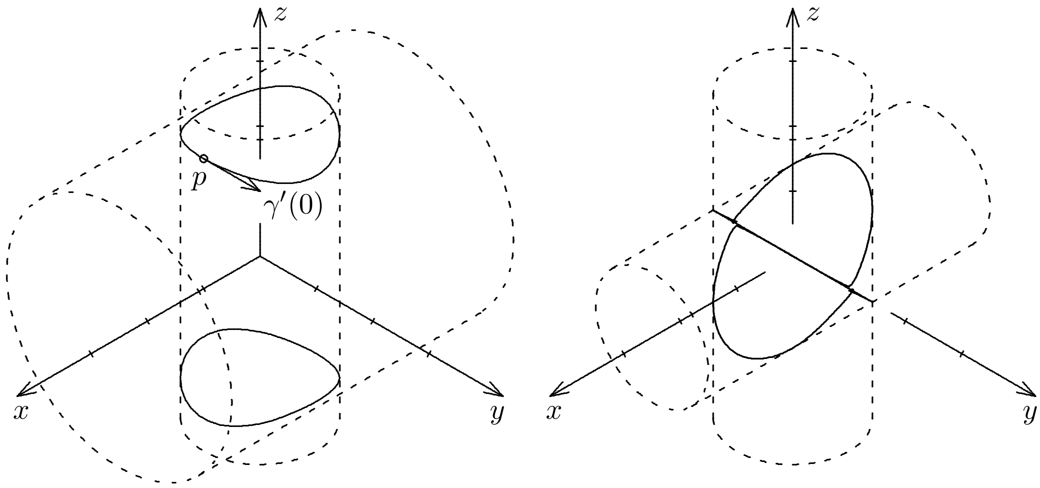
Ist also $w \in \mathbb{R}^2$ von keiner der drei Gestalten $(x^2, 0)^{\mathrm{T}}$, $(y^2, y^2)^{\mathrm{T}}$ oder $(0, z^2)^{\mathrm{T}}$, so enthält $f^{-1}(\{w\})$ keinen Punkt aus A_x, A_y oder A_z, d. h. $f^{-1}(\{w\}) \subset \operatorname{RegP} f$. Für die Menge der regulären Werte gilt also

$$\operatorname{RegW} f = \left\{ (u, v)^{\mathrm{T}} \in \mathbb{R}^2 \,\middle|\, u < 0 \text{ oder } v < 0 \text{ oder } (u, v > 0 \text{ und } u \neq v) \right\}.$$

Wegen $(1, 4)^{\mathrm{T}} \in \operatorname{RegW} f$ ist $M = f^{-1}(\{(1, 4)^{\mathrm{T}}\})$ nach dem Satz vom regulären Wert eine Untermannigfaltigkeit des $\mathbb{R}^3$ mit Dimension $3 - 2 = 1$.

Im linken Diagramm sind die beiden Zylindermantel-Mengen $\{(x, y, z)^{\mathrm{T}} \in \mathbb{R}^3 \mid x^2 + y^2 = 1\}$ und $\{(x, y, z)^{\mathrm{T}} \in \mathbb{R}^3 \mid y^2 + z^2 = 4\}$ gestrichelt und deren Schnittmenge M durchgezogen dargestellt. Rechts wurde der zweite Zylinder durch $\{(x, y, z)^{\mathrm{T}} \in \mathbb{R}^3 \mid y^2 + z^2 = 1{,}001\}$ ersetzt. Wählt man die rechte Seite gleich 1, dann berühren sich die vier entstehenden Spitzen, so dass zwei Überkreuzungspunkte entstehen.

Zusatz: Wir betrachten die Menge $M' := f^{-1}(\{(1, 1)^{\mathrm{T}}\})$ und einen der beiden Überkreuzungspunkte $p' := (0, 1, 0)^{\mathrm{T}} \in M'$. Wegen $\operatorname{Rang} J_f(p') = 1$ kann der Satz vom regulären Wert nicht angewendet werden. Und tatsächlich ist M' keine Untermannigfaltigkeit des $\mathbb{R}^3$, weil die Eigenschaften aus der Definition der Untermannigfaltigkeit nicht erfüllt sind, denn: Nehmen wir an, es gäbe offene (o. B. d. A. zusammenhängende) Mengen $V \subset \mathbb{R}^n$ und $U \subset \mathbb{R}^k$ mit $p' \in V$ und einer bijektiven stetigen Funktion $g^{-1}\colon M \cap V \to U$. Die Einschränkung $g^{-1}|_{(M\cap V)\setminus\{p'\}}$ bildet dann bijektiv und stetig von $(M \cap V) \setminus \{p'\}$ nach $U \setminus \{g^{-1}(p')\}$ ab. $(M \cap V) \setminus \{p'\}$ besteht aus vier Zusammenhangskomponenten, $U \setminus \{g^{-1}(p')\}$ jedoch für $k = 1$ nur aus zwei, für $k \geq 2$ aus einer. Dies steht im Widerspruch zur Stetigkeit von g^{-1}.



(b) Wir definieren

$$\gamma(t) := \begin{pmatrix} \sqrt{1-t^2} \\ t \\ \sqrt{4-t^2} \end{pmatrix}$$

und stellen fest:

- Es gilt $\gamma(0) = p = (1, 0, 2)^{\mathrm{T}}$.
- Es gilt $\gamma(t) \in M$ für alle t in einem geeigneten Definitionsbereich.
- γ ist eine $\mathcal{C}^1$-Funktion in einem geeigneten Definitionsbereich.
- Für $\varepsilon := \sqrt{15}/4 < 1$ gilt $\gamma(\pm\varepsilon) = (1/4, \pm\varepsilon, 7/4)^{\mathrm{T}}$ und

$$\big|p - \gamma(\pm\varepsilon)\big|_2 = \sqrt{\Big(1 - \frac{1}{4}\Big)^2 + (\pm\varepsilon)^2 + \Big(2 - \frac{7}{4}\Big)^2} = \frac{5}{4}\,.$$

Wir wählen als Definitionsbereich $I := (-\varepsilon, \varepsilon) \subset \mathbb{R}$ und als Umgebung von p die offene Menge $V := U_{5/4}(p)$. Dann gilt nach Konstruktion $\gamma(I) = M \cap V$.

(c) Da M eine eindimensionale Untermannigfaltigkeit ist, gilt auch $\dim T_pM = 1$. Der Tangentialraum besteht aus den Ableitungen aller Kurven in M durch p. Für die Kurve γ aus (b) gilt $\gamma(0) = p$, d. h. $\gamma'(0) = (0, 1, 0)^{\mathrm{T}}$ ist ein Tangentenvektor. Der Tangentialraum ist also

$$T_pM = \operatorname{span}\gamma'(0) = \left\{ \begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix} \,\middle|\, y \in \mathbb{R} \right\}.$$

Dasselbe ergibt sich, wenn man $T_pM = \operatorname{Kern} \mathrm{d}f(p)$ berechnet.

36. (a) Wähle irgendeine Basis von W und ergänze sie zu einer Basis vom $\mathbb{R}^n$ (Basisergänzungssatz). Wende das Gram-Schmidt-Verfahren auf diese Basis an, wodurch eine ONB $\{a_1, \ldots, a_k\}$ von W und eine ONB $\{a_1, \ldots, a_k\} \cup \{a_{k+1}, \ldots, a_n\}$ vom $\mathbb{R}^n$

entsteht. Ferner seien $v \in \mathbb{R}^n$ und $w \in W$ mit $v = \sum_{i=1}^n \lambda_i a_i$ und $w = \sum_{i=1}^k \mu_i a_i$. Dann gilt (in [...] sehr ausführlich)

$$
\begin{aligned}
|v-w|_2^2 &= \left| \sum_{i=1}^n \lambda_i a_i - \sum_{i=1}^k \mu_i a_i \right|_2^2 = \left| \sum_{i=1}^k (\lambda_i - \mu_i) a_i + \sum_{i=k+1}^n \lambda_i a_i \right|_2^2 \\
&\Bigg[= \left\langle \sum_{i=1}^k (\lambda_i - \mu_i) a_i, \sum_{i=1}^k (\lambda_i - \mu_i) a_i \right\rangle + \left\langle \sum_{i=k+1}^n \lambda_i a_i, \sum_{i=k+1}^n \lambda_i a_i \right\rangle \\
&\quad + 2 \left\langle \sum_{i=1}^k (\lambda_i - \mu_i) a_i, \sum_{i=k+1}^n \lambda_i a_i \right\rangle \\
&= \sum_{i=1}^k \sum_{j=1}^k (\lambda_i - \mu_i)(\lambda_j - \mu_j) \langle a_i, a_j \rangle + \sum_{i=k+1}^n \sum_{j=k+1}^n \lambda_i \lambda_j \langle a_i, a_j \rangle \\
&\quad + 2 \sum_{i=1}^k \sum_{j=k+1}^n (\lambda_i - \mu_i) \lambda_j \langle a_i, a_j \rangle \Bigg] = \sum_{i=1}^k (\lambda_i - \mu_i)^2 + \sum_{i=k+1}^n \lambda_i^2 .
\end{aligned}
$$

Zu gegebenem $v \in \mathbb{R}^n$ wird $|v - w|_2$ also genau dann minimal, wenn $\mu_i = \lambda_i$. Mit

$$
\langle a_i, v \rangle = \left\langle a_i, \sum_{j=1}^n \lambda_j a_j \right\rangle = \sum_{j=1}^n \lambda_j \langle a_i, a_j \rangle = \lambda_i
$$

gilt also

$$
w = \sum_{i=1}^k \langle a_i, v \rangle a_i .
$$

Dies zeigt die Existenz und die Eindeutigkeit. Ferner lesen wir in der Rechnung $v - w = \sum_{i=k+1}^n \langle a_i, v \rangle a_i$ ab. Mit (b) folgt daraus $v - w \in W^\perp$.

(b) Wir zeigen, dass $B^\perp$ ein Unterraum ist: Klarerweise gilt $0 \in B^\perp$. Seien $v, w \in B^\perp$ und $\lambda, \mu \in \mathbb{R}$. Für jedes $b \in B$ gilt dann $\langle b, v \rangle = \langle b, w \rangle = 0$, und wegen der Linearität des Skalarprodukts im zweiten Faktor auch $\langle b, \lambda v + \mu w \rangle = 0$, d. h. $\lambda v + \mu w \in B^\perp$.

Wir zeigen $\mathbb{R}^n = \operatorname{span} B \oplus B^\perp$: Analog zu (a) seien $\{a_1, \dots, a_k\}$ eine ONB von $\operatorname{span} B$ und $\{a_1, \dots, a_k\} \cup \{a_{k+1}, \dots, a_n\}$ eine ONB vom $\mathbb{R}^n$. Ferner sei $v \in \mathbb{R}^n$ mit $v = \sum_{i=1}^n \lambda_i a_i = v' + v''$, $v' = \sum_{i=1}^k \lambda_i a_i$ und $v'' = \sum_{i=k+1}^n \lambda_i a_i$. Nach Konstruktion gilt $v' \in \operatorname{span} B$ und $v'' \in B^\perp$, denn jede Linearkombination der $a_{k+1}, \dots, a_n$ steht senkrecht auf jeder Linearkombination der $a_1, \dots, a_k$. Damit folgt $\mathbb{R}^n = \operatorname{span} B + B^\perp$. Ist $v \in \operatorname{span} B$, so gilt $v'' = 0$, und ist $v \in B^\perp$, so gilt $v' = 0$. Zusammen erhalten wir $v = v' + v'' = 0$, d. h. die Summe ist direkt.

Schließlich zu $(B^\perp)^\perp = \operatorname{span} B$: „$\subset$": Die vorige Rechnung hat insbesondere auch $\operatorname{span} B \perp B^\perp$ gezeigt. Das bedeutet aber, dass von den beiden Vektorräume $\operatorname{span} B$ und $B^\perp$ der eine jeweils das orthogonale Komplement des anderen ist, d. h. es gilt $(B^\perp)^\perp = \operatorname{span} B$.

37. (a) i. Wir definieren die Nebenbedingung $g(x) = x_1 \cdots x_n - 1$. Dann sind $f, g \in \mathcal{C}^1(\mathbb{R}^n)$ und $M = g^{-1}(\{0\}) \cap [0, \infty)^n$, und es gilt

$$
J_g(x) = \left(\prod_{i=1,\, i \neq 1}^n x_i \quad \cdots \quad \prod_{i=1,\, i \neq n}^n x_i \right) .
$$

Jeder Eintrag dieser Jacobimatrix ist von 0 verschieden, denn aus $x_i = 0$ für ein $i \in \{1,\dots,n\}$ folgt $g(x) = -1$, d. h. $x \notin M$. Somit hat $J_g(x)$ für alle $x \in M$ maximalen Rang, d. h. 0 ist regulärer Wert von g, so dass folgt: Besitzt f in $\bar{x}$ ein Maximum unter der Nebenbedingung $\bar{x} \in M$, so gibt es einen Lagrange-Multiplikator λ, so dass

$$\nabla f(\bar{x}) - \lambda \nabla g(\bar{x}) = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} - \lambda \begin{pmatrix} \prod_{i=1,\, i\neq 1}^{n} \bar{x}_i \\ \vdots \\ \prod_{i=1,\, i\neq n}^{n} \bar{x}_i \end{pmatrix} = 0\,.$$

Wenn wir die k-te Komponente mit $\bar{x}_k > 0$ durchmultiplizieren, muss also für alle $k \in \{1,\dots,n\}$

$$\lambda \prod_{i=1}^{n} \bar{x}_i = \bar{x}_k$$

gelten. Da die linke Seite nicht von k abhängt, müssen alle Komponenten von $\bar{x}$ gleich sein, und mit $g(\bar{x}) = 0$ folgt $\bar{x} = (1,\dots,1)^{\mathrm{T}}$ und $f(\bar{x}) = n$.
Nun definieren wir für $m \in \mathbb{N}$ den Quader $Q_m := [0,m]^n \subset \mathbb{R}^n$. Dieser ist abgeschlossen und beschränkt, also kompakt; ferner ist M abgeschlossen, d. h. $M_m := M \cap Q_m$ für jedes $m \in \mathbb{N}$ kompakt (siehe Zusatz von Aufgabe 19. (a)). Die Funktion $f_m := f|_{M_m}$ nimmt also ein globales Minimum an. Der Rand von M_m als Teilmenge von M ist $M \cap \partial Q_m = \{x \in Q_m \mid x_i = m \text{ für ein } i\}$. Für ein solches x gilt jedoch $f_m(x) > m = f(\bar{x})$, d. h. f_m nimmt ihr globales Minimum nicht auf $M \cap \partial Q_m$ an. Damit muss sie es in $\bar{x}$ annehmen, genauso wie schließlich auch f selbst.

ii. Für $x_1,\dots,x_n > 0$ gilt nach i.

$$\prod_{i=1}^{n} x_i = 1 \quad \Longrightarrow \quad \sum_{i=1}^{n} x_i \geq n\,.$$

Seien nun $a_1,\dots,a_n > 0$ beliebig und definiere $x_i := a_i / \sqrt[n]{a_1 \cdots a_n}$. Dann ist

$$\prod_{i=1}^{n} x_i = \prod_{i=1}^{n} \frac{a_i}{\sqrt[n]{a_1 \cdots a_n}} = \frac{a_1 \cdots a_n}{a_1 \cdots a_n} = 1\,,$$

also

$$\sum_{i=1}^{n} \frac{a_i}{\sqrt[n]{a_1 \cdots a_n}} \geq n \quad \text{bzw.} \quad \sqrt[n]{a_1 \cdots a_n} \leq \frac{1}{n}(a_1 + \ldots + a_n)\,.$$

(b) Aufgrund der Symmetrie hat ein solcher Quader Q acht Eckpunkte der Form $(\pm x, \pm y, \pm z)^{\mathrm{T}}$ mit $x,y,z \geq 0$, und einer der Eckpunkte besitzt nur nicht-negative Koordinaten. Ist jedoch eine Koordinate 0, so ist das Volumen ebenfalls 0, also sicher nicht maximal. Gesucht ist daher das Maximum der Funktion

$$f(x,y,z) = 8xyz\,, \quad x,y,z > 0\,,$$

die das Volumen des Quaders beschreibt, unter der Nebenbedingung

$$g(x,y,z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0\,,$$

nämlich dass die Eckpunkte auf dem Ellipsoid liegen. Es ist $f, g \in \mathcal{C}^1$ und

$$J_g(x,y,z) = \begin{pmatrix} \frac{2x}{a^2} & \frac{2y}{b^2} & \frac{2z}{c^2} \end{pmatrix}$$

mit maximalem Rang, denn $x = y = z = 0$ erfüllt nicht die Nebenbedingung. Somit ist 0 ein regulärer Wert von g, und es gilt: Besitzt f in $(\bar{x}, \bar{y}, \bar{z})^{\mathrm{T}}$ ein Maximum unter der Nebenbedingung $g(\bar{x}, \bar{y}, \bar{z}) = 0$, so gibt es einen Lagrange-Multiplikator λ, so dass

$$\nabla f(\bar{x}, \bar{y}, \bar{z}) - \lambda \nabla g(\bar{x}, \bar{y}, \bar{z}) = \begin{pmatrix} 8\bar{y}\bar{z} - 2\lambda\bar{x}/a^2 \\ 8\bar{x}\bar{z} - 2\lambda\bar{y}/b^2 \\ 8\bar{x}\bar{y} - 2\lambda\bar{z}/c^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Wenn wir die erste Komponente mit $\bar{x}$, die zweite mit $\bar{y}$ und die dritte mit $\bar{z}$ multiplizieren und alles addieren, erhalten wir

$$0 = 8\bar{x}\bar{y}\bar{z} - \frac{2\lambda\bar{x}^2}{a^2} + 8\bar{x}\bar{y}\bar{z} - \frac{2\lambda\bar{y}^2}{b^2} + 8\bar{x}\bar{y}\bar{z} - \frac{2\lambda\bar{z}^2}{c^2} = 24\bar{x}\bar{y}\bar{z} - 2\lambda\,,$$

wobei wir die Nebenbedingung verwendet haben. Setzen wir also $2\lambda = 24\bar{x}\bar{y}\bar{z}$ in die ursprüngliche Gleichung ein, so erhalten wir

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 8\bar{y}\bar{z} - 24\bar{x}^2\bar{y}\bar{z}/a^2 \\ 8\bar{x}\bar{z} - 24\bar{x}\bar{y}^2\bar{z}/b^2 \\ 8\bar{x}\bar{y} - 24\bar{x}\bar{y}\bar{z}^2/c^2 \end{pmatrix} = \begin{pmatrix} 8\bar{y}\bar{z}(1 - 3\bar{x}^2/a^2) \\ 8\bar{x}\bar{z}(1 - 3\bar{y}^2/b^2) \\ 8\bar{x}\bar{y}(1 - 3\bar{z}^2/c^2) \end{pmatrix},$$

d. h. es folgt $\bar{x} = a/\sqrt{3}$, $\bar{y} = b/\sqrt{3}$ und $\bar{z} = c/\sqrt{3}$. Da f stetig auf der kompakten Menge $\{(x,y,z)^{\mathrm{T}} \in \mathbb{R}^3 \mid x,y,z \geq 0\,,\ g(x,y,z) = 0\}$ und auf dem Rand der Menge 0 ist, muss f im gefundenen kritischen Punkt ihr Maximum annehmen. Das maximale Volumen ist also

$$f(\bar{x}, \bar{y}, \bar{z}) = \frac{8abc}{3\sqrt{3}}\,.$$

38. Da f stetig und K kompakt ist, nimmt f ein globales Minimum und ein globales Maximum an. Wir suchen zuerst nach Minima und Maxima von $f|_{K^\circ}$ und danach nach solchen von $f|_{\partial K}$.

- Es gilt $f|_{K^\circ} \in \mathcal{C}^1$. Für $x \in K^\circ$ ist

  $$\nabla f(x) = \begin{pmatrix} 18x_1 - 4x_2 \\ -4x_1 + 12x_2 \end{pmatrix} = \begin{pmatrix} 18 & -4 \\ -4 & 12 \end{pmatrix} x\,.$$

  Da die Matrix regulär ist, hat $\nabla f(\bar{x}) = 0$ nur die triviale Lösung $\bar{x} = 0$. Einsetzen liefert sofort $f(\bar{x}) = 0$.

- ∂K wird durch die Nebenbedingung $g(x) = |x|_2^2 - 1 = 0$ charakterisiert. Es gilt $f, g \in \mathcal{C}^1$ und

  $$J_g(x) = \begin{pmatrix} 2x_1 & 2x_2 \end{pmatrix}$$

  mit maximalem Rang, da $x_1 = x_2 = 0$ nicht die Nebenbedingung erfüllt. Also ist 0 ein regulärer Wert von g, und es gilt: Besitzt f in $\bar{x}$ ein Extremum unter der Nebenbedingung $g(\bar{x}) = 0$, so gibt es einen Lagrange-Multiplikator λ, so dass

  $$\nabla f(\bar{x}) - \lambda \nabla g(\bar{x}) = \begin{pmatrix} 18 & -4 \\ -4 & 12 \end{pmatrix} \bar{x} - 2\lambda\bar{x} = \begin{pmatrix} 18 - 2\lambda & -4 \\ -4 & 12 - 2\lambda \end{pmatrix} \bar{x} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Damit dieses Gleichungssystem nicht-triviale Lösungen hat, muss

$$\begin{vmatrix} 18-2\lambda & -4 \\ -4 & 12-2\lambda \end{vmatrix} = (18-2\lambda)(12-2\lambda) - 16 = 4(\lambda^2 - 15\lambda + 50) = 0$$

gelten. Dies ist wegen

$$\lambda = \frac{15}{2} \pm \sqrt{\frac{225}{4} - 50} = \frac{15}{2} \pm \frac{5}{2}$$

für $\lambda = 5$ oder $\lambda = 10$ erfüllt.
Für $\lambda = 5$ rechnen wir

$$\begin{pmatrix} 8 & -4 \\ -4 & 2 \end{pmatrix} \bar{x} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Longleftrightarrow \quad \bar{x} = a \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad a \in \mathbb{R}.$$

Normieren (so dass $g(\bar{x}) = 0$ gilt) ergibt

$$\bar{x} = \pm \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \text{mit} \quad f(\bar{x}) = \frac{9 - 8 + 24}{5} = 5.$$

Für $\lambda = 10$ rechnen wir

$$\begin{pmatrix} -2 & -4 \\ -4 & -8 \end{pmatrix} \bar{x} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Longleftrightarrow \quad \bar{x} = a \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \quad a \in \mathbb{R}.$$

Normieren (so dass $g(\bar{x}) = 0$ gilt) ergibt

$$\bar{x} = \pm \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ -1 \end{pmatrix} \quad \text{mit} \quad f(\bar{x}) = \frac{36 + 8 + 6}{5} = 10.$$

Vergleichen der Funktionswerte ergibt: f besitzt ein globales Minimum vom Wert 0 in $(0,0)^{\mathrm{T}}$ und zwei globale Maxima vom Wert 10 in $\pm(2/\sqrt{5}, -1/\sqrt{5})^{\mathrm{T}}$. Es gilt also $\max\{f(x) \mid x \in K\} = 10$ und $\min\{f(x) \mid x \in K\} = 0$.

Bemerkung: Schreibt man f als quadratische Form einer symmetrischen Matrix $A \in \mathbb{R}^{2\times 2}$, also

$$f(x) = x^{\mathrm{T}} A x \quad \text{mit} \quad A = \begin{pmatrix} 9 & -2 \\ -2 & 6 \end{pmatrix},$$

so rechnet man sofort nach, dass A positiv definit ist, d. h. $x_3 = f(x_1, x_2)$ beschreibt ein Paraboloid im $\mathbb{R}^3$. Dies stimmt mit der Tatsache überein, dass f in $(0,0)^{\mathrm{T}}$ ein Minimum besitzt.

39. (a) Wir zeigen die Aussage i., also dass (U, V_+, ψ) für jedes $p \in S \setminus \{e_n\}$ eine lokale Parametrisierung von S um p ist:

- $p \in V_+$ ist klar, da $S \setminus \{e_n\} \subset \mathbb{R}^n \setminus \{e_n\}$.
- Wegen $S \cap V_+ = S \setminus \{e_n\}$ ist zu zeigen, dass $\psi\colon \mathbb{R}^{n-1} \to S \setminus \{e_n\}$ bijektiv ist. Dazu zeigen wir, dass

$$h(y) := \frac{1}{1 - y_n} \begin{pmatrix} y_1 \\ \vdots \\ y_{n-1} \end{pmatrix}, \quad y \in S \setminus \{e_n\},$$

die Umkehrfunktion von ψ ist. Es ist

$$\big|h(y)\big|_2^2 = \frac{y_1^2 + \ldots + y_{n-1}^2}{(1-y_n)^2} = \frac{1-y_n^2}{(1-y_n)^2} \quad \text{wegen} \quad |y|_2 = 1\,,$$

d. h. für $k \in \{1, \ldots, n-1\}$ gilt

$$\psi_k\big(h(y)\big) = \frac{2h_k(y)}{1+|h(y)|_2^2} = \frac{2(1-y_n)y_k}{(1-y_n)^2 + (1-y_n^2)} = y_k$$

sowie

$$\psi_n\big(h(y)\big) = \frac{|h(y)|_2^2 - 1}{1+|h(y)|_2^2} = \frac{(1-y_n^2) - (1-y_n)^2}{(1-y_n)^2 + (1-y_n^2)} = y_n\,.$$

Dies zeigt, dass $\psi \circ h = \mathrm{id}_{S\setminus\{e_n\}}$. Andersherum gilt für $x \in \mathbb{R}^{n-1}$

$$h\big(\psi(x)\big) = \frac{1}{1-\psi_n(x)} \begin{pmatrix} \psi_1(x) \\ \vdots \\ \psi_{n-1}(x) \end{pmatrix} = \left(1 - \frac{|x|_2^2 - 1}{1+|x|_2^2}\right)^{-1} \frac{2x}{1+|x|_2^2} = x\,,$$

d. h. $h \circ \psi = \mathrm{id}_{\mathbb{R}^{n-1}}$. Damit ist ψ bijektiv mit $\psi^{-1} = h$.

- Die angegebene Funktion $h\colon S \cap V_+ \to U$ ist offensichtlich stetig.
- Aus (b) folgt $\det(J_\psi(x)^{\mathrm{T}} J_\psi(x)) > 0$ für alle $x \in \mathbb{R}^{n-1}$, d. h. $J_\psi(x)$ hat vollen Rang, so dass $\mathrm{d}\psi(x) \in \mathcal{L}(\mathbb{R}^{n-1}, \mathbb{R}^n)$ injektiv für alle $x \in \mathbb{R}^{n-1}$ ist.

(b) Für $x \in \mathbb{R}^{n-1}$ und $k \in \{1, \ldots, n-1\}$ rechnen wir

$$\partial_k \frac{2x}{1+|x|_2^2} = \frac{2(1+|x|_2^2)\partial_k x - 2x\partial_k(1+|x|_2^2)}{(1+|x|_2^2)^2} = \frac{2(1+|x|_2^2)e_k - 4x_k x}{(1+|x|_2^2)^2}$$

und

$$\begin{aligned} \partial_k \frac{|x|_2^2 - 1}{1+|x|_2^2} &= \frac{(1+|x|_2^2)\partial_k(|x|_2^2 - 1) - (|x|_2^2 - 1)\partial_k(1+|x|_2^2)}{(1+|x|_2^2)^2} \\ &= \frac{2(1+|x|_2^2)x_k - 2(|x|_2^2 - 1)x_k}{(1+|x|_2^2)^2} = \frac{4x_k}{(1+|x|_2^2)^2}\,. \end{aligned}$$

Damit folgt

$$\begin{aligned} J_\psi(x) &= \begin{pmatrix} \partial_1 \begin{pmatrix} \dfrac{2x}{1+|x|_2^2} \\ \dfrac{|x|_2^2 - 1}{1+|x|_2^2} \end{pmatrix} & \cdots & \partial_{n-1} \begin{pmatrix} \dfrac{2x}{1+|x|_2^2} \\ \dfrac{|x|_2^2 - 1}{1+|x|_2^2} \end{pmatrix} \end{pmatrix} \\ &= \frac{1}{(1+|x|_2^2)^2} \begin{pmatrix} 2(1+|x|_2^2)\mathbb{1} - 4xx^{\mathrm{T}} \\ 4x^{\mathrm{T}} \end{pmatrix} \end{aligned}$$

und analog

$$J_\varphi(x) = \frac{1}{(1+|x|_2^2)^2} \begin{pmatrix} 2(1+|x|_2^2)\mathbb{1} - 4xx^{\mathrm{T}} \\ -4x^{\mathrm{T}} \end{pmatrix},$$

d. h.

$$J_\psi(x)^{\mathrm{T}} = \frac{1}{(1+|x|_2^2)^2}\Big(2(1+|x|_2^2)\mathbb{1} - 4xx^{\mathrm{T}} \quad 4x\Big)\,,$$
$$J_\varphi(x)^{\mathrm{T}} = \frac{1}{(1+|x|_2^2)^2}\Big(2(1+|x|_2^2)\mathbb{1} - 4xx^{\mathrm{T}} \quad -4x\Big)\,.$$

Multiplikation ergibt

$$\begin{aligned}
J_\psi(x)^{\mathrm{T}} J_\psi(x) &= \frac{1}{(1+|x|_2^2)^4}\Big(\big(2(1+|x|_2^2)\mathbb{1} - 4xx^{\mathrm{T}}\big)^2 + 16xx^{\mathrm{T}}\Big) \\
&= \frac{1}{(1+|x|_2^2)^4}\Big(4(1+|x|_2^2)^2\mathbb{1} - 16(1+|x|_2^2)xx^{\mathrm{T}} + 16xx^{\mathrm{T}}xx^{\mathrm{T}} + 16xx^{\mathrm{T}}\Big) \\
&= \frac{1}{(1+|x|_2^2)^4}\Big(4(1+|x|_2^2)^2\mathbb{1}\Big) = \frac{4}{(1+|x|_2^2)^2}\mathbb{1}\,.
\end{aligned}$$

Die Rechnung für $J_\varphi(x)^{\mathrm{T}} J_\varphi(x)$ verläuft exakt genauso.

40. (a) Falls $I \setminus (I_1 \cup \cdots \cup I_k) = \emptyset$, wählen wir $\ell = 0$. Ansonsten seien $a_\kappa, b_\kappa \in \mathbb{R}^n$ für $\kappa \in \{0,\ldots,k\}$ definiert durch $I = (a_0, b_0]$ und $I_\kappa = (a_\kappa, b_\kappa]$ für $\kappa \in \{1,\ldots,k\}$. Daraus bilden die Mengen (Zerlegungen) $Z_\nu = \{a_0^\nu,\ldots,a_k^\nu, b_0^\nu,\ldots,b_k^\nu\}$ für $\nu \in \{1,\ldots,n\}$. Durch Sortieren und Entfernen doppelter Elemente schreiben wir $Z_\nu = \{z_{\nu,0},\ldots,z_{\nu,m_\nu}\}$ mit $m_\nu = |Z_\nu| - 1$ und $z_{\nu,\mu-1} < z_{\nu,\mu}$ für $\mu \in \{1,\ldots,m_\nu\}$. Ferner definieren wir $\tilde{J}_{\mu_1,\ldots,\mu_n} = (z_{1,\mu_1-1}, z_{1,\mu_1}] \times \cdots \times (z_{n,\mu_n-1}, z_{n,\mu_n}] \in Q_n$ für $\mu_\nu \in \{1,\ldots,m_\nu\}$, $\nu \in \{1,\ldots,n\}$.
Mit dieser Konstruktion gilt: Für alle $\kappa \in \{0,\ldots,k\}$ gibt es Indexmengen $M_\kappa \subset \{1,\ldots,m_1\} \times \cdots \times \{1,\ldots,m_n\}$, so dass

$$I = \dot{\bigcup_{(\mu_1,\ldots,\mu_n)\in M_0}} \tilde{J}_{\mu_1,\ldots,\mu_n} \quad \text{und} \quad I_\kappa = \dot{\bigcup_{(\mu_1,\ldots,\mu_n)\in M_\kappa}} \tilde{J}_{\mu_1,\ldots,\mu_n} \quad \forall\, \kappa \in \{1,\ldots,k\}\,.$$

Damit folgt

$$I \setminus (I_1 \cup \cdots \cup I_k) = \dot{\bigcup_{(\mu_1,\ldots,\mu_n)\in M_0\setminus(M_1\cup\cdots\cup M_k)}} \tilde{J}_{\mu_1,\ldots,\mu_n}\,,$$

was zu zeigen war.

Bemerkung: Ein anderer, dem Skript ähnlicherer Beweis ist der folgende: Nach de Morgan gilt

$$I \setminus (I_1 \cup \cdots \cup I_k) = \bigcap_{\kappa=1}^{k} I \setminus I_\kappa\,.$$

Nun ist aber nach Skript $I \setminus I_\kappa$, sofern $I, I_\kappa \in Q_n$, eine disjunkte Vereinigung endlich vieler halboffener Quader, d. h.

$$I \setminus (I_1 \cup \cdots \cup I_k) = \bigcap_{\kappa=1}^{k} \dot{\bigcup_{\mu=1}^{m_\kappa}} \tilde{J}_{\kappa,\mu}\,.$$

Mit dem Distributivgesetz folgt weiter

$$I \setminus (I_1 \cup \cdots \cup I_k) = \bigcup_{\mu_1\in\{1,\ldots,m_1\},\,\ldots,\,\mu_k\in\{1,\ldots,m_k\}} \bigcap_{\kappa=1}^{k} \tilde{J}_{\kappa,\mu_\kappa}\,.$$

Ebenfalls nach Skript ist $J_{\mu_1,\ldots,\mu_k} := \bigcap_{\kappa=1}^{k} \tilde{J}_{\kappa,\mu_\kappa}$ wieder ein halboffener Quader. Nun argumentieren wir noch, warum die Vereinigung disjunkt ist: Ist nämlich $(\mu_1,\ldots,\mu_k) \neq (\nu_1,\ldots,\nu_k)$, etwa $\mu_\kappa \neq \nu_\kappa$, so gilt $J_{\mu_1,\ldots,\mu_k} \subset \tilde{J}_{\kappa,\mu_\kappa}$ und $J_{\nu_1,\ldots,\nu_k} \subset \tilde{J}_{\kappa,\nu_\kappa}$, und da $\tilde{J}_{\kappa,\mu_\kappa}$ und $\tilde{J}_{\kappa,\nu_\kappa}$ disjunkt sind, folgt die Behauptung.

(b) An der nebenstehenden Skizze lesen wir für $I \setminus (I_1 \cup I_2)$ die mögliche Zerlegung

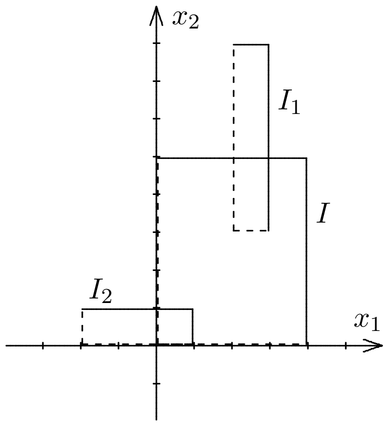
$$I \setminus (I_1 \cup I_2) = J_1 \,\dot{\cup}\, J_2 \,\dot{\cup}\, J_3 \,\dot{\cup}\, J_4$$

mit

$$J_1 = (0,1] \times (1,5]\,, \quad J_2 = (1,2] \times (0,5]\,,$$
$$J_3 = (2,3] \times (0,3]\,, \quad J_4 = (3,4] \times (0,5]$$

ab. Selbstverständlich ist die Zerlegung nicht eindeutig und wurde auch nicht dem Beweis aus (a) folgend konstruiert. Letzteres hätte deutlich mehr Quader erzeugt.

Skizze:



41. Für alle $n \in \mathbb{N}_0$ und $k, \ell \in \{1, \ldots, 2^n\}$ definieren wir den halboffenen Quader

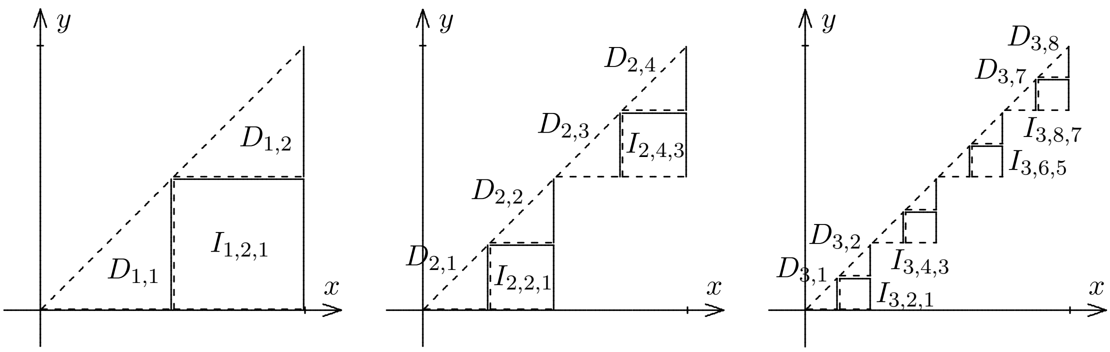
$$I_{n,k,\ell} = \big((k-1)\cdot 2^{-n}, k \cdot 2^{-n}\big] \times \big((\ell-1)\cdot 2^{-n}, \ell \cdot 2^{-n}\big] \in Q_2\,.$$

Ferner definieren wir für alle $n \in \mathbb{N}_0$ und $k \in \{1, \ldots, 2^n\}$ die Dreiecke

$$D_{n,k} = \big\{(x,y)^{\mathrm{T}} \in I_{n,k,k} \;\big|\; y < x\big\} \subset M\,.$$

Dann gilt $D_{0,1} = M$, und für jedes $n \in \mathbb{N}_0$ und $k \in \{1, \ldots, 2^n\}$ ist

$$D_{n,k} = I_{n+1,2k,2k-1} \,\dot{\cup}\, D_{n+1,2k-1} \,\dot{\cup}\, D_{n+1,2k}\,.$$



Damit ist rekursiv eine Zerlegung von M in abzählbar viele disjunkte Quader definiert, nämlich

$$M = \dot{\bigcup_{n=1}^{\infty}} \, \dot{\bigcup_{k=1}^{2^{n-1}}} I_{n,2k,2k-1}\,,$$

und es gilt

$$|M| = \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} |I_{n,2k,2k-1}| = \sum_{n=1}^{\infty} \sum_{k=1}^{2^{n-1}} 2^{-2n} = \sum_{n=1}^{\infty} 2^{-n-1} = \frac{1}{4} \sum_{n=0}^{\infty} \Big(\frac{1}{2}\Big)^n = \frac{1}{2}\,.$$

Bemerkung: Natürlich ist dieses Ergebnis keine große Überraschung, denn schließlich muss der Flächeninhalt gleich dem Wert des klassischen Riemann-Integrals $\int_0^1 x\,\mathrm{d}x = 1/2$ sein.

42. (a) In Aufgabe 37 (a) i. wurde für $c_1, \dots, c_n > 0$ die Ungleichung

$$\sqrt[n]{c_1 \cdots c_n} \leq \frac{1}{n}(c_1 + \ldots + c_n)$$

bewiesen. (Für den etwas allgemeineren Fall $c_1, \dots, c_n \geq 0$ ist sie offenbar auch wahr.) Ist $I = (a, b] \in Q_n$ mit $a, b \in \mathbb{R}^n$, so setzen wir $c_k = (b_k - a_k)^2 \geq 0$ und erhalten

$$\left(\prod_{k=1}^{n}(b_k - a_k)^2\right)^{1/n} \leq \frac{1}{n}\sum_{k=1}^{n}(b_k - a_k)^2 .$$

Potenzieren mit $n/2$ liefert

$$|I| = \prod_{k=1}^{n}(b_k - a_k) \leq n^{-n/2}\left(\sum_{k=1}^{n}(b_k - a_k)^2\right)^{n/2} = n^{-n/2}(\operatorname{diam} I)^n ,$$

denn offenbar ist $\operatorname{diam} I$ die Länge der Raumdiagonalen. Für einen Würfel W gilt $c_1 = \ldots = c_n = \ell^2$, und beide Seiten der Ungleichung sind ℓ^2. Damit folgt auch die zu zeigende Gleichheit.

(b) Ist $I = (a, b] \in Q_n$ mit $a, b \in \mathbb{R}^n$, so setze $\ell = \min\{b_j - a_j \mid j \in \{1, \dots, n\}\}$. Für $\ell = 0$ ist $I = \emptyset$, und es ist nichts zu zeigen. Ansonsten definiere $m_j = \lceil (b_j - a_j)/\ell \rceil$ für $j \in \{1, \dots, n\}$. Dann kann I von $k = m_1 \cdots m_n$ Würfeln mit Kantenlänge ℓ überdeckt werden, also von verschobenen Kopien des Würfels $\tilde{W} = (a_1, a_1 + \ell] \times \cdots \times (a_n, a_n + \ell]$. Wir bezeichnen sie mit $\{W_i \mid i \in \{1, \dots, k\}\}$, und es gilt

$$\begin{aligned}\sum_{i=1}^{k}|W_i| = k\ell^n = \prod_{j=1}^{n} m_j\ell = \prod_{j=1}^{n}\left\lceil\frac{b_j - a_j}{\ell}\right\rceil\ell &\leq \prod_{j=1}^{n}\left(\frac{b_j - a_j}{\ell} + 1\right)\ell \\ = \prod_{j=1}^{n}(b_j - a_j + \ell) &\leq \prod_{j=1}^{n}(b_j - a_j + b_j - a_j) = 2^n|I| .\end{aligned}$$

(c) Nach Definition ist $A \subset \mathbb{R}^n$ genau dann eine Nullmenge, wenn $\lambda^*(A) = 0$, d. h.

$$\inf\left\{\sum_{k=1}^{\infty}|A_k| \;\middle|\; A_k \in Q_n\,,\ A \subset \bigcup_{k=1}^{\infty} A_k\right\} = 0 .$$

Das ist wegen der Infimumseigenschaft äquivalent zu: Zu jedem $\varepsilon > 0$ gibt es $I_k \in Q_n$ für $k \in \mathbb{N}$ mit $A \subset \bigcup_{k=1}^{\infty} I_k$ und $\sum_{k=1}^{\infty}|I_k| < \varepsilon$.

„$\Longrightarrow$": Seien also A eine Nullmenge und $\varepsilon > 0$. Dann gibt es $I_k \in Q_n$ für $k \in \mathbb{N}$ mit $A \subset \bigcup_{k=1}^{\infty} I_k$ und $\sum_{k=1}^{\infty}|I_k| < \varepsilon$. Nach (b) gibt es zu jedem I_k endlich viele Würfel $W_{k,1}, \dots, W_{k,m_k} \in Q_n$ mit $I_k \subset \bigcup_{i=1}^{m_k} W_{k,i}$ und

$$\sum_{i=1}^{m_k} n^{-n/2}\big(\operatorname{diam}(W_{k,i})\big)^n \overset{\text{(a)}}{=} \sum_{i=1}^{m_k}|W_{k,i}| \leq 2^n|I_k| .$$

Multiplikation mit $n^{n/2}$ und Summation über k liefert

$$\sum_{k=1}^{\infty}\sum_{i=1}^{m_k}\big(\operatorname{diam}(W_{k,i})\big)^n \leq \sum_{k=1}^{\infty} 2^n n^{n/2}|I_k| \leq 2^n n^{n/2}\varepsilon =: \varepsilon' .$$

Die abzählbare Menge $\{W_{k,i}\}$ erfüllt also die Behauptung für ε'.

„$\Longleftarrow$“: Seien also $A \in \mathbb{R}^n$, $\varepsilon > 0$ und $A_k \subset \mathbb{R}^n$ für $k \in \mathbb{N}$ mit $A \subset \bigcup_{k=1}^{\infty} A_k$ und $\sum_{k=1}^{\infty}(\operatorname{diam}(A_k))^n < \varepsilon$. Wähle $x_k \in A_k$ beliebig; dann gilt $A_k \subset B_{\operatorname{diam}(A_k)}(x_k) =: K_k$ und $\operatorname{diam}(K_k) = 2\operatorname{diam}(A_k)$. Nun definiere den Würfel

$$W_k := \prod_{\nu=1}^{n}\Big(x_k^\nu - \operatorname{diam}(K_k), x_k^\nu + \operatorname{diam}(K_k)\Big] \in Q_n$$

mit $W_k \supset K_k$. Dann gilt $A \subset \bigcup_{k=1}^{\infty} W_k$ und

$$\sum_{k=1}^{\infty}|W_k| = \sum_{k=1}^{\infty}\big(2\operatorname{diam}(K_k)\big)^n = 4^n\sum_{k=1}^{\infty}\big(\operatorname{diam}(A_k)\big)^n < 4^n\varepsilon =: \varepsilon'\,.$$

Da ε' beliebig ist, ist A eine Nullmenge.

43. (a) Wegen $f \in \mathcal{C}^1$ sind alle partiellen Ableitungen $\partial_k f_i\colon D \to \mathbb{R}$ stetig. Daher ist auch $J_f\colon D \to \mathbb{R}^{n\times n}$, also $x \mapsto J_f(x)$ stetig; ebenso natürlich das zugehörige Differential $\mathrm{d}f\colon D \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, also $x \mapsto \mathrm{d}f(x)$. Die Einschränkung $\mathrm{d}f|_K$ auf die kompakte Menge K ist immer noch stetig und dort beschränkt, es existiert also die Operatornorm $L := \|\mathrm{d}f|_K\|$. Mit dem Schrankensatz 9.21 folgt dann

$$\big|f(x+v) - f(x)\big|_2 \le \max\Big\{\big\|\mathrm{d}f(x+tv)\big\| \;\Big|\; t \in [0,1]\Big\}\,|v|_2 \le L|v|_2\,,$$

sofern $x \in K$ und $x + v \in K$. (Da K konvex ist, gilt dann auch $x + tv \in K$ für alle $t \in [0,1]$.) Ersetzt man v durch $y - x$, so entsteht die Lipschitz-Stetigkeit mit Konstante L.

(b) Die Inklusion „$\supset$“ ist trivial, da $D \supset K$ für alle $K \in \hat{Q}$ nach Definition. Um „$\subset$“ zu zeigen, sei $x \in D$. Da D offen ist, gibt es ein $\varepsilon > 0$, so dass $U_\varepsilon(x) \subset D$ (bezüglich d_2). Sei nun $\tilde{K} = [x_1-\delta, x_1+\delta]\times\cdots\times[x_n-\delta, x_n+\delta]$. Damit $\tilde{K} \subset U_\varepsilon(x)$ erfüllt ist, muss $\delta > 0$ klein genug gewählt werden: Die Raumdiagonale von $\tilde{K}$ hat die Länge $2\sqrt{n}\,\delta$, der Durchmesser von $U_\varepsilon(x)$ ist 2ε. Wähle also δ mit $0 < \delta < \varepsilon/\sqrt{n}$.

Da $\mathbb{Q}$ dicht in $\mathbb{R}$ liegt, gibt es für alle $i \in \{1,\dots,n\}$ eine Zahl $a_i \in \mathbb{Q}$ mit $x_i - \delta < a_i < x_i$ und analog $b_i \in \mathbb{Q}$ mit $x_i < b_i < x_i + \delta$. Dadurch werden zwei Vektoren $a, b \in \mathbb{Q}^n$ definiert, und es gilt $\hat{Q} \ni K := [a,b] \subset \tilde{K} \subset U_\varepsilon(x) \subset D$ mit $x \in K$.

(c) Seien $A \subset D$ eine Nullmenge und $K \in \hat{Q}$. Dann ist auch $A \cap K$ eine Nullmenge, d. h. für jedes $\varepsilon > 0$ gibt es nach Aufgabe 42. (c) abzählbar viele Mengen $A_k \in D$ mit $A \cap K \subset \bigcup_{k=1}^{\infty}(A_k \cap K)$ und $\sum_{k=1}^{\infty}(\operatorname{diam}(A_k\cap K))^n < \varepsilon$. Für je zwei Punkte $x, y \in A_k \cap K$ gilt

$$\big|f(x) - f(y)\big|_2 \le L|x-y|_2 \le L\operatorname{diam}(A_k \cap K)\,,$$

da f auf $A_k \cap K$ Lipschitz-stetig ist. Übergang zum Supremum liefert schließlich $\operatorname{diam}(f(A_k \cap K)) \le L\operatorname{diam}(A_k \cap K)$.

Damit folgt nun $\sum_{k=1}^{\infty}(\operatorname{diam}(f(A_k\cap K)))^n \le L^n\varepsilon =: \varepsilon'$, d. h. $\bigcup_{k=1}^{\infty} f(A_k \cap K)$ ist eine Nullmenge. Da Funktionsauswertung und Vereinigung miteinander vertauscht werden können, ist ist letzte Menge gerade $f(A\cap K)$. $f(A\cap K)$ ist also für jeden Quader $K \in \hat{Q}$ eine Nullmenge.

Nach (b) gilt jedoch $D = \bigcup_{K\in\hat{Q}} K$. Wegen $A \subset D$ folgt damit

$$f(A) = f(A\cap D) = f\left(A \cap \bigcup_{K\in\hat{Q}} K\right) = f\left(\bigcup_{K\in\hat{Q}} (A\cap K)\right) = \bigcup_{K\in\hat{Q}} f(A\cap K)\,.$$

Nun ist aber $\hat{Q}$ eine Teilmenge von $\mathbb{Q}^n \times \mathbb{Q}^n$, und dies ist ein kartesisches Produkt abzählbarer Mengen, also selbst abzählbar. $f(A)$ ist damit die abzählbare Vereinigung der Nullmengen $f(A\cap K)$, d. h. wieder eine Nullmenge.

44. (a) Für $\varepsilon > 0$ seien $I_\ell \in Q_k$, $\ell \in \mathbb{N}$, mit $A \subset \bigcup_{\ell=1}^\infty I_\ell$ und $\sum_{\ell=1}^\infty |I_\ell| < \varepsilon$. Für $R \in \mathbb{N}$ definieren wir $B_R = A \times (-R,R]^m \subset \mathbb{R}^{k+m}$ und $J_{R,\ell} = I_\ell \times (-R,R]^m \in Q_{k+m}$. Dann gilt $B_R \subset \bigcup_{\ell=1}^\infty J_{R,\ell}$ und $\sum_{\ell=1}^\infty |J_{R,\ell}| = \sum_{\ell=1}^\infty |I_\ell|(2R)^m = (2R)^m\varepsilon = \varepsilon'$. Also ist B_R für jedes $R \in \mathbb{N}$ eine Nullmenge. Nun ist aber $A \times \mathbb{R}^m = \bigcup_{R=1}^\infty B_R$ eine abzählbare Vereinigung von Nullmengen, also selbst eine Nullmenge.

(b) Seien $I = (a,b] \in Q_n$ mit $a,b \in \mathbb{R}^n$ und $e = (1,\dots,1)^{\mathrm{T}} \in \mathbb{R}^n$. Wir definieren $A_\delta = (a, b+\delta e)$ und stellen fest: A_δ ist offen; $I \subset A_\delta$ für alle $\delta > 0$; $\lambda(A_\delta) = \prod_{i=1}^n (b_i + \delta - a_i)$. (Falls letzteres nicht bekannt sein sollte: $J = (c,d)$ mit $c,d \in \mathbb{R}^n$ ist offen und daher messbar, und es gilt $\lambda(J) = \prod_{i=1}^n (d_i - c_i)$, wie die Beziehung $J = \bigcup_{k\in\mathbb{N}} (c, d-(1/k)e]$ und Aufgabe 45. (a) iii. zeigen.)
Mit Aufgabe 45. (a) ii. folgt dann $\lambda(A_\delta \setminus I) = \prod_{i=1}^n (b_i+\delta-a_i) - \prod_{i=1}^n (b_i - a_i)$. Dieser Ausdruck konvergiert aber für $\delta \to 0$ gegen 0, d. h. zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$ mit $\lambda(A_\delta \setminus I) < \varepsilon$. Dann ist $\mathcal{O} = A_\delta$ eine solche gesuchte offene Menge.

45. (a) i. Es ist

$$\bigcap_{k\in\mathbb{N}} A_k = \mathbb{R}^n \setminus \left(\mathbb{R}^n \setminus \bigcap_{k\in\mathbb{N}} A_k\right) = \mathbb{R}^n \setminus \bigcup_{k\in\mathbb{N}} (\mathbb{R}^n \setminus A_k) \in \mathcal{M}\,,$$

denn nach dem Hauptsatz sind Komplemente und abzählbare Vereinigungen messbarer Mengen messbar. (Damit folgt sofort die Messbarkeit von $A\cap B$.) Ferner gilt $A\setminus B = A \cap (\mathbb{R}^n \setminus B) \in \mathcal{M}$.

ii. Aus $A = (A\cap B)\,\dot\cup\,(A\setminus B)$ und $A \cap B = B$ folgt mit der Additivität des Maßes $\lambda(A) = \lambda(B) + \lambda(A\setminus B)$ und daraus die Behauptung, da das Subtrahieren von $\lambda(B) \in \mathbb{R}$ eine Äquivalenzumformung ist.

iii. Wir setzen $B_1 = A_1$ und $B_{k+1} = A_{k+1} \setminus A_k$ für alle $k \in \mathbb{N}$. Dann sind die messbaren Mengen B_i paarweise disjunkt, und es gilt

$$\bigcup_{k=1}^\infty A_k = \bigcup_{k=1}^\infty B_k \quad \text{und} \quad \bigcup_{k=1}^m B_k = A_m\,.$$

Damit folgt

$$\lambda\left(\bigcup_{k\in\mathbb{N}} A_k\right) = \lambda\left(\bigcup_{k\in\mathbb{N}} B_k\right) = \sum_{k\in\mathbb{N}} \lambda(B_k)$$
$$= \lim_{m\to\infty} \sum_{k=1}^m \lambda(B_k) = \lim_{m\to\infty} \lambda\left(\bigcup_{k=1}^m B_k\right) = \lim_{m\to\infty} \lambda(A_m)\,.$$

iv. Aus $A_k \supset A_{k+1}$ folgt $A_1 \setminus A_k \subset A_1 \setminus A_{k+1}$ und mit iii.

$$\lambda\left(\bigcup_{k\in\mathbb{N}} (A_1 \setminus A_k)\right) = \lim_{m\to\infty} \lambda(A_1 \setminus A_m)\,.$$

Formt man die linke Seite nach de Morgan um und benutzt ii., so folgt

$$\lambda(A_1) - \lambda\left(\bigcap_{k\in\mathbb{N}} A_k\right) = \lim_{m\to\infty} \big(\lambda(A_1) - \lambda(A_m)\big)\,.$$

(Man überlege sich, dass die Voraussetzungen für ii. erfüllt sind.) Subtraktion von $\lambda(A_1) < \infty$ und Multiplikation mit -1 liefert die Behauptung.

(b) Wir wählen $A_k = (k, \infty) \subset \mathbb{R}$, also $A_k \supset A_{k+1}$. Diese Mengen sind offen, also $A_k \in \mathcal{M}$, und es gilt $\lambda(A_k) = \infty$ für alle $k \in \mathbb{N}$. Daraus folgt natürlich sofort $\lim_{k\to\infty} \lambda(A_k) = \infty$. Allerdings ist $\bigcap_{k\in\mathbb{N}} A_k = \emptyset$, denn: Für ein x aus dem Schnitt müsste $x \in A_k$ für jedes k gelten, d. h. x müsste größer als jede natürliche Zahl sein, was ein Widerspruch ist. Wegen $\lambda(\emptyset) = 0$ folgt

$$\lambda\left(\bigcap_{k\in\mathbb{N}} A_k\right) \neq \lim_{m\to\infty} \lambda(A_m)\,.$$

46. (a) Für $\lambda = 0$ ist λf die Nullfunktion, d. h. $\{\lambda f > c\}$ ist messbar, nämlich B für $c < 0$ und $\emptyset$ für $c \geq 0$. Ferner ist g genau dann messbar, wenn $-g$ messbar ist, denn $\{g > c\} = \{-g < -c\}$. Im Folgenden sei also o. B. d. A. $\lambda > 0$. Es ist aber $\{\lambda f > c\} = \{f > c/\lambda\}$, und da f messbar ist, ist es auch λf.

(b) Gemäß den Rechenregeln auf $\bar{\mathbb{R}}$ setzen wir

$$(f+g)(x) := \begin{cases} f(x)+g(x) & \text{für } f(x), g(x) \in \mathbb{R}\,, \\ \infty & \text{für } f(x) \in \mathbb{R} \text{ und } g(x) = \infty\,, \\ \infty & \text{für } f(x) = \infty \text{ und } g(x) \in \mathbb{R}\,, \\ \infty & \text{für } f(x) = g(x) = \infty\,, \\ -\infty & \text{für } f(x) \in \mathbb{R} \text{ und } g(x) = -\infty\,, \\ -\infty & \text{für } f(x) = -\infty \text{ und } g(x) \in \mathbb{R}\,, \\ -\infty & \text{für } f(x) = g(x) = -\infty\,. \end{cases}$$

Dies zeigt die Wohldefiniertheit von $f + g$.

Ist $f(x) < g(x)$ für $x \in B$, so gibt es ein $q \in \mathbb{Q}$ mit $f(x) < q < g(x)$, da $\mathbb{Q}$ dicht in $\mathbb{R}$ liegt. Daher gilt $\{f < g\} = \bigcup_{q\in\mathbb{Q}}(\{f < q\} \cap \{q < g\})$. Da f und g messbare Funktionen sind, sind $\{f < q\}$ und $\{q < g\}$ messbare Mengen, ebenso deren Schnitt. $\{f < g\}$ ist also die abzählbare Vereinigung messbarer Mengen und damit selbst messbar.

Nach (a) ist aber auch $-g$ messbar, d. h. $\{-g < c'\}$ ist messbar für alle $c' \in \bar{\mathbb{R}}$, ebenso $\{-g + c < c' + c\}$ für alle $c' + c$, d. h. die Funktion $-g + c$ ist messbar. Also ist $\{f < -g + c\} = \{f + g < c\}$ messbar, und damit auch $f + g$.

47. Seien $f = \sum_{\alpha=1}^{k} y_\alpha \mathbb{1}_{Y_\alpha}$ und $g = \sum_{\beta=1}^{\ell} z_\beta \mathbb{1}_{Z_\beta}$ mit jeweils paarweise verschiedenen Werten $y_1, \ldots, y_k \geq 0$ und $z_1, \ldots, z_\ell \geq 0$ und Zerlegungen $Y_1 \dot\cup \cdots \dot\cup Y_k = Z_1 \dot\cup \cdots \dot\cup Z_\ell = B$ in messbare Mengen. Dann ist

$$\begin{aligned} F = f + g &= \sum_{\alpha=1}^{k} y_\alpha \mathbb{1}_{Y_\alpha} + \sum_{\beta=1}^{\ell} z_\beta \mathbb{1}_{Z_\beta} \\ &= \sum_{\alpha=1}^{k} y_\alpha \sum_{\beta=1}^{\ell} \mathbb{1}_{Y_\alpha \cap Z_\beta} + \sum_{\beta=1}^{\ell} z_\beta \sum_{\alpha=1}^{k} \mathbb{1}_{Y_\alpha \cap Z_\beta} = \sum_{\alpha=1}^{k} \sum_{\beta=1}^{\ell} (y_\alpha + z_\beta) \mathbb{1}_{Y_\alpha \cap Z_\beta} . \end{aligned}$$

Die so entstandenen Funktionswerte $y_\alpha + z_\beta$ von F müssen nicht zwangsläufig paarweise verschieden sein. Um dies zu erreichen, wählen wir $w_1, \ldots, w_m \geq 0$ paarweise verschieden, so dass

$$\bigcup_{\alpha=1}^{k} \bigcup_{\beta=1}^{\ell} \{y_\alpha + z_\beta\} = \{w_1, \ldots, w_m\} ,$$

und setzen für alle $\gamma \in \{1, \ldots, m\}$

$$W_\gamma = \{F = w_\gamma\} = \dot{\bigcup_{\substack{(\alpha,\beta) \in \{1,\ldots,k\} \times \{1,\ldots,\ell\} \\ y_\alpha + z_\beta = w_\gamma}}} (Y_\alpha \cap Z_\beta) .$$

Dann ist $W_1 \dot\cup \cdots \dot\cup W_m = B$, und $F = \sum_{\gamma=1}^{m} w_\gamma \mathbb{1}_{W_\gamma}$ ist eine Elementarfunktion, weil W_γ als endliche Vereinigung von Schnitten endlich vieler messbarer Mengen messbar ist.

Es folgt (mit sehr ausführlicher Rechnung)

$$\begin{aligned} &\sum_{w \in F(B)} w \lambda\big(F^{-1}(w)\big) = \sum_{\gamma=1}^{m} w_\gamma \lambda\big(F^{-1}(w_\gamma)\big) = \sum_{\gamma=1}^{m} w_\gamma \lambda(W_\gamma) \\ &= \sum_{\gamma=1}^{m} w_\gamma \sum_{\substack{(\alpha,\beta) \in \{1,\ldots,k\} \times \{1,\ldots,\ell\} \\ y_\alpha + z_\beta = w_\gamma}} \lambda(Y_\alpha \cap Z_\beta) = \sum_{(\alpha,\beta) \in \{1,\ldots,k\} \times \{1,\ldots,\ell\}} (y_\alpha + z_\beta) \lambda(Y_\alpha \cap Z_\beta) \\ &= \sum_{\alpha=1}^{k} \sum_{\beta=1}^{\ell} y_\alpha \lambda(Y_\alpha \cap Z_\beta) + \sum_{\alpha=1}^{k} \sum_{\beta=1}^{\ell} z_\beta \lambda(Y_\alpha \cap Z_\beta) \\ &= \sum_{\alpha=1}^{k} y_\alpha \lambda\left(\dot{\bigcup_{\beta=1}^{\ell}} (Y_\alpha \cap Z_\beta) \right) + \sum_{\beta=1}^{\ell} z_\beta \lambda\left(\dot{\bigcup_{\alpha=1}^{k}} (Y_\alpha \cap Z_\beta) \right) \\ &= \sum_{\alpha=1}^{k} y_\alpha \lambda(Y_\alpha) + \sum_{\beta=1}^{\ell} z_\beta \lambda(Z_\beta) = \sum_{\alpha=1}^{k} y_\alpha \lambda\big(f^{-1}(y_\alpha)\big) + \sum_{\beta=1}^{\ell} z_\beta \lambda\big(g^{-1}(z_\beta)\big) \\ &= \sum_{y \in f(B)} y \lambda\big(f^{-1}(y)\big) + \sum_{z \in g(B)} z \lambda\big(g^{-1}(z)\big) . \end{aligned}$$

Obwohl alle vorkommenden Summationen endlich sind, können die Werte der Summen unendlich werden, weil die Maße der Mengen unendlich sein können. Da aber alle y_α, z_β und w_γ nichtnegativ sind, treten keine undefinierten Ausdrücke wie z. B. $\infty - \infty$ auf.

Bemerkung: Die Nichtnegativität von f und g ist notwendig, wie das Gegenbeispiel $B = \mathbb{R}$, $f \equiv a > 0$ und $g \equiv b < 0$ beweist. Rechts entsteht dann der undefinierte Ausdruck $\infty - \infty$.

48. (a) Wir betrachten die Mengen $U(f) = \{(x^{\mathrm{T}}, y)^{\mathrm{T}} \in \mathbb{R}^{n+1} \mid x \in \mathbb{R}^n\,,\ 0 < y < f(x)\}$ und $U(g) = \{(x^{\mathrm{T}}, y)^{\mathrm{T}} \in \mathbb{R}^{n+1} \mid x \in \mathbb{R}^n\,,\ 0 < y < g(x)\}$. Wegen $f \le g$ fast überall ist $N = U(f) \setminus U(g)$ eine Nullmenge. Offenbar ist $U(f) \subset U(g) \,\dot\cup\, N$, d. h.

$$\int f = \lambda\big(U(f)\big) \le \lambda\big(U(g) \,\dot\cup\, N\big) = \lambda\big(U(g)\big) + \lambda(N) = \lambda\big(U(g)\big) = \int g\,.$$

(b) Aus $f = g$ fast überall folgt $f \le g$ fast überall und $g \le f$ fast überall, also mit (a) $\int f \le \int g$ und $\int g \le \int f$, d. h. zusammen $\int f = \int g$.

(c) Für $\mu = 0$ ist nichts zu zeigen; gelte also $\mu > 0$. Sei zunächst $f \in \mathcal{E}(\mathbb{R}^n)$ eine nicht-negative Elementarfunktion, d. h. $f = \sum_{k=1}^m y_k \mathbb{1}_{B_k}$ mit paarweise verschiedenen Zahlen $y_1, \ldots, y_m \ge 0$ und paarweise disjunkten, messbaren Mengen $B_1, \ldots, B_m \subset \mathbb{R}^n$. Dann ist $\mu f = \sum_{k=1}^m (\mu y_k) \mathbb{1}_{B_k}$ ebenfalls eine Elementarfunktion, und es gilt

$$\int \mu f = \sum_{k=1}^m (\mu y_k) \lambda(B_k) = \mu \sum_{k=1}^m y_k \lambda(B_k) = \mu \int f\,.$$

Sei nun $(f_\nu)_\nu$ eine Folge nicht-negativer Elementarfunktionen $f_\nu \in \mathcal{E}(\mathbb{R}^n)$, die monoton wachsend gegen f konvergiert. Dann konvergiert $(\mu f_\nu)_\nu$ monoton wachsend gegen μf, d. h. μf ist messbar. Mit dem Satz von der monotonen Konvergenz konvergiert schließlich die linke Seite von $\int \mu f_\nu = \mu \int f_\nu$ gegen $\int \mu f$ und die rechte gegen $\mu \int f$, was die Behauptung beweist.

49. (a) i. Dies ist gerade die Definition von relativ offen. Aber wir zeigen es noch einmal explizit: Seien also $x \in \mathbb{R}^k$ und $y \in A_x$. (Selbstverständlich kann A_x leer sein, aber dann ist nichts zu zeigen.) Dann ist $z = (x^{\mathrm{T}}, y^{\mathrm{T}})^{\mathrm{T}} \in A$, und da A offen ist, gibt es eine offene Umgebung um z in A, die o. B. d. A. von der Form $U \times V$ ist, d. h. $U \times V \subset A$ mit $U \subset \mathbb{R}^k$ und $V \subset \mathbb{R}^m$. Dann ist aber U eine offene Umgebung von x in A_x, d. h. A_x ist offen im $\mathbb{R}^m$.

ii. Sei A eine G_δ-Menge, d. h. es gebe offene Mengen $U_1, U_2, \ldots \subset \mathbb{R}^n$ mit $A = \bigcap_{i \in \mathbb{N}} U_i$. Nach i. ist dann aber auch jedes $(U_i)_x$ offen im $\mathbb{R}^m$, d. h. $\bigcap_{i \in \mathbb{N}} (U_i)_x = A_x$ ist ebenfalls eine G_δ-Menge.

(b) Sei zunächst $f \in \mathcal{E}(\mathbb{R}^n)$ nicht-negativ, d. h. es gebe paarweise verschiedene Zahlen $w_1, \ldots, w_\ell \ge 0$ und paarweise disjunkte, messbare Mengen $B_1, \ldots, B_\ell$, so dass $f = \sum_{\alpha=1}^\ell w_\alpha \mathbb{1}_{B_\alpha}$. Dann gilt $f_x = \sum_{\alpha=1}^\ell w_\alpha \mathbb{1}_{(B_\alpha)_x}$. Nach dem Prinzip von Cavalieri ist nun $(B_\alpha)_x$ für fast alle $x \in \mathbb{R}^k$ messbar, d. h. $(B_\alpha)_x$ ist messbar für alle $x \in \mathbb{R}^k \setminus N_\alpha$ mit einer Nullmenge $N_\alpha \subset \mathbb{R}^k$. Damit ist f_x messbar für alle $x \in \bigcap_{\alpha=1}^\ell (\mathbb{R}^k \setminus N_\alpha) = \mathbb{R}^k \setminus N$ mit einer Nullmenge $N = \bigcup_{\alpha=1}^\ell N_\alpha \subset \mathbb{R}^k$.

Wir erhalten also: Ist $f \in \mathcal{E}(\mathbb{R}^n)$ eine nicht-negative Elementarfunktion, so ist die durch $f_x(y) = f(x, y)$ gegebene Funktion $f_x\colon \mathbb{R}^m \to [0, \infty]$ für fast alle $x \in \mathbb{R}^k$ messbar und wegen des endlichen Bildes eine Elementarfunktion.

Seien nun $f\colon \mathbb{R}^n \to [0, \infty]$ messbar und $(f_\nu)_\nu$ eine Folge von nicht-negativen Elementarfunktionen $f_\nu \in \mathcal{E}(\mathbb{R}^n)$ mit $f_\nu \to f$ monoton wachsend. Dann gibt es gemäß der vorigen Rechnung Nullmengen $N_\nu \subset \mathbb{R}^k$, so dass $(f_\nu)_x \in \mathcal{E}(\mathbb{R}^m)$ für alle $x \in \mathbb{R}^k \setminus N_\nu$. Das bedeutet aber, dass $(f_\nu)_x \to f_x$ monoton wachsend für alle

$x \in \mathbb{R}^k \setminus N$ mit $N = \bigcup_{\nu\in\mathbb{N}} N_\nu$. Da N als abzählbare Vereinigung von Nullmengen wieder eine Nullmenge ist, sind fast alle f_x die Grenzfunktion monoton wachsender Elementarfunktionen. Also ist f_x messbar für fast alle $x \in \mathbb{R}^k$.

50. (a) Seien $I_k \in Q_n$ für $k \in \mathbb{N}$ halboffene Quader mit $\bigcup_{k\in\mathbb{N}} I_k \supset A$. Dann sind auch rI_k halboffene Quader, und diese überdecken die Menge rA, denn zu $x \in A$ gibt es ein $k \in \mathbb{N}$, so dass $x \in I_k$, und dann folgt $rx \in rI_k$. Ferner gilt

$$\sum_{k=1}^{\infty} |rI_k| = \sum_{k=1}^{\infty}\prod_{i=1}^{n}(rb_{k,i} - ra_{k,i}) = r^n \sum_{k=1}^{\infty}\prod_{i=1}^{n}(b_{k,i} - a_{k,i}) = r^n \sum_{k=1}^{\infty} |I_k|\,,$$

wobei beide Seiten auch unendlich sein dürfen.

Daraus erhalten wir: Gibt es zu einem $M \in [0,\infty]$ eine Überdeckung $(I_k)_k$ von A mit $\sum_{k\in\mathbb{N}} |I_k| \leq M$, dann gibt es eine Überdeckung von rA mit Inhalt höchstens $r^n M$, nämlich z. B. $(rI_k)_k$. Das bedeutet $\lambda^*(rA) \leq r^n\lambda^*(A)$. Damit gilt

$$\lambda^*(A) = \lambda^*\left(\frac{1}{r}rA\right) \leq \left(\frac{1}{r}\right)^n \lambda^*(rA) \leq \frac{1}{r^n} r^n \lambda^*(A) = \lambda^*(A)\,,$$

wobei wieder beide Seiten unendlich sein dürfen. Aus der Rechnung folgt Gleichheit an jeder Stelle, also $\lambda^*(rA) = r^n\lambda^*(A)$.

(b) Sind die I_k wie in (a) gewählt, so sind auch $y + I_k$ halboffene Quader, diese überdecken die Menge $y + A$, und es gilt

$$\sum_{k=1}^{\infty} |y + I_k| = \sum_{k=1}^{\infty}\prod_{i=1}^{n}\big((y_i + b_{k,i}) - (y_i + a_{k,i})\big) = \sum_{k=1}^{\infty}\prod_{i=1}^{n}(b_{k,i} - a_{k,i}) = \sum_{k=1}^{\infty} |I_k|\,,$$

wobei beide Seiten auch unendlich sein dürfen. Mit derselben Argumentation wie in (a) folgt $\lambda^*(y + A) \leq \lambda^*(A)$, und die Gleichheit folgt ebenfalls wie in (a) aus der Rechnung $\lambda^*(A) = \lambda^*(y + (-y) + A) \leq \lambda^*((-y) + A) \leq \lambda^*(A)$.

Bemerkung: Diese Eigenschaft nennt man Translationsinvarianz des äußeren Maßes bzw. des Lebesgue-Maßes. Man kann noch – aber aufwendiger – zeigen, dass $\lambda^*(RA) = \lambda^*(A)$ für jede Matrix $R \in \mathrm{SO}(n)$. Beides zusammen bezeichnet man dann als Bewegungsinvarianz.

(c) Sind die $I_k = (a_k, b_k]$ mit $a_k, b_k \in \mathbb{R}^n$ wiederum wie in (a) gewählt, so ist auch

$$(I_k)_\pi = \prod_{i=1}^{n}\big(a_{k,\pi(i)}, b_{k,\pi(i)}\big] \in Q_n$$

für alle $\pi \in S_n$ ein halboffener Quader, und es gilt $|(I_k)_\pi| = |I_k|$ wegen der Kommutativität des Produkts. Mit derselben Argumentation wie in (a) folgt $\lambda^*(A_\pi) \leq \lambda^*(A)$, und die Gleichheit folgt ebenfalls wie in (a) aus der Rechnung $\lambda^*(A) = \lambda^*(A_{\pi^{-1}\circ\pi}) \leq \lambda^*(A_\pi) \leq \lambda^*(A)$.

(d) Mit Cavalieri gilt zunächst

$$\lambda^{n+1}(K_A) = \int \lambda^n\big((K_A)_x\big)\,\mathrm{d}x\,,$$

wobei die Menge $(K_A)_x$ für fast alle $x \in \mathbb{R}$ messbar ist. Weiter ist

$$(K_A)_x = \big\{y \in \mathbb{R}^n \;\big|\; (x, y^{\mathrm{T}})^{\mathrm{T}} \in K_A\big\}$$

$$\begin{aligned}
&= \{y \in \mathbb{R}^n \mid x = x' \text{ und } y = (1-x')y' \text{ für ein } y' \in A \text{ und } 0 \le x' \le 1\} \\
&= \{y \in \mathbb{R}^n \mid y = (1-x)y' \text{ für ein } y' \in A\} \\
&= (1-x)A
\end{aligned}$$

mit der Notation aus (a). Es gilt also $\lambda^*((K_A)_x) = (1-x)^n \lambda^*(A)$, und für fast alle $x \in [0,1]$ gilt die Gleichheit auch für λ^n statt λ^*. Wir erhalten also

$$\lambda^{n+1}(K_A) = \int_0^1 (1-x)^n \lambda^n(A)\,\mathrm{d}x = -\frac{(1-x)^{n+1}}{n+1}\lambda^n(A)\Big|_{x=0}^{x=1} = \frac{1}{n+1}\lambda^n(A)\,.$$

51. (a) Es sei $A_{x,y} = \{z \in \mathbb{R} \mid (x,y,z)^{\mathrm{T}} \in A\} = [0, 1-x-y]$ mit $\lambda^1(A_{x,y}) = 1-x-y$. Setzen wir weiter $A_x = \{(y,z)^{\mathrm{T}} \in \mathbb{R}^2 \mid (x,y,z)^{\mathrm{T}} \in A\}$, so folgt mit Cavalieri

$$\begin{aligned}
\lambda^2(A_x) &= \int \lambda^1(A_{x,y})\,\mathrm{d}y = \int_0^{1-x} (1-x-y)\,\mathrm{d}y = \left((1-x)y - \frac{y^2}{2}\right)\Big|_{y=0}^{y=1-x} \\
&= (1-x)(1-x) - \frac{(1-x)^2}{2} = \frac{(1-x)^2}{2}\,.
\end{aligned}$$

Ein weiteres Mal mit Cavalieri ergibt sich

$$\lambda^3(A) = \int \lambda^2(A_x)\,\mathrm{d}x = \int_0^1 \frac{(1-x)^2}{2}\,\mathrm{d}x = -\frac{(1-x)^3}{6}\Big|_{x=0}^{x=1} = \frac{1}{6}\,.$$

Bemerkung: Statt dieser ausführlichen Darstellung schreibt man häufig nur

$$\lambda^3(A) = \int_0^1 \int_0^{1-x} \int_0^{1-x-y} 1\,\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x \overset{\text{s.o.}}{=} \frac{1}{6}$$

und argumentiert mit dem Satz von Fubini, der ein mehrdimensionales Integral in ein sog. iteriertes Integral aus eindimensionalen Integralen üerführt.

(b) i. Es sei

$$K_2(r)_x = \{y \in \mathbb{R} \mid (x,y)^{\mathrm{T}} \in K_2(r)\} = \left[-\sqrt{r^2-x^2}, \sqrt{r^2-x^2}\,\right]$$

mit $\lambda^1(K_2(r)_x) = 2\sqrt{r^2-x^2}$. Dann folgt mit Cavalieri

$$\lambda^2\big(K_2(r)\big) = \int \lambda^1\big(K_2(r)_x\big)\,\mathrm{d}x = \int_{-r}^{r} 2\sqrt{r^2-x^2}\,\mathrm{d}x = 4\int_0^r \sqrt{r^2-x^2}\,\mathrm{d}x$$

aus Symmetriegründen. Mit der Substitution $x = r\sin t$ ergibt sich

$$\lambda^2\big(K_2(r)\big) = 4\int_0^{\pi/2} \sqrt{r^2 - r^2\sin^2 t}\; r\cos t\,\mathrm{d}t = 2r^2 \int_0^{\pi/2} 2\cos^2 t\,\mathrm{d}t\,.$$

Eine Stammfunktion von $2\cos^2 t$ ist $t + \cos t \sin t$, wie man durch Ableiten schnell nachrechnet. Damit folgt schließlich

$$\lambda^2\big(K_2(r)\big) = 2r^2(t + \cos t\sin t)\Big|_{t=0}^{t=\pi/2} = \pi r^2\,.$$

ii. Analog sei

$$\begin{aligned} K_3(r)_x &= \big\{ (y,z)^{\mathrm{T}} \in \mathbb{R}^2 \bigm| (x,y,z)^{\mathrm{T}} \in K_3(r) \big\} \\ &= \big\{ (y,z)^{\mathrm{T}} \in \mathbb{R}^2 \bigm| y^2 + z^2 \le r^2 - x^2 \big\} = K_2\big(\sqrt{r^2 - x^2}\,\big)\,. \end{aligned}$$

Mit Cavalieri und dem Ergebnis aus i. folgt dann

$$\begin{aligned} \lambda^3\big(K_3(r)\big) &= \int \lambda^2\big(K_3(r)_x\big)\,\mathrm{d}x = \int_{-r}^{r} \pi(r^2 - x^2)\,\mathrm{d}x = 2\pi \int_0^r (r^2 - x^2)\,\mathrm{d}x \\ &= 2\pi\left(r^2 x - \frac{x^3}{3}\right)\Bigg|_{x=0}^{x=r} = \frac{4}{3}\pi r^3\,. \end{aligned}$$