Analysis 2

Franz Merkl

(sehr vorläufige Version [1], 26. Juni 2017)

Inhaltsverzeichnis

[1]Dies ist nur ein Entwurf eines Analysis 2 Skripts. Ohne jede Garantie. Für Hinweise auf Fehler aller Art ist der Autor dankbar. Vielen Dank an Frau E. Roth und Herrn T. Simonis für Korrekturhinweise!

Die Analysis 2 und 3 sind gewissermaßen parallel zur Analysis 1 aufgebaut: Erst Topologie, dann Differentialrechnung (in der Analysis 2), schließlich (in der Analysis 3) Integralrechnung. Dabei werden die meisten Begriffe, die Sie in der Analysis 1 kennengelernt haben, weitgehend verallgemeinert und abstrahiert, insbesondere auf höhere Dimensionen. Naturgemäß brauchen wir bei der analytischen Arbeit in höherdimensionalen Vektorräumen Hilfsmittel aus der Linearen Algebra.

1 Metrische und topologische Räume

1.1 Definition metrischer und normierter Räume

In der Analysis 1 spielte der Abstand $|x - y|$ zweier reeller oder komplexer Zahlen x und y sowie die dafür geltende Dreiecksungleichung

$$|x - z| \leq |x - y| + |y - z| \quad (x, y, z \in \mathbb{C})$$

eine zentrale Rolle, z.B. bei Konvergenz- und Stetigkeitsargumenten. Die meisten dieser Argumente lassen sich auch auf folgenden verallgemeinerten Abstandsbegriff anwenden:

Definition 1.1 (Halbmetrische und metrische Räume) Es sei M eine Menge. Eine Abbildung $d : M \times M \to \mathbb{R}$ heißt *Halbmetrik* (synonym: *Semimetrik*) auf M, wenn gilt:

1. *Nichtnegativität:* $\forall x, y \in M : \ d(x, y) \geq 0$.

2. *Symmetrie:* $\forall x, y \in M : \ d(x, y) = d(y, x)$.

3. *Dreiecksungleichung:* $\forall x, y, z \in M : \ d(x, z) \leq d(x, y) + d(y, z)$.

4. *Selbstabstand Null:* $\forall x \in M : \ d(x, x) = 0$.

Das Paar (M, d) heißt dann ein *halbmetrischer Raum* (synonym: *semimetrischer Raum*). Die Zahl $d(x, y)$ wird *Abstand* zwischen den zwei Punkten x und y in M bezüglich d genannt.
Gilt zusätzlich

5. $\forall x, y \in M : (d(x, y) = 0 \Rightarrow x = y)$,

so nennen wir d eine *Metrik* auf M und das Paar (M, d) einen *metrischen Raum.*

Beispiel 1.2

1. Ist $M = \mathbb{R}$ oder $M = \mathbb{C}$, so wird durch den in der Analysis 1 verwendeten Abstand $d : M \times M \to \mathbb{R}$, $d(x, y) = |x - y|$, eine Metrik auf M definiert. Wir nennen sie die "Standardmetrik" auf $\mathbb{R}$ bzw. auf $\mathbb{C}$.

2. Bezeichnet

$$\mathrm{Arctan} : \mathbb{R} \cup \{\pm\infty\} \to [-\frac{\pi}{2}, \frac{\pi}{2}], \ \mathrm{Arctan}\, x = \begin{cases} \arctan x & \text{für } x \in \mathbb{R}, \\ \frac{\pi}{2} & \text{für } x = +\infty, \\ -\frac{\pi}{2} & \text{für } x = -\infty \end{cases}$$

die auf $\mathbb{R} \cup \{\pm\infty\}$ stetig fortgesetzte Arcustangensfunktion, so wird durch

$$d : (\mathbb{R} \cup \{\pm\infty\}) \times (\mathbb{R} \cup \{\pm\infty\}) \to \mathbb{R},\ d(x, y) := |\operatorname{Arctan} x - \operatorname{Arctan} y|$$

eine Metrik auf $\mathbb{R} \cup \{\pm\infty\}$ definiert.

3. **p-adische Metrik auf $\mathbb{Z}$.** Ist p eine Primzahl und $k \in \mathbb{Z} \setminus \{0\}$, so bezeichne $v_p(k) \in \mathbb{N}_0$ die Vielfachheit, mit der p in der Primfaktorzerlegung von k vorkommt. Das bedeutet: k kann in der Form $k = p^{v_p(k)}m$ mit einer Zahl $m \in \mathbb{Z}$, die nicht durch p teilbar ist, geschrieben werden. Zum Beispiel besitzt die Zahl 360 die Primfaktorzerlegung $360 = 2^3 \cdot 3^2 \cdot 5^1$, also ist $v_2(360) = 3$, $v_3(360) = 2$, $v_5(360) = 1$ und $v_p(360) = 0$ für alle anderen Primzahlen p. Wir setzen noch formal $v_p(0) := \infty$ und $p^{-\infty} := 0$. Durch
$$d_p : \mathbb{Z} \times \mathbb{Z} \to \mathbb{R},\ d_p(x, y) = p^{-v_p(x-y)}$$
wird eine Metrik auf $\mathbb{Z}$ gegeben, die *p-adische Metrik* genannt wird. Diese Metrik spielt in der Zahlentheorie eine wichtige Rolle. Um die Dreiecksungleichung in diesem Beispiel einzusehen, beachte man, dass für $a, b \in \mathbb{Z}$ gilt:
$$v_p(a + b) \geq \min\{v_p(a), v_p(b)\},$$
denn wenn jede der Zahlen a und b mindestens v-mal durch p teilbar ist, so ist auch $a + b$ mindestens v-mal durch p teilbar. Es folgt
$$p^{-v_p(a+b)} \leq \max\{p^{-v_p(a)}, p^{-v_p(b)}\}$$
und damit für $x, y, z \in \mathbb{Z}$ die "verschärfte Dreiecksungleichung", auch "ultrametrische Ungleichung" genannt:
$$d_p(x, z) \leq \max\{d_p(x, y), d_p(y, z)\} \leq d_p(x, y) + d_p(y, z)$$
für $x, y, z \in \mathbb{Z}$.

Im Folgenden steht das Symbol $\mathbb{K}$ stets entweder für $\mathbb{R}$ oder für $\mathbb{C}$. In Verallgemeinerung des Absolutbetrags über $\mathbb{K}$ führen wir den Begriff der *Norm* über einem Vektorraum über $\mathbb{K}$ ein:

Definition 1.3 (Halbnormen und Normen) Es sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}$ oder über $\mathbb{K} = \mathbb{C}$. Eine Abbildung

$$\|\cdot\| : V \to \mathbb{R}, \quad x \mapsto \|x\|$$

heißt *Halbnorm* (synonym: *Seminorm*) auf V, wenn gilt:

1. *Nichtnegativität:* $\forall x \in V : \ \|x\| \geq 0$,

2. *Homogenität:* $\forall x \in V\ \forall \lambda \in \mathbb{K} : \ \|\lambda x\| = |\lambda|\, \|x\|$,

3. *Dreiecksungleichung:* $\forall x, y \in V: \; \|x+y\| \leq \|x\| + \|y\|$.

In diesem Fall nennen wir das Paar $(V, \|\cdot\|)$ auch einen *halbnormierten Raum*.
Gilt zusätzlich

4. $\forall x \in V: \; (\|x\| = 0 \Rightarrow x = 0)$,

so nennen wir die Abbildung $\|\cdot\| : V \to \mathbb{R}$ eine *Norm* auf V. In diesem Fall heißt das Paar $(V, \|\cdot\|)$ ein *normierter Raum*.

Ähnlich, wie der Absolutbetrag zur Definition der Standardmetrik auf $\mathbb{K}$ verwendet wird, erzeugt eine Halbnorm eine Halbmetrik:

Lemma 1.4 (von einer Norm erzeugte Metrik) *Ist $\|\cdot\|$ eine Halbnorm auf V, so wird durch*

$$d : V \times V \to \mathbb{R}, \quad d(x,y) = \|x-y\|$$

eine Halbmetrik auf V gegeben. Sie wird die durch $\|\cdot\|$ erzeugte Halbmetrik genannt. Die von $\|\cdot\|$ erzeugte Halbmetrik d ist genau dann eine Metrik, wenn $\|\cdot\|$ eine Norm ist.

Beweis: Die Nichtnegativität von d folgt unmittelbar aus der Nichtnegativität von $\|\cdot\|$. Zur Symmetrie von d verwenden wir die Homogenität der Halbnorm wie folgt: Für $x, y \in V$ gilt:

$$d(x,y) = \|x-y\| = \|-1 \cdot (y-x)\| = |-1| \, \|y-x\| = \|y-x\| = d(y,x).$$

Die Dreiecksungleichung für d folgt so aus der Dreiecksungleichung der Halbnorm: Für $x, y, z \in V$ gilt:

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

Zum Selbstabstand 0 verwenden wir nochmal die Homogenität der Halbnorm: Für $x \in V$ gilt:

$$d(x,x) = \|x-x\| = \|0\| = \|0 \cdot 0\| = |0| \, \|0\| = 0.$$

Damit ist gezeigt, dass d eine Halbmetrik ist. Ist nun $\|\cdot\|$ sogar eine Norm, so gilt für $x, y \in V$ mit $d(x,y) = 0$ auch $\|x-y\| = 0$, also $x - y = 0$ und daher $x = y$; also ist hier d eine Metrik. Ist umgekehrt d eine Metrik, so folgt für alle $x \in V$: mit $\|x\| = 0$ auch

$$d(x,0) = \|x-0\| = \|x\| = 0,$$

also $x = 0$. Dann ist $\|\cdot\|$ sogar eine Norm.

$\square$

Lemma 1.5 (Variante der Dreiecksungleichung) *Ist $(V, \|\cdot\|)$ einen halbnormierter Raum und gilt $x, y \in V$, so folgt*

$$\big| \|x\| - \|y\| \big| \leq \|x-y\|$$

Beweis: Die Dreiecksungleichung liefert

$$\|x\| - \|y\| \leq \|x - y\|$$

und

$$\|y\| - \|x\| \leq \|y - x\| = \|x - y\|,$$

also zusammen die Behauptung.

□

Beispiel 1.6 1. Der Absolutbetrag $|\cdot| : \mathbb{K} \to \mathbb{R}$ ist eine Norm auf $\mathbb{K}$, aufgefasst als Vektorraum über $\mathbb{K}$. Die davon erzeugte Metrik ist die Standardmetrik auf $\mathbb{K}$.

2. Ist V ein $\mathbb{K}$-Vektorraum und ist $f : V \to \mathbb{K}$ eine Linearform, also eine $\mathbb{K}$-lineare Abbildung mit Werten im Körper $\mathbb{K}$, so ist

$$\|\cdot\| : V \to \mathbb{R}, \quad \|x\| := |f(x)|$$

eine Seminorm auf V. Ist V mindestens zweidimensional, so ist diese Seminorm jedoch keine Norm. Zum Beispiel ist für $n \in \mathbb{N}$ mit $n \geq 2$ und $1 \leq i \leq n$ die Abbildung

$$N_i : \mathbb{K}^n \ni (x_1, \ldots, x_n) \mapsto |x_i|$$

eine Seminorm, aber keine Norm. Zum Beispiel bildet N_i den j-ten kanonischen Einheitsvektor $e_j = (\delta_{jm})_{m=1,\ldots,n}$ für $j \neq i$ auf 0 ab.

3. Für $n \in \mathbb{N}$ werden durch

$$\|\cdot\|_1 : \mathbb{K}^n \to \mathbb{R},\ \|(x_1, \ldots, x_n)\|_1 := \sum_{j=1}^{n} |x_j|$$

und

$$\|\cdot\|_\infty : \mathbb{K}^n \to \mathbb{R},\ \|(x_1, \ldots, x_n)\|_\infty := \max\{|x_1|, \ldots, |x_n|\}$$

zwei Normen auf $\mathbb{K}^n$ definiert. (Übung)

4. *Das folgende Beispiel ist eine abstraktere Variante des vorhergehenden Beispiels: Statt Vektoren in $\mathbb{K}^n$ betrachten wir Funktionen auf einem Intervall $[a, b]$ und statt Summen Integrale. Wir arbeiten hier also mit einem Funktionenraum, also höher in der Mengenhierarchie als bisher. Der höhere Abstraktionsgrad ist vielleicht gewöhnungsbedürftig.*
Ist $[a, b]$ ein abgeschlossenes Intervall mit reellen Zahlen $a < b$, so bezeichnet $C([a, b], \mathbb{K})$ die Menge aller stetigen Funktionen $f : [a, b] \to \mathbb{K}$. Durch

$$\|\cdot\|_\infty : C([a, b], \mathbb{K}) \to \mathbb{R},\ \|f\|_\infty := \max\{|f(x)| \mid x \in [a, b]\}$$

wird eine Norm auf $C([a, b], \mathbb{K})$ definiert. Zur Wohldefiniertheit beachte man, dass für jedes $f \in C([a, b], \mathbb{K})$ auch $|f|$ stetig ist, also ein Maximum auf der nichtleeren kompakten Menge $[a, b]$ annimmt.

Lemma 1.7 *Es sei* $\|\cdot\|$ *eine beliebige Norm auf* $\mathbb{K}^n$, $n \in \mathbb{N}$. *Dann gilt*

$$\forall x \in \mathbb{K}^n : \ \|x\| \le C\|x\|_\infty$$

wobei

$$C := \sum_{j=1}^{n} \|e_j\| > 0$$

mit den kanonischen Einheitsvektoren $e_j = (\delta_{j,k})_{k=1,\dots,n}$ *in* $\mathbb{K}^n$ *definiert wird.*

Beweis: Für $x = (x_1, \dots, x_n) \in \mathbb{K}^n$ folgt:

$$\|x\| = \left\|\sum_{j=1}^{n} x_j e_j\right\| \le \sum_{j=1}^{n} \|x_j e_j\| = \sum_{j=1}^{n} |x_j| \ \|e_j\| \le \max_{j=1,\dots,n} |x_j| \sum_{j=1}^{n} \|e_j\| = \|x\|_\infty \sum_{j=1}^{n} \|e_j\|.$$

□

Erst später (in Satz 1.169 unten) werden wir sehen, dass auch umgekehrt

$$\exists C' > 0 \ \forall x \in \mathbb{K}^n : \ \|x\|_\infty \le C'\|x\|$$

gilt.

Übung 1.8 (Integralversion der Dreiecksungleichung) Es seien $a \le b$ zwei reelle Zahlen, $n \in \mathbb{N}$, $f_1, \dots, f_n : [a, b] \to \mathbb{K}$ stetige Funktionen und $f : [a, b] \to \mathbb{K}^n$, $f(x) = (f_k(x))_{k=1,\dots,n}$. Wir kürzen ab:

$$\int_a^b f(x)\,dx := \left(\int_a^b f_k(x)\,dx\right)_{k=1,\dots,n}.$$

Weiter bezeichne $\|\cdot\|$ eine beliebige Norm auf $\mathbb{K}^n$. Zeigen Sie: Die Abbildung $[a, b] \ni x \mapsto \|f(x)\| \in \mathbb{R}$ ist stetig, und es gilt

$$\left\|\int_a^b f(x)\,dx\right\| \le \int_a^b \|f(x)\|\,dx.$$

Approximieren Sie dazu die Integrale durch Integrale von Treppenfunktionen, also durch Summen und wenden Sie die Dreiecksungleichung an.
(Eine analoge Integralversion der Dreiecksungleichung gilt auch für Riemann-integrierbare Funktionen.)

"Abstand 0" zu haben, ist eine Äquivalenzrelation. In der Tat:

Lemma 1.9 (Verkleben von Punkten mit Abstand 0**)** *1. Es sei* (M, d) *ein halbmetrischer Raum. Die zweistellige Relation* $\sim$ *auf* M, *gegeben durch*

$$x \sim y :\Leftrightarrow d(x, y) = 0 \ \ \textit{für } x, y \in M$$

ist eine Äquivalenzrelation.

2. Bezeichnet

$$[x]_\sim := \{y \in M | \, y \sim x\}$$

die Äquivalenzklasse von $x \in M$ und

$$M/{\sim} = \{[x]_\sim | \, x \in M\}$$

den Quotientenraum von M nach $\sim$, so ist die Abbildung

$$d' : (M/{\sim}) \times (M/{\sim}) \to \mathbb{R}, \; d'([x]_\sim, [y]_\sim) := d(x, y)$$

wohldefiniert. Die Abbildung d' ist eine Metrik auf $M/{\sim}$.

3. Ist M sogar ein $\mathbb{K}$-Vektorraum und wird die Halbmetrik d auf M von einer Halbnorm $\|\cdot\|$ erzeugt, so ist $N := [0]_\sim$ ein Untervektorraum von M. Er wird "Nullraum" zu $\|\cdot\|$ genannt. In diesem Fall ist $[x]_\sim = x + N$ für alle $x \in M$ ein affiner Unterraum von M, und es wird damit $M/{\sim} = M/N$.

Beweis: 1. *Reflexivität:* $x \sim x$ gilt für $x \in M$ wegen $d(x, x) = 0$.
Symmetrie: Für $x, y \in M$ gilt $x \sim y \Rightarrow y \sim x$ wegen $d(x, y) = d(y, x)$.
Transitivität: Es seien $x, y, z \in M$ mit $x \sim y$ und $y \sim z$ gegeben. Dann folgt mit der Dreiecksungleichung:

$$0 \le d(x, z) \le d(x, y) + d(y, z) = 0 + 0 = 0,$$

also $d(x, z) = 0$, d. h. $x \sim z$.
2. Es seien $x, x', y, y' \in M$ mit $x \sim x'$ und $y \sim y'$ gegeben. Zur Wohldefiniertheit von d' ist zu zeigen: $d(x, y) = d(x', y')$. Wir zeigen $d(x, y) \le d(x', y')$; die umgekehrte Ungleichung $d(x', y') \le d(x, y)$ folgt ebenso durch Vertauschen von x, y mit x', y'. Mit der Dreiecksungleichung und mit $d(x, x') = 0 = d(y, y') = d(y', y)$ erhalten wir:

$$d(x, y) \le d(x, x') + d(x', y) = d(x', y) \le d(x', y') + d(y', y) = d(x', y'),$$

Damit ist die Behauptung $d(x, y) \le d(x', y')$ und damit die Wohldefiniertheit von d' gezeigt. Nichtnegativität, Symmetrie, Dreiecksungleichung und Selbstabstand Null für d' folgt unmittelbar aus den entsprechenden Eigenschaften von d. Also ist d' eine Halbmetrik. Sind nun $x, y \in M$ mit $d'([x]_\sim, [y]_\sim) = 0$ gegeben, so folgt $d(x, y) = 0$, also $x \sim y$ und daher $[x]_\sim = [y]_\sim$. Daher ist d' sogar eine Metrik.
3. Aus der Definition von N folgt $N = \{x \in M | \, \|x\| = 0\}$. Es ist $0 \in N$ wegen $\|0\| = 0$. Weiter gilt für $x \in N$ und $\lambda \in \mathbb{K}$ auch

$$\|\lambda x\| = |\lambda| \, \|x\| = |\lambda| 0 = 0,$$

also $\lambda x \in N$. Schließlich folgt aus $x, y \in N$ auch

$$0 \le \|x + y\| \le \|x\| + \|y\| = 0 + 0 = 0,$$

also $\|x+y\| = 0$ und damit $x+y \in N$. Damit ist gezeigt, dass N ein Untervektorraum von M ist. Für $x \in M$ ist

$$[x]_\sim = \{y \in M | \, \|y-x\| = 0\} = \{y \in M | \, y-x \in N\} = x + N$$

ein affiner Unterraum von M.

□

Die kanonische Abbildung $f : M \to M/\!\sim$, $f(x) = [x]_\sim$ erhält offensichtlich den Abstand. Man nennt sie daher auch eine *Isometrie* im Sinne der folgenden Definition:

Definition 1.10 (Isometrien) Es seien (M, d) und (N, d') zwei halbmetrische Räume. Eine Abbildung $f : M \to N$ heißt *isometrisch* (synonym: eine *Isometrie*) von (M, d) nach (N, d'), wenn für alle $x, y \in M$ gilt:

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

Übung 1.11 (Produktmetrik) *Zeigen Sie: Sind (M_i, d_i), $i = 1, \ldots, n$, halbmetrische Räume, so wird eine Halbmetrik auf dem kartesischen Produkt $M = M_1 \times \ldots \times M_n$ wie folgt definiert:*

$$d : M \times M \to \mathbb{R}, \quad d((x_1, \ldots, x_n), (y_1, \ldots, y_n)) := \max_{i=1,\ldots,n} d_i(x_i, y_i).$$

Sind alle d_i Metriken, so ist auch d eine Metrik. Sie wird Produkthalbmetrik bzw. Produktmetrik genannt.

Zum Beispiel ist die von $\|\cdot\|_\infty : \mathbb{R}^n \to \mathbb{R}$ erzeugte Metrik auf $\mathbb{R}^n = \mathbb{R} \times \ldots \times \mathbb{R}$ die Produktmetrik zur Standardmetrik auf $\mathbb{R}$.

1.2 Prähilberträume

Normierte Räume mit besonders schönen Eigenschaften erhält man aus Vektorräumen mit Skalarprodukt. Erinnern Sie sich dazu an den Begriff des Skalarprodukts aus der Linearen Algebra:

Definition 1.12 (reelle Prähilberträume) Es sei V ein $\mathbb{R}$-Vektorraum. Eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ heißt eine *Bilinearform*, wenn gilt:

1. *Linearität im ersten Argument:* Für alle $x, y, z \in V$ und $\lambda \in \mathbb{R}$ gilt:

$$\langle x+y, z \rangle = \langle x, z \rangle + \langle y, z \rangle\,,$$
$$\langle \lambda x, y \rangle = \lambda \langle x, y \rangle\,.$$

2. *Linearität im zweiten Argument:* Für alle $x, y, z \in V$ und $\lambda \in \mathbb{R}$ gilt:

$$\langle z, x+y \rangle = \langle z, x \rangle + \langle z, y \rangle\,,$$
$$\langle y, \lambda x \rangle = \lambda \langle y, x \rangle\,.$$

Eine solche Bilinearform $\langle \cdot, \cdot \rangle$ heißt *symmetrisch*, wenn für alle $x, y \in V$ gilt:

$$\langle x, y \rangle = \langle y, x \rangle .$$

Eine symmetrische Bilinearform heißt *positiv semidefinit*, wenn für alle $x \in V$ gilt: $\langle x, x \rangle \geq 0$. Gilt sogar $\langle x, x \rangle > 0$ für alle $x \in V \setminus \{0\}$, so heißt sie *positiv definit*. Eine positiv definite, symmetrische Bilinearform auf einem reellen Vektorraum wird (reelles) *Skalarprodukt* genannt. Das Paar $(V, \langle \cdot, \cdot \rangle)$ heißt dann ein (reeller) *Prähilbertraum*.

Für komplexe Vektorräume unterscheidet sich die Definition davon leicht durch das Auftreten der Komplexkonjugation, ähnlich wie man für $z \in \mathbb{R}$ schreiben kann: $|z|^2 = z \cdot z$, für $z \in \mathbb{C}$ jedoch nur $|z|^2 = \overline{z} \cdot z$.

Definition 1.13 (komplexe Prähilberträume) Es sei V ein $\mathbb{C}$-Vektorraum. Eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ heißt eine *Sesquilinearform* ("sesquilinear" = "$1\frac{1}{2}$-fach linear"), wenn gilt:[2]

1. *Antilinearität im ersten Argument:* Für alle $x, y, z \in V$ und $\lambda \in \mathbb{C}$ gilt:

$$\begin{aligned} \langle x + y, z \rangle &= \langle x, z \rangle + \langle y, z \rangle , \\ \langle \lambda x, y \rangle &= \overline{\lambda} \langle x, y \rangle . \end{aligned}$$

2. *Linearität im zweiten Argument:* Für alle $x, y, z \in V$ und $\lambda \in \mathbb{C}$ gilt:

$$\begin{aligned} \langle z, x + y \rangle &= \langle z, x \rangle + \langle z, y \rangle \\ \langle y, \lambda x \rangle &= \lambda \langle y, x \rangle . \end{aligned}$$

Eine solche Sesquilinearform $\langle \cdot, \cdot \rangle$ heißt *hermitesch*, wenn für alle $x, y \in V$ gilt:

$$\langle x, y \rangle = \overline{\langle y, x \rangle}.$$

Insbesondere gilt dann $\langle x, x \rangle \in \mathbb{R}$ für $x \in V$. Eine hermitesche Sesquilinearform heißt *positiv semidefinit*, wenn für alle $x \in V$ gilt: $\langle x, x \rangle \geq 0$. Gilt sogar $\langle x, x \rangle > 0$ für alle $x \in V \setminus \{0\}$, so heißt sie *positiv definit*. Eine positiv definite, hermitesche Sesquilinearform auf einem komplexen Vektorraum wird (komplexes) *Skalarprodukt* genannt. Das Paar $(V, \langle \cdot, \cdot \rangle)$ heißt dann ein (komplexer) *Prähilbertraum*.

In der Literatur werden auch manchmal die Rollen von 1. und 2. Argument des Skalarprodukts vertauscht, also Linearität des Skalarprodukts im 1. Argument, Antilinearität im 2. Argument gefordert. Wir verwenden diese Konvention hier nicht.

Im Folgenden sei wieder entweder $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Um nicht diese beiden Fälle getrennt behandeln zu müssen, wird im Folgenden die Komplexkonjugation meist dazugeschrieben, auch wenn sie im reellen Fall überflüssig ist.

---

[2] Vorsicht: Manche Autoren bevorzugen genau die umgekehrte Reihenfolge der Argumente: Linearität im *ersten* Argument und Antilinearität im *zweiten* Argument. Die hier gewählte Reihenfolge ist insbesondere in der (Mathematischen) Physik üblich und hat den Vorteil, dass die Linearform $\langle \ell, \cdot \rangle$ in Analogie zu Funktionsauswertungen $L(\cdot)$ ihr Argument *rechts* stehen hat.

Definition 1.14 (vom Skalarprodukt induzierte Norm) *Für eine positiv semidefinite symmetrische Bilinearform bzw. positiv semidefinite hermitesche Sesquilinearform* $\langle\cdot,\cdot\rangle$ *auf einem* $\mathbb{K}$*-Vektorraum* V *setzen wir für* $x \in V$*:*

$$\|x\| := \sqrt{\langle x, x\rangle}.$$

Lemma 1.15 *Die Abbildung* $\|\cdot\| : V \to \mathbb{R}$ *aus Definition 1.14 ist eine Seminorm. Ist* $\langle\cdot,\cdot\rangle$ *sogar positiv definit, also ein Skalarprodukt, so ist* $\|\cdot\|$ *eine Norm.*

Beweis: Offensichtlich ist $\|\cdot\|$ nichtnegativ und wegen

$$\|\lambda x\| = \sqrt{\langle \lambda x, \lambda x\rangle} = \sqrt{\overline{\lambda}\lambda \langle x, x\rangle} = |\lambda|\|x\| \qquad (x \in V, \lambda \in \mathbb{K})$$

auch homogen. Ist $\langle\cdot,\cdot\rangle$ sogar positiv definit, so folgt für $x \in V$ mit $x \neq 0$: $\langle x, x\rangle > 0$, also auch $\|x\| = \langle x, x\rangle^{1/2} > 0$.
Um zu zeigen, dass $\|\cdot\|$ die Dreiecksungleichung erfüllt, verwenden wir die folgende *Cauchy-Schwarz-Ungleichung*. Sie ist die wichtigste Ungleichung zur Abschätzung von Skalarprodukten und eine der wichtigsten Ungleichungen der gesamten Mathematik.

Satz 1.16 (Cauchy-Schwarz-Ungleichung) *Es sei* $\langle\cdot,\cdot\rangle$ *eine positiv semidefinite symmetrische Bilinearform bzw. eine positiv semidefinite hermitesche Sesquilinearform auf einem* $\mathbb{K}$*-Vektorraum* V*. Dann gilt für alle* $x, y \in V$*:*

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

Beweis: Gegeben $x, y \in V$, setzen wir

$$z := \|y\|^2 x - \langle y, x\rangle y$$

und erhalten:

$$\begin{aligned}
0 \leq \|z\|^2 &= \left\langle \|y\|^2 x - \langle y, x\rangle y, \|y\|^2 x - \langle y, x\rangle y \right\rangle \\
&= \left\langle \|y\|^2 x, \|y\|^2 x \right\rangle - \left\langle \|y\|^2 x, \langle y, x\rangle y \right\rangle - \left\langle \langle y, x\rangle y, \|y\|^2 x \right\rangle + \left\langle \langle y, x\rangle y, \langle y, x\rangle y \right\rangle \\
&= \|y\|^4 \langle x, x\rangle - \|y\|^2 \langle y, x\rangle \langle x, y\rangle - \overline{\langle y, x\rangle}\|y\|^2 \langle y, x\rangle + \overline{\langle y, x\rangle} \langle y, x\rangle \langle y, y\rangle \\
&= \|x\|^2\|y\|^4 - \|y\|^2 |\langle x, y\rangle|^2 - \|y\|^2 |\langle x, y\rangle|^2 + \|y\|^2 |\langle x, y\rangle|^2 \\
&= \|y\|^2(\|x\|^2\|y\|^2 - |\langle x, y\rangle|^2).
\end{aligned}$$

Im Fall $\|y\| \neq 0$, also $\|y\|^2 > 0$, folgt hieraus die Behauptung. Indem wir die Rollen von x und y vertauschen und $|\langle x, y\rangle| = |\langle y, x\rangle|$ beachten, folgt die Behauptung auch im Fall $\|x\| \neq 0$. Es bleibt der Fall $\|x\| = \|y\| = 0$ zu behandeln: Für $\lambda \in \mathbb{K}$ gilt in diesem Fall:

$$0 \leq \|x + \lambda y\|^2 = \|x\|^2 + \lambda \langle x, y\rangle + \overline{\lambda} \langle y, x\rangle + \overline{\lambda}\lambda\|y\|^2 = \lambda \langle x, y\rangle + \overline{\lambda \langle x, y\rangle} = 2\,\mathrm{Re}(\lambda \langle x, y\rangle).$$

Speziell für $\lambda = -\langle y, x\rangle = -\overline{\langle x, y\rangle}$ folgt:

$$0 \leq -\overline{\langle x, y\rangle} \langle x, y\rangle = -|\langle x, y\rangle|^2,$$

also

$$|\langle x, y\rangle| = 0 = \|x\|\|y\|.$$

$\square$

Geometrische Interpretation: Im Fall eines Skalarprodukts ist für $\|y\| \neq 0$

$$\|y\|^{-2}z = x - \frac{\langle y, x\rangle}{\|y\|^2}y$$

die orthogonale Projektion von x auf den Orthogonalraum von y.

Es bleibt die Dreiecksungleichung für $\|\cdot\|$ zu zeigen. Diese folgt nun für $x, y \in V$ aus der Cauchy-Schwarz-Ungleichung so:

$$\begin{aligned}\|x+y\|^2 &= \|x\|^2 + \langle x, y\rangle + \langle y, x\rangle + \|y\|^2 = \|x\|^2 + 2\operatorname{Re}\langle x, y\rangle + \|y\|^2 \leq \|x\|^2 + 2|\langle x, y\rangle| + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2\end{aligned}$$

und damit

$$\|x+y\| \leq \|x\| + \|y\|.$$

Damit ist Lemma 1.15 gezeigt.

Beispiel 1.17 1. Durch

$$\langle x, y\rangle := \sum_{j=1}^{n} \overline{x_j} y_j$$

für Vektoren $x = (x_1, \ldots, x_n)$ und $y = (y_1, \ldots, y_n)$ in $\mathbb{K}^n$, $n \in \mathbb{N}$, wird ein Skalarprodukt auf $\mathbb{K}^n$ gegeben. Es heißt *euklidisches Skalarprodukt* oder auch *Standardskalarprodukt* auf $\mathbb{K}^n$. Die zugehörige Norm wird mit

$$\|x\|_2 := \sqrt{\sum_{j=1}^{n} |x_j|^2}$$

bezeichnet und *euklidische Norm* auf $\mathbb{K}^n$ genannt. Sie verallgemeinert den elementargeometrischen Abstand vom Nullpunkt in $\mathbb{R}^2$ und in $\mathbb{R}^3$.

2. *Das folgende Beispiel ist eine abstraktere Variante des vorhergehenden Beispiels in einem Funktionenraum.*
Für gegebene $a, b \in \mathbb{R}$ mit $a < b$ bezeichne $V = \mathcal{R}([a, b], \mathbb{K})$ den $\mathbb{K}$-Vektorraum der Riemann-integrierbaren Funktionen $f : [a, b] \to \mathbb{K}$. Auf V wird durch

$$\langle f, g\rangle := \int_a^b \overline{f(x)} g(x)\, dx, \quad (f, g \in V)$$

eine positiv semidefinite symmetrische Bilinearform bzw. eine positiv semidefinite hermitesche Sesquilinearform definiert.[3] Die zugehörige Halbnorm wird ebenfalls mit

$$\|f\|_2 := \sqrt{\int_a^b |f(x)|^2\, dx}$$

[3] Man beachte, dass $\overline{f}g$ wieder Riemann-integrierbar ist (Übung). Um dies zu zeigen, verwende man, dass Riemann-Integrierbarkeit Beschränktheit impliziert.

bezeichnet. Sie ist jedoch keine Norm, wie das Beispiel $f(x) = 1_{\{x=a\}}$ zeigt: Hier ist zwar $f \neq 0$, aber $\|f\|_2 = 0$. Schränken wir das obige Skalarprodukt jedoch auf den Untervektorraum $C([a,b],\mathbb{K}) \subset V$ der *stetigen* Funktionen $f : [a,b] \to \mathbb{K}$ ein, so wird es positiv definit und damit $\|\cdot\|_2 : C([a,b],\mathbb{K}) \to \mathbb{R}$ zu einer Norm. In der Tat: Ist $f : [a,b] \to \mathbb{K}$ stetig, aber $f \neq 0$, so finden wir ein $y \in [a,b]$ mit $f(y) \neq 0$. Wegen der Stetigkeit von f und wegen $f(y) \neq 0$ finden wir $\epsilon > 0$ und $c, d \in \mathbb{R}$ mit $a \leq c < d \leq b$, so dass $|f(x)| \geq \epsilon$ für alle $x \in [c,d]$ gilt. Es folgt

$$\|f\|_2^2 = \int_a^b |f(x)|^2\,dx \geq \int_c^d |f(x)|^2\,dx \geq \int_c^d \epsilon^2\,dx = (d-c)\epsilon^2 > 0$$

und damit $\|f\|_2 > 0$.
Alternativ kann man natürlich aus $(\mathcal{R}[a,b], \langle\cdot,\cdot\rangle)$ einen Prähilbertraum machen, indem man Funktionen mit Abstand 0 wie in Lemma 1.9 miteinander verklebt.

1.3 Die Räume ℓ^p

Wir besprechen nun ein wichtiges, aber etwas komplizierteres Beispiel normierter Räume. Im Folgenden sei wieder entweder $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Weiter sei I eine endliche oder abzählbar unendliche Menge. $\mathbb{K}^I$ bezeichnet die Menge aller Abbildungen von I nach $\mathbb{K}$.

Definition 1.18 (p-Normen) *1. Für $1 \leq p < \infty$ definieren wir:*

$$\|\cdot\|_p : \mathbb{K}^I \to \mathbb{R} \cup \{+\infty\}, \quad \|f\|_p := \left(\sum_{i \in I} |f(i)|^p\right)^{1/p},$$

mit der Konvention $(+\infty)^{1/p} := +\infty$.

2. Weiter setzen wir

$$\|\cdot\|_\infty : \mathbb{K}^I \to \mathbb{R} \cup \{+\infty\}, \quad \|f\|_\infty := \sup_{i \in I} |f(i)|.$$

3. Für $1 \leq p \leq \infty$ setzen wir

$$\ell^p(I) := \{f \in \mathbb{K}^I \mid \|f\|_p < +\infty\}.$$

Wenn nicht aus dem Kontext klar ist, welcher Grundkörper $\mathbb{K}$ gemeint ist, schreiben wir auch genauer $\ell^p(I,\mathbb{K})$ dafür.

Die früher definierten Normen $\|\cdot\|_1$, $\|\cdot\|_2$ und $\|\cdot\|_\infty$ auf $\mathbb{K}^n$ sind Spezialfälle hiervon. Insbesondere gilt $\ell^p(\{1,\dots,n\}) = \mathbb{K}^n$.
Die Notation $\|\cdot\|_\infty$ wird durch folgende Übung motiviert:

Übung 1.19 *Zeigen Sie für $n \in \mathbb{N}$ und $f \in \mathbb{K}^n$:*

$$\|f\|_p \xrightarrow{p\to\infty} \|f\|_\infty.$$

Ein Ziel dieses Abschnitts ist es, zu zeigen, dass die Räume $\ell^p(I)$ Untervektorräume von $\mathbb{K}^I$ sind, und dass die Einschränkungen $\|\cdot\|_p : \ell^p(I) \to \mathbb{R}$ Normen sind.
Offensichtlich gilt $\|f\|_p \geq 0$ und $f \neq 0 \Rightarrow \|f\|_p \neq 0$ sowie $\|\lambda f\|_p = |\lambda| \|f\|_p$ für $f \in \mathbb{K}^I$ und $\lambda \in \mathbb{K}$. Der Nachweis der Dreiecksungleichung für $\|\cdot\|_p$ braucht als wesentliche Zutat die nun zu besprechende *Hölder-Ungleichung*.

Wir nennen zwei erweitert reelle Zahlen $p, q \in [1, \infty]$ *konjugiert zueinander*, wenn gilt: Entweder $1 < p < \infty$ und $q = \frac{p}{p-1}$, also

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

oder $p = 1$ und $q = \infty$ oder $p = \infty$ und $q = 1$.

Satz 1.20 (Hölder-Ungleichung für ℓ^p) *Es seien $p, q \in [1, \infty]$ konjugiert zueinander. Dann gilt für alle $f, g \in \mathbb{K}^I$:*

$$\sum_{j \in I} |f(j)g(j)| \leq \|f\|_p \|g\|_q. \tag{1}$$

Insbesondere ist für $f \in \ell^p(I)$ und $g \in \ell^q(I)$ die Reihe auf der linken Seite in (1) endlich. In diesem Fall folgt:

$$\left| \sum_{j \in I} f(j)g(j) \right| \leq \|f\|_p \|g\|_q < \infty. \tag{2}$$

Im Spezialfall $p = q = 2$ reduziert sich die Hölder-Ungleichung auf die Cauchy-Schwarz-Ungleichung für $\|\cdot\|_2$.
Beweis der Hölder-Ungleichung: Wir betrachten zuerst den Fall $1 < p < \infty$. Für $f = 0$ oder $g = 0$ ist die Behauptung offensichtlich. Wir dürfen also $f \neq 0$ und $g \neq 0$ und damit $\|f\|_p > 0$ und $\|g\|_q > 0$ annehmen. Im Fall $\|f\|_p = \infty$ oder $\|g\|_q = \infty$ wird die rechte Seite der Behauptung (1) unendlich und damit die Behauptung trivial. Daher dürfen wir $0 < \|f\|_p < \infty$ und $0 < \|g\|_q < \infty$ annehmen.
Wir zeigen nun für $a, b \in [0, \infty[$:

$$ab \leq \frac{1}{p} a^p + \frac{1}{q} b^q. \tag{3}$$

Dies ist klar für $a = 0$ oder $b = 0$. Deshalb nehmen wir nun $a > 0$ und $b > 0$ an. Weil die Exponentialfunktion auf $\mathbb{R}$ konvex ist, erhalten wir mit $1/p + 1/q = 1$ und $0 < 1/p < 1$:

$$ab = \exp\left(\frac{1}{p}\log(a^p) + \frac{1}{q}\log(b^q)\right) \leq \frac{1}{p}\exp(\log(a^p)) + \frac{1}{q}\exp(\log(b^q)) = \frac{1}{p}a^p + \frac{1}{q}b^q,$$

wie behauptet.
Setzen wir in der Ungleichung (3) Folgendes ein:

$$a = \frac{|f(j)|}{\|f\|_p} \quad \text{und} \quad b = \frac{|g(j)|}{\|g\|_q}$$

und summieren über $j \in I$, so folgt:

$$\begin{aligned}
\frac{\sum_{j\in I}|f(j)g(j)|}{\|f\|_p\|g\|_q} &= \sum_{j\in I}\frac{|f(j)|}{\|f\|_p}\frac{|g(j)|}{\|g\|_q} \\
&\le \sum_{j\in I}\left(\frac{1}{p}\frac{|f(j)|^p}{\|f\|_p^p}+\frac{1}{q}\frac{|g(j)|^q}{\|g\|_q^q}\right) \\
&= \frac{1}{p}\frac{1}{\|f\|_p^p}\sum_{j\in I}|f(j)|^p+\frac{1}{q}\frac{1}{\|g\|_q^q}\sum_{j\in I}|g(j)|^q \\
&= \frac{1}{p}+\frac{1}{q}=1.
\end{aligned}$$

Hieraus folgt die Hölder-Ungleichung (1).
Beweisen wir nun die Hölder-Ungleichung (1) im Fall $p=1$ und $q=\infty$. Wegen $|g(j)| \le \|g\|_\infty$ für $j \in I$ folgt

$$\sum_{j\in I}|f(j)g(j)| \le \sum_{j\in I}|f(j)|\|g\|_\infty = \|f\|_1\|g\|_\infty.$$

Ebenso, mit vertauschten Rollen von f und g, wird der Fall $p=\infty$, $q=1$ behandelt.
Für beliebige konjugierte p, q und $\|f\|_p < \infty$ sowie $\|g\|_q < \infty$ liefert die Hölder-Ungleichung (1), dass $\sum_{j\in I}|f(j)g(j)|$ endlich ist. In diesem Fall ist $\sum_{j\in I} f(j)g(j) \in \mathbb{K}$ wohldefiniert, und es folgt die Behauptung (2):

$$\left|\sum_{j\in I} f(j)g(j)\right| \le \sum_{j\in I}|f(j)g(j)| \le \|f\|_p\|g\|_q < \infty.$$

$\square$

Übung 1.21 (Integralversion der Hölder-Ungleichung) *Für $1 \le p < \infty$ und $f \in C([a,b],\mathbb{K})$ mit $a, b \in \mathbb{R}$, $a < b$ definieren wir*

$$\|f\|_p := \left(\int_a^b |f(x)|^p\,dx\right)^{1/p}.$$

Erinnern Sie sich auch an die Norm $\|\cdot\|_\infty : C([a,b],\mathbb{K}) \to \mathbb{R}$ aus Beispiel 1.6 4. Zeigen Sie für konjugierte $p, q \in [1,\infty]$ und für $f, g \in C([a,b],\mathbb{K})$:

$$\left|\int_a^b f(x)g(x)\,dx\right| \le \|f\|_p\|g\|_q.$$

Imitieren Sie dazu den Beweis der Hölder-Ungleichung in ℓ^p.

Korollar 1.22 (Dualität zwischen p-Norm und q-Norm) *Für konjugierte $p, q \in [1, \infty]$ und für $f \in \mathbb{K}^I$ gilt*

$$\|f\|_p = \sup \left\{ \sum_{j \in I} |f(j)g(j)| \,\middle|\, g \in \ell^q(I), \|g\|_q \leq 1 \right\}. \tag{4}$$

Für $f \in \ell^p(I)$ gilt:

$$\|f\|_p = \sup \left\{ \left| \sum_{j \in I} f(j)g(j) \right| \,\middle|\, g \in \ell^q(I), \|g\|_q \leq 1 \right\}. \tag{5}$$

Beweis: Wir zeigen zunächst den Teil "$\geq$" der Behauptung. Hierzu seien $f \in \mathbb{K}^I$ und $g \in \ell^q(I)$ mit $\|g\|_q \leq 1$ gegeben. Die Hölder-Ungleichung liefert

$$\|f\|_p \geq \|f\|_p \|g\|_q \geq \sum_{j \in I} |f(j)g(j)|. \tag{6}$$

Nehmen wir hier von der rechten Seite das Supremum über g, folgt

$$\|f\|_p \geq \sup \left\{ \sum_{j \in I} |f(j)g(j)| \,\middle|\, g \in \ell^q(I), \|g\|_q \leq 1 \right\}.$$

Für gegebenes g wie oben und $f \in \ell^p(I)$, also $\|f\|_p < \infty$, erhalten wir aus (6), dass $\sum_{j \in I} f(j)g(j) \in \mathbb{K}$ existiert, sowie

$$\infty > \|f\|_p \geq \sum_{j \in I} |f(j)g(j)| \geq \left| \sum_{j \in I} f(j)g(j) \right|.$$

Nehmen wir auch hier wieder von der rechten Seite das Supremum über g, folgt

$$\|f\|_p \geq \sup \left\{ \left| \sum_{j \in I} f(j)g(j) \right| \,\middle|\, g \in \ell^q(I), \|g\|_q \leq 1 \right\}.$$

Nun zeigen wir den Teil "$\leq$" der Behauptung.
1. Fall: $1 < p < \infty$. Es sei $f \in \mathbb{K}^I$ gegeben. Im Fall $f = 0$ ist die Behauptung trivial; deshalb dürfen wir $f \neq 0$, also $\|f\|_p > 0$ annehmen. Wir betrachten

$$h \in \mathbb{K}^I, \quad h(j) := \begin{cases} |f(j)|^{p-2}\overline{f(j)} & \text{für } f(j) \neq 0 \\ 0 & \text{für } f(j) = 0 \end{cases} \tag{7}$$

Aufgrund der Annahme $f \neq 0$ gilt $h \neq 0$, also $\|h\|_q > 0$. Weiter gilt $|h| = |f|^{p-1}$ und daher

$$|h|^q = |f|^{(p-1)q} = |f|^p$$

sowie wegen $f\overline{f} = |f|^2$:

$$fh = |f|^p \geq 0$$

Insbesondere erhalten wir

$$\|f\|_p^p = \|h\|_q^q = \sum_{j\in I} f(j)h(j).$$

Wir unterscheiden nochmal zwei Fälle:
1.1. Fall: $f \in \ell^p(I)$. Hier folgt $h \in \ell^q(I)$, so dass $g_* := h/\|h\|_q \in \ell^q(I)$ wohldefiniert ist und $\|g_*\|_q = 1$ erfüllt. Wir erhalten in diesem Fall: $f(j)g_*(j) \geq 0$ für alle $j \in I$, und damit

$$\begin{aligned}\sum_{j\in I} f(j)g_*(j) &= \sum_{j\in I} |f(j)g_*(j)| = \left|\sum_{j\in I} f(j)g_*(j)\right| \\ &= \|h\|_q^{-1} \sum_{j\in I} f(j)h(j) = \frac{\|f\|_p^p}{\|h\|_q} = \frac{\|f\|_p^p}{\|f\|_p^{p/q}} = \|f\|_p^{p(1-1/q)} = \|f\|_p.\end{aligned}$$

Die Suprema in (4) und (5) sind also im betrachteten Fall sogar Maxima und werden an der Stelle $g = g_*$ angenommen.
1.2. Fall: $f \in \mathbb{K}^I \setminus \ell^p(I)$, also $\|f\|_p = \infty$ Hier sei

$$I_1 \subseteq I_2 \subseteq I_3 \subseteq \dots$$

eine aufsteigende Folge von endlichen Teilmengen von I mit $\bigcup_{n\in\mathbb{N}} I_n = I$. Es sei $1_{I_n} \in \mathbb{K}^I$ die Indikatorfunktion von I_n (Wert 1 auf I_n, Wert 0 auf $I \setminus I_n$). Nach dem Satz von der monotonen Konvergenz für Reihen gilt

$$\|f1_{I_n}\|_p^p = \sum_{j\in I} |f(j)|^p 1_{I_n}(j) \stackrel{n\to\infty}{\longrightarrow} \sum_{j\in I} |f(j)|^p = \|f\|_p^p = \infty. \tag{8}$$

Nun sei $n \in \mathbb{N}$ gegeben. Wegen $f1_{I_n} \in \ell^p(I)$ wissen wir (4) schon für $f1_{I_n}$ statt f:

$$\begin{aligned}\|f1_{I_n}\|_p &= \sup\left\{\sum_{j\in I} |f(j)1_{I_n}(j)g(j)| \,\middle|\, g \in \ell^q(I), \|g\|_q \leq 1\right\} \\ &\leq \sup\left\{\sum_{j\in I} |f(j)g(j)| \,\middle|\, g \in \ell^q(I), \|g\|_q \leq 1\right\}\end{aligned}$$

Im Limes $n \to \infty$ folgt hieraus der Teil "$\leq$" der Behauptung (4) für f wegen (8).
2. Fall: $p = 1$, also $q = \infty$. Wir setzen

$$g_* : I \to \mathbb{K}, \quad g_*(j) = \begin{cases} \frac{\overline{f(j)}}{|f(j)|} & \text{für } f(j) \neq 0 \\ 1 & \text{für } f(j) = 0 \end{cases}$$

so dass $|g_*| = 1$ und $fg = |f|$ gilt. Insbesondere folgt $\|g_*\|_\infty = 1$ und

$$\|f\|_1 = \sum_{j\in I} |f(j)| = \sum_{j\in I} f(j)g_*(j) = \sum_{j\in I} |f(j)g_*(j)|.$$

Hieraus folgt "$\leq$" sowohl in (4) als auch in (5).
3. Fall: $p = \infty$, also $q = 1$. Wir setzen für $k \in I$:

$$g_k : I \to \mathbb{K}, \quad g_k(j) = \begin{cases} \frac{\overline{f(k)}}{|f(k)|} & \text{für } j = k \text{ und } f(j) \neq 0 \\ 1 & \text{für } j = k \text{ und } f(j) = 0 \\ 0 & \text{für } j \neq k \end{cases}$$

Dann gilt $\|g_k\|_1 = |g_k(k)| = 1$ und

$$\sum_{j\in I} |f(j)g_k(j)| = \sum_{j\in I} f(j)g_k(j) = f(k)g_k(k) = |f(k)|,$$

also

$$\sup\left\{ \sum_{j\in I} |f(j)g(j)| \,\middle|\, g \in \ell^1(I), \|g\|_1 \leq 1 \right\} \geq \sup_{k\in I} \sum_{j\in I} |f(j)g_k(j)| = \sup_{k\in I} |f(k)| = \|f\|_\infty$$

und im Fall $\|f\|_\infty < \infty$ ebenso

$$\sup\left\{ \left|\sum_{j\in I} f(j)g(j)\right| \,\middle|\, g \in \ell^1(I), \|g\|_1 \leq 1 \right\} \geq \sup_{k\in I} \left|\sum_{j\in I} f(j)g_k(j)\right| = \sup_{k\in I} |f(k)| = \|f\|_\infty.$$

$\square$

Übung 1.23 (Gleichheit in der Hölder-Ungleichung in ℓ^p) *Es seien I eine abzählbare Menge, $p, q \in]1, \infty[$ konjugiert zueinander, $f \in \ell^p(I)$, $f \neq 0$, und $g \in \ell^q(I)$. Zeigen Sie, dass folgende beiden Aussagen äquivalent sind:*

1. $\sum_{j\in I} f(j)g(j) = \|f\|_p \|g\|_q$
2. *Es gilt $g = \alpha h$ für ein $\alpha \geq 0$ mit $h \in \mathbb{K}^I$ wie in Formel (7) im Beweis von Korollar 1.22.*

Übung 1.24 (Dualität zwischen p- und q-Norm für stetige Funktionen) *Mit den Notationen der Übung 1.21 seien $p, q \in [1, \infty]$ konjugiert zueinander und $f \in C([a, b], \mathbb{K})$. Zeigen Sie:*

$$\|f\|_p = \sup\left\{ \left|\int_a^b f(x)g(x)\,dx\right| \,\middle|\, g \in C([a, b], \mathbb{K}), \|g\|_q \leq 1 \right\}. \tag{9}$$

Lassen Sie sich dazu vom Beweis des Korollars 1.22 inspirieren.

Korollar 1.25 (Dreiecksungleichung in ℓ^p, synonym: Minkowskische Ungleichung) *Für $1 \leq p \leq \infty$ und $f, g \in \mathbb{K}^I$ gilt*

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p.$$

Insbesondere ist der Raum $\ell^p(I)$ ein Untervektorraum von $\mathbb{K}^I$ und $\|\cdot\|_p : \ell^p(I) \to \mathbb{R}$ eine Norm.

Beweis: Es sei $q \in [1, \infty]$ konjugiert zu p. Wir kürzen ab: $\mathcal{B} := \{h \in \ell^q(I) | \, \|h\|_q \leq 1\}$. Mit Korollar 1.22 folgt

$$\begin{aligned}
\|f+g\|_p &= \sup\left\{ \sum_{j\in I} |(f(j)+g(j))h(j)| \,\middle|\, h \in \mathcal{B} \right\} \\
&\leq \sup\left\{ \sum_{j\in I} |f(j)h(j)| + \sum_{j\in I} |g(j))h(j)| \,\middle|\, h \in \mathcal{B} \right\} \\
&\leq \sup\left\{ \sum_{j\in I} |f(j)h(j)| \,\middle|\, h \in \mathcal{B} \right\} + \sup\left\{ \sum_{j\in I} |g(j))h(j)| \,\middle|\, h \in \mathcal{B} \right\} \\
&= \|f\|_p + \|g\|_p.
\end{aligned}$$

Insbesondere ist $\ell^p(I)$ abgeschlossen unter Addition. Wegen $0 \in \ell^p(I)$, $\|\lambda f\|_p = |\lambda| \|f\|_p$ für $\lambda \in \mathbb{K}$, $f \in \ell^p(I)$ folgt: $\ell^p(I) \subseteq \mathbb{K}^I$ ist ein Untervektorraum. Zusammen mit $\|f\|_p = 0 \Rightarrow f = 0$ für $f \in \ell^p(I)$ folgt: $\|\cdot\|_p : \ell^p(I) \to \mathbb{R}$ ist eine Norm.

□

Übung 1.26 (Abstraktion des Beweises der Minkowskiungleichung) *Es seien V ein $\mathbb{K}$-Vektorraum und $\|\cdot\| : V \to \mathbb{R}$ eine Abbildung. $V' = \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$ bezeichne den Dualraum von V, also die Menge aller linearen Abbildungen von V nach $\mathbb{K}$ (synonym: Linearformen auf V). Weiter sei $\mathcal{B} \subseteq V'$ eine nichtleere Menge von Linearformen, und es gelte für alle $x \in V$:*

$$\|x\| = \sup\{|H(x)| \mid H \in \mathcal{B}\}.$$

Beweisen Sie, dass $\|\cdot\|$ eine Halbnorm auf V ist.

Übung 1.27 *Es seien $1 \leq p < \infty$ und $-\infty < a < b < \infty$. Beweisen Sie die Dreiecksungleichung für $\|\cdot\|_p : C([a,b], \mathbb{K}) \to \mathbb{R}$.*

1.4 Topologie metrischer Räume

In der Analysis 1 spielten ϵ-Umgebungen und offene Mengen eine fundamentale Rolle bei der Einführung von Konvergenz und Stetigkeit. Diese Begriffe lassen sich analog auf halbmetrische Räume übertragen:

Definition 1.28 (ϵ-Umgebungen und offene Mengen) Es sei (M, d) ein halbmetrischer Raum. Für $x \in M$ und $\epsilon > 0$ nennen wir

$$U_\epsilon^d(x) := \{y \in M |\ d(x, y) < \epsilon\}$$

die ϵ-Umgebung von x in (M, d). [4] Wir nennen eine Teilmenge $U \subseteq M$ *offen* (bezüglich d), wenn gilt:

$$\forall x \in U\ \exists \epsilon > 0 : U_\epsilon^d(x) \subseteq U.$$

Zum Beispiel sind ϵ-Umgebungen $U_\epsilon^d(x)$ offen. Um dies zu sehen, sei $y \in U_\epsilon^d(x)$ gegeben. Dann ist $d(x, y) < \epsilon$, also $\delta := \epsilon - d(x, y) > 0$. Es folgt $U_\delta^d(y) \subseteq U_\epsilon^d(x)$, denn für beliebig gegebene $z \in U_\delta^d(y)$ folgt: $d(y, z) < \delta$, also

$$d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + \delta = \epsilon$$

und damit $z \in U_\epsilon^d(x)$.
Erinnern Sie sich an den Begriff der *Topologie*:

Definition 1.29 (Topologie) Ein Mengensystem $\mathcal{T} \subseteq \mathcal{P}(M)$ über einer Menge M heißt eine *Topologie* (über M), wenn gilt:

1. $\emptyset \in \mathcal{T}$ und $M \in \mathcal{T}$.
2. *Abgeschlossenheit unter endlicher Durchschnittsbildung:*
$$\forall U, V \in \mathcal{T} :\ U \cap V \in \mathcal{T}.$$
3. *Abgeschlossenheit unter beliebiger Vereinigungsbildung:* Für jede Familie $(U_i)_{i \in I}$ von Mengen $U_i \in \mathcal{T}$ gilt:
$$\bigcup_{i \in I} U_i \in \mathcal{T}.$$

Das Paar $(M, \mathcal{T})$ nennen wir dann einen *topologischen Raum* . Die Elemente $U \in \mathcal{T}$ einer Topologie $\mathcal{T}$ werden *offen* (in $(M, \mathcal{T})$) genannt.

Lemma 1.30 (von einer Halbmetrik erzeugte Topologie) *Ist (M, d) ein halbmetrischer Raum, so bildet das Mengensystem*

$$\mathcal{T}_d := \{U \subseteq M |\ U \text{ ist offen bzgl. } d\}$$

aller in (M, d) offenen Mengen eine Topologie über M. Sie wird die von d erzeugte Topologie genannt.

Beweis:

[4] Wenn klar ist, welche Halbmetrik d gemeint ist, schreiben wir auch $U_\epsilon(x)$ statt $U_\epsilon^d(x)$.

1. $\emptyset \in \mathcal{T}_d$ und $M \in \mathcal{T}_d$ folgen unmittelbar aus der Definition.

2. Es seien $U, V \in \mathcal{T}_d$ gegeben. Um $U \cap V \in \mathcal{T}_d$ zu beweisen, müssen wir zeigen:

$$\forall x \in U \cap V \ \exists \epsilon > 0 : \ U_\epsilon(x) \subseteq U \cap V.$$

Hierzu sei $x \in U \cap V$ gegeben. Wegen $x \in U$ finden wir ein $\delta > 0$ mit $U_\delta(x) \subseteq U$. Ebenso finden wir wegen $x \in V$ ein $\eta > 0$ mit $U_\eta(x) \subseteq V$. Wir wählen $\epsilon := \min\{\delta, \eta\} > 0$. Es folgt

$$U_\epsilon(x) = U_\delta(x) \cap U_\eta(x) \subseteq U \cap V.$$

3. Es sei eine Familie $(U_i)_{i \in I}$ von offenen Mengen $U_i \in \mathcal{T}_d$ gegeben. Wir müssen zeigen:

$$\forall x \in \bigcup_{i \in I} U_i \ \exists \epsilon > 0 : \ U_\epsilon(x) \subseteq \bigcup_{i \in I} U_i.$$

Hierzu sei $x \in \bigcup_{i \in I} U_i$ gegeben. Wir finden also ein $j \in I$ mit $x \in U_j$. Weil U_j offen in (M, d) ist, finden wir ein $\epsilon > 0$ mit $U_\epsilon(x) \subseteq U_j$. Mit $U_j \subseteq \bigcup_{i \in I} U_i$ folgt hieraus die Behauptung $U_\epsilon(x) \subseteq \bigcup_{i \in I} U_i$.

□

Wird eine Metrik d auf einem $\mathbb{K}$-Vektorraum V von einer Halbnorm $\|\cdot\|$ erzeugt, so nennen wir $\mathcal{T}_{\|\cdot\|} := \mathcal{T}_d$ auch die von der Halbnorm erzeugte Topologie.

Beispiel 1.31 *Der Absolutbetrag $|\cdot| : \mathbb{K} \to \mathbb{R}$ erzeugt die aus der Analysis 1 bekannte Topologie auf $\mathbb{K}$. Wir nennen sie die Standardtopologie $\mathcal{T}_\mathbb{K}$ auf $\mathbb{K}$. Die in Beispiel 1.2 2. betrachtete Metrik auf $\mathbb{R} \cup \{\pm\infty\}$ erzeugt die in der Analysis 1 eingeführte Topologie auf $\mathbb{R} \cup \{\pm\infty\}$.*

Übung 1.32 *Geben Sie eine Metrik auf $\mathbb{C} \cup \{\infty\}$ an, die die aus der Analysis 1 bekannte Topologie auf $\mathbb{C} \cup \{\infty\}$ erzeugt.*

Übung 1.33 *Es sei (M, d) ein halbmetrischer Raum und $a \in \mathbb{R}^+$. Zeigen Sie, dass durch*

$$d_1 : M \times M \to \mathbb{R}, \quad d_1(x, y) = \frac{d(x, y)}{a + d(x, y)}$$

und durch

$$d_2 : M \times M \to \mathbb{R}, \quad d_2(x, y) = \min\{d(x, y), a\}$$

Halbmetriken auf M gegeben werden, die die gleiche Topologie wie d erzeugen.

In der Analysis 1 haben Sie schon zahlreiche topologische Begriffe am Beispiel der Räume $\mathbb{R}$, $\mathbb{R}\cup\{\pm\infty\}$, $\mathbb{C}$ und $\mathbb{C}\cup\{\infty\}$ kennengelernt. Wir wiederholen hier einige wichtige dieser Begriffe in allgemeinerem Rahmen:
Es sei ein topologischer Raum $(M, \mathcal{T})$ gegeben.

Definition 1.34 (abgeschlossene Mengen) *Eine Menge $A \subseteq M$ heißt abgeschlossen (in $(M, \mathcal{T})$), wenn ihr Komplement $M \setminus A$ offen (in $(M, \mathcal{T})$) ist.*

In Analogie zu Definition 1.29 erhalten wir:

Lemma 1.35 (Fundamentale Eigenschaften der Abgeschlosenheit)

1. *$\emptyset$ und M sind abgeschlossen.*
2. *Abgeschlossenheit unter endlicher Vereinigungsbildung: Sind $A, B \subseteq M$ abgeschlossen, so ist auch $A \cup B$ abgeschlossen.*
3. *Abgeschlossenheit unter beliebiger Durchschnittsbildung: Für jede Familie $(A_i)_{i\in I}$ abgeschlossener Mengen $A_i \subseteq M$ ist $\bigcap_{i\in I} A_i$ bezüglich $\mathcal{T}$ abgeschlossen.*

Beweis: Dies folgt unmittelbar aus den Definitionen 1.29 und 1.34 durch Komplementbildung unter Verwendung der de-Morganschen Regeln

$$A \cup B = M \setminus ((M \setminus A) \cap (M \setminus B)),$$
$$\bigcap_{i\in I} A_i = M \setminus \bigcup_{i\in I} (M \setminus A_i).$$

□

Definition 1.36 (Grundlegende topologische Begriffe) Es sei $(M, \mathcal{T})$ ein topologischer Raum und $N \subseteq M$.

1. Die Vereinigung aller offenen Mengen, die Teilmengen von N sind, heißt *Inneres* von N und wird mit N° bezeichnet. Die Elemente von N° werden *innere Punkte* von N genannt. Die Menge N heißt eine *Umgebung* von einem Punkt $x \in M$, wenn x ein innerer Punkt von N ist.
2. Der Durchschnitt aller abgeschlossenen Mengen $A \subseteq M$ mit $N \subseteq A$ heißt *Abschluss* von N und wird mit $\overline{N}$ bezeichnet. Die Elemente von $\overline{N}$ werden *Berührpunkte* von N genannt. Gilt $\overline{N} = M$, so nennen wir N *dicht* in M.
3. Der *Rand* ∂N von N wird durch

   $$\partial N := \overline{N} \cap \overline{M \setminus N}$$

   definiert. Die Elemente von ∂N werden *Randpunkte* von N genannt.

Bemerkung 1.37 1. Aus der Definition folgt unmittelbar: $N^\circ \subseteq N \subseteq \overline{N}$. Weiter sieht man leicht $\partial N = \overline{N} \setminus N^\circ$ und $M \setminus N^\circ = \overline{M \setminus N}$. (Übung)

2. Mit Definition 1.29 und Lemma 1.35 folgt: Das Innere N° von N ist offen, und der Abschluss $\overline{N}$ ist abgeschlossen.

3. Wird die Topologie $\mathcal{T} = \mathcal{T}_d$ von einer Halbmetrik d erzeugt, so erhält man folgende "metrische Charakterisierung" der topologischen Grundbegriffe: Für $x \in M$ und $N \subseteq M$ gilt:

$$x \in N^\circ \Leftrightarrow \exists \epsilon > 0 : \; U_\epsilon^d(x) \subseteq N,$$
$$x \in \overline{N} \Leftrightarrow \forall \epsilon > 0 : \; U_\epsilon^d(x) \cap N \neq \emptyset.$$

Übung 1.38 Zeigen Sie, dass $x \in M$ genau dann ein Berührpunkt von $N \subseteq M$ ist, wenn jede Umgebung von x die Menge N trifft: $U \cap N \neq \emptyset$. Zeigen Sie auch, dass $N \subseteq M$ genau dann dicht in M ist, wenn für jede nichtleere offene Menge $U \subseteq M$ gilt: $U \cap N \neq \emptyset$.

Übung 1.39 Es sei $(M, \mathcal{T})$ ein topologischer Raum und $N \subseteq M$. Zeigen Sie:

1. N ist genau dann abgeschlossen, wenn $N = \overline{N}$ gilt. Folgern Sie: $\overline{\overline{N}} = \overline{N}$.

2. N ist genau dann offen, wenn $N = N^\circ$ gilt. Folgern Sie: $N^{\circ\circ} = N^\circ$.

Übung 1.40 Es sei (M, d) ein halbmetrischer Raum, $x \in M$ und $r \in \mathbb{R}^+$. Zeigen Sie, dass $B_r^d(x) := \{y \in M |\; d(x, y) \leq r\}$ eine abgeschlossene Menge mit $B_r^d(x) \supseteq \overline{U_r^d(x)}$ ist. Zeigen Sie an einem Gegenbeispiel, dass $B_r^d(x) \neq \overline{U_r^d(x)}$ gelten kann. Zeigen Sie jedoch, dass in einem *halbnormierten* Raum $B_r^d(x) = \overline{U_r^d(x)}$ gilt.

Übung 1.41 *Es sei p eine Primzahl und d_p die p-adische Metrik auf $\mathbb{Z}$ aus Beispiel 1.2.3. Beweisen Sie, dass für $k \in \mathbb{Z}$ die Menge $k + p\mathbb{Z} = \{k + pz | z \in \mathbb{Z}\}$ offen und abgeschlossen bezüglich d_p ist.*

Übung 1.42 Es sei I eine abzählbar unendliche Menge und $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Zeigen Sie, dass der Raum

$$\mathbb{C}^{(I)} := \{(\alpha_j)_{j \in I} \in \mathbb{K}^I |\; \alpha_j = 0 \text{ für alle bis auf endlich viele } j \in I\}$$

dicht in $(\ell^p(I, \mathbb{K}), \|\cdot\|_p)$ für $1 \leq p < \infty$, aber nicht dicht in $(\ell^\infty(I, \mathbb{K}), \|\cdot\|_\infty)$ ist.

Übung 1.43 (Approximation Riemann-integrierbarer Funktionen durch stetige Funktionen) Es sei $(\mathcal{R}([a, b], \mathbb{K}), \langle\cdot, \cdot\rangle)$ der Prähilbertraum der Riemann-integrierbaren Funktionen aus Beispiel 1.17. Zeigen Sie, dass $C([a, b], \mathbb{K})$ dicht in $(\mathcal{R}([a, b], \mathbb{K}), \|\cdot\|_2)$ ist. Zeigen Sie dazu zuerst, dass sich jede Indikatorfunktion 1_I eines Intervalls $I \subseteq [a, b]$ in der 2-Norm beliebig genau durch stetige Funktionen approximieren lässt. Betrachten Sie dazu die Funktionen

$$f_n(x) := \max\{1 - n \operatorname{dist}(x, I), 0\} \text{ für } n \in \mathbb{N}, x \in [a, b],$$

wobei $\mathrm{dist}(x, I) = \inf\{|x - y| \; |y \in I\}$ den Abstand von x zu I bezeichnet. Folgern Sie daraus, dass jede Treppenfunktion im Abschluss von $C([a, b], \mathbb{K})$ in $(\mathcal{R}([a, b], \mathbb{K}), \|\cdot\|_2)$ liegt. Zeigen Sie dann, dass der Raum der Treppenfunktionen dicht in $(\mathcal{R}([a, b], \mathbb{K}), \|\cdot\|_2)$ liegt.

1.5 Konvergente Folgen und Stetigkeit

Erinnern Sie sich aus der Analysis 1 an den Konvergenzbegriff bei Folgen:

Definition 1.44 (Konvergenz von Folgen) Eine Folge $(a_n)_{n\in\mathbb{N}}$ in einem topologischen Raum $(M, \mathcal{T})$ heißt *konvergent* in $(M, \mathcal{T})$ gegen einen *Grenzwert* $x \in M$, in Zeichen $a_n \overset{n\to\infty}{\longrightarrow}_{\mathcal{T}} x$ oder auch kurz $a_n \overset{n\to\infty}{\longrightarrow} x$, wenn es zu jeder Umgebung U von x in $(M, \mathcal{T})$ ein $m \in \mathbb{N}$ gibt, so dass für alle $n \geq m$ gilt: $a_n \in U$. Ist der Grenzwert x einer konvergenten Folge $(a_n)_{n\in\mathbb{N}}$ eindeutig bestimmt, schreiben wir auch $x = \lim^{\mathcal{T}}_{n\to\infty} a_n$ oder kurz $x = \lim_{n\to\infty} a_n$ dafür.

Konvergenz in einem halbmetrischen Raum (M, d) oder halbnormierten Raum $(V, \|\cdot\|)$ wird als Konvergenz bezüglich der zugehörigen Topologie $\mathcal{T}_d$ bzw. $\mathcal{T}_{\|\cdot\|}$ verstanden.

Lemma 1.45 (Metrische Charakterisierung der Konvergenz) *Es seien (M, d) ein halbmetrischer Raum, $(a_n)_{n\in\mathbb{N}}$ eine Folge in M und $x \in M$. Dann sind äquivalent:*

1. $a_n \overset{n\to\infty}{\longrightarrow}_{\mathcal{T}_d} x$
2. $\lim_{n\to\infty} d(x, a_n) = 0$, *d.h.* $\forall \epsilon > 0 \; \exists m \in \mathbb{N} \; \forall n \geq m : d(x, a_n) < \epsilon$.

Beweis: Die Richtung "1.⇒ 2." ist trivial, da jede ϵ-Umgebung $U^d_\epsilon(x)$ von x auch eine Umgebung von x bezüglich $\mathcal{T}_d$ ist. Zu "2.⇒ 1.": Es gelte 2., und es sei U eine Umgebung von x bezüglich $\mathcal{T}_d$. Dann gibt es ein $\epsilon > 0$ mit $U^d_\epsilon(x) \subseteq U$. Wegen 2. finden wir ein $m \in \mathbb{N}$, für das gilt:

$$\forall n \geq m : \; a_n \in U^d_\epsilon(x) \subseteq U.$$

Es folgt $a_n \overset{n\to\infty}{\longrightarrow}_{\mathcal{T}_d} x$.

□

In allgemeinen topologischen Räumen oder auch in halbmetrischen Räumen, die keine metrischen Räume sind, braucht der Grenzwert x nicht eindeutig zu sein. Als ein Beispiel diene die triviale Topologie $T = \{\emptyset, M\}$ auf M: Bezüglich dieser trivialen Topologie konvergiert jede Folge in M gegen jeden Punkt in M. Die folgende Eigenschaft topologischer Räume garantiert jedoch die Eindeutigkeit von Grenzwerten:

Definition 1.46 (Hausdorffeigenschaft) Ein topologischer Raum $(M, \mathcal{T})$ heißt *Hausdorffraum*, wenn es zu je zwei verschiedenen Punkten $x, y \in M$, $x \neq y$, eine Umgebung U von x und eine Umgebung V von y bezüglich $\mathcal{T}$ mit $U \cap V = \emptyset$ gibt.

Lemma 1.47 (Eindeutigkeit des Limes in Hausdorffräumen) *Der Grenzwert konvergenter Folgen $(a_n)_{n\in\mathbb{N}}$ in einem Hausdorffraum ist eindeutig bestimmt.*

Beweis: Gilt sowohl $a_n \overset{n\to\infty}{\longrightarrow} x$ als auch $a_n \overset{n\to\infty}{\longrightarrow} y$ und sind U und V Umgebungen von x bzw. von y, so enthält sowohl U als auch V alle a_n für genügend große $n \in \mathbb{N}$. Insbesondere folgt $U \cap V \neq \emptyset$. Da dies für beliebige Umgebungen U von x und V von y gilt, folgt $x = y$ wegen der Haussdorffeigenschaft.

□

Lemma 1.48 (Hausdorffeigenschaft bei metrischen Räumen) *Es sei (M, d) ein halbmetrischer Raum. Dann sind äquivalent:*

1. *(M, d) ist ein metrischer Raum.*
2. *$(M, \mathcal{T}_d)$ ist ein Hausdorffraum.*

Beweis: "1. ⇒ 2.": Ist d eine Metrik und sind x, y zwei verschiedene Punkte in M, so ist $\epsilon := d(x, y)/2 > 0$. Dann sind $U_\epsilon(x)$ und $U_\epsilon(y)$ disjunkte Umgebungen von x bzw. y. In der Tat: Wäre $z \in U_\epsilon(x) \cap U_\epsilon(y)$, so folgte der Widerspruch

$$2\epsilon = d(x, y) \leq d(x, z) + d(y, z) < \epsilon + \epsilon = 2\epsilon.$$

"2. ⇒ 1.": Ist $(M, \mathcal{T}_d)$ ein Hausdorffraum und sind $x, y \in M$ mit $d(x, y) = 0$, so folgt $y \in U_\epsilon(x)$ für alle $\epsilon > 0$. Da jede Umgebung von x eine geeignete ϵ-Umgebung von x enthält, ist dann y in jeder Umgebung von x enthalten. Trivialerweise ist y aber auch in jeder Umgebung von y enthalten. Die Punkte x und y besitzen also keine disjunkten Umgebungen. Aus der Hausdorffeigenschaft folgt $x = y$. Also ist d eine Metrik.

□

Lemma 1.49 (Berührpunkte in halbmetrischen Räumen) *Es seien (M, d) ein halbmetrischer Raum, $x \in M$ und $N \subseteq M$. Dann sind äquivalent:*

1. *x ist ein Berührpunkt von N in $(M, \mathcal{T}_d)$.*
2. *Es gibt eine gegen x bzgl. $\mathcal{T}_d$ konvergente Folge $(a_n)_{n\in\mathbb{N}}$ in N.*

Beweis: "1. ⇒ 2.": Ist x ein Berührpunkt von N, so ist für alle $n \in \mathbb{N}$ die Menge $U_{1/n}(x) \cap N$ nichtleer. Wir können dann (mit dem Auswahlaxiom) eine Folge $(a_n)_{n\in\mathbb{N}}$ in N auswählen, so dass für alle $n \in \mathbb{N}$ gilt: $a_n \in U_{1/n}(x) \cap N$. Gegeben $\epsilon > 0$, finden wir $m \in \mathbb{N}$ mit $1/m < \epsilon$. Für alle $n \geq m$ erhalten wir:

$$d(a_n, x) < \frac{1}{n} \leq \frac{1}{m} < \epsilon.$$

Damit ist $a_n \overset{n\to\infty}{\longrightarrow} x$ gezeigt.
"2. ⇒ 1.": Konvergiert $(a_n)_{n\in\mathbb{N}} \in N^{\mathbb{N}}$ gegen $x \in M$ und ist U eine Umgebung von x, so enthält $U \cap N$ alle a_n für genügend große n. Insbesondere ist $U \cap N$ nichtleer. Es folgt $x \in \overline{N}$.

□

Bemerkung: In allgemeinen topologischen Räumen, die nicht von einer Halbmetrik d erzeugt werden, gibt es *keine* analoge Charakterisierung von Berührpunkten mit konvergenten Folgen. Das Analogon zu "2. $\Rightarrow$ 1." bleibt zwar (inklusive Beweis) allgemein richtig, aber das Analogon zu "1. $\Rightarrow$ 2." kann falsch werden.

Erinnern Sie sich an den Stetigkeitsbegriff aus der Analysis 1:

Definition 1.50 (Stetigkeit) Es seien $(M, \mathcal{T}_M)$ und $(N, \mathcal{T}_N)$ topologische Räume.

1. Es seien $x \in M$ und $M' \subseteq M$ eine Umgebung von x in $(M, \mathcal{T}_M)$. Eine Abbildung $f : M' \to N$ heißt *stetig* in x (bzgl. der Topologien $\mathcal{T}_M$ und $\mathcal{T}_N$), wenn für jede Umgebung U von $f(x)$ in $(N, \mathcal{T}_N)$ das Urbild $f^{-1}[U]$ eine Umgebung von x in $(M, \mathcal{T}_M)$ ist.

2. Eine Abbildung $f : M \to N$ heißt *stetig* (bzgl. der Topologien $\mathcal{T}_M$ und $\mathcal{T}_N$), wenn für jede offene Menge U in $(N, \mathcal{T}_N)$ das Urbild $f^{-1}[U]$ offen in $(M, \mathcal{T}_M)$ ist.

Notation: Wir schreiben auch kurz dafür: "$f : (M, \mathcal{T}_M) \to (N, \mathcal{T}_N)$ ist stetig." Stetigkeit für Abbildungen zwischen (halb-)metrischen oder (halb-)normierten Räumen ist als Stetigkeit bezüglich der zugehörigen Topologien zu verstehen.
Es ist leicht zu sehen, dass eine Abbildung f genau dann stetig ist, wenn sie stetig in allen Punkten ihres Definitionsbereichs ist.
Stetigkeit von f in einem Punkt x ist eine *lokale* Eigenschaft: Es kommt nur auf Einschränkung von f auf eine (beliebig kleine) Umgebung von x an.

Übung 1.51 *Zeigen Sie, dass eine Abbildung $f : (M, \mathcal{T}_M) \to (N, \mathcal{T}_N)$ genau dann stetig in $x \in M$ ist, wenn für alle Mengen $A \subseteq M$, die x als Berührpunkt bzgl. $\mathcal{T}_M$ besitzen, gilt: $f(x)$ ist ein Berührpunkt von $f[A]$ bzgl. $\mathcal{T}_N$. Folgern Sie: $f : (M, \mathcal{T}_M) \to (N, \mathcal{T}_N)$ ist genau dann stetig, wenn für alle Mengen $A \subseteq M$ gilt: $f[\overline{A}] \subseteq \overline{f[A]}$.*

Übung 1.52 *Zeigen Sie, dass eine Abbildung $f : (M, \mathcal{T}_M) \to (N, \mathcal{T}_N)$ zwischen zwei topologischen Räumen genau dann stetig ist, wenn für alle bzgl. $\mathcal{T}_N$ abgeschlossenen Mengen $A \subseteq N$ das Urbild $f^{-1}[A]$ abgeschlossen bzgl. $\mathcal{T}_M$ ist.*

Wie in der Analysis 1 führen wir noch folgende praktische Sprechweise ein:

Definition 1.53 (Konvergenz für $x \to x_0$) *Es seien $(M, \mathcal{T}_M)$ und $(N, \mathcal{T}_N)$ topologische Räume, $x_0 \in M$, U eine Umgebung von x_0, $f : U \setminus \{x_0\} \to N$ eine Abbildung und $y \in N$. Wir sagen, $f(x)$ konvergiert gegen y für $x \to x_0$, in Zeichen $f(x) \stackrel{x \to x_0}{\longrightarrow} y$, wenn es eine in x0 stetige Fortsetzung $F : U \to N$ von f in den Punkt x_0 mit dem Wert $F(x_0) = y$ gibt. Ist $(N, \mathcal{T}_N)$ ein Hausdorffraum, so kann es nur höchstens ein solches y geben; in diesem Fall schreiben wir auch $\lim_{x \to x_0} f(x) = y$.*

Beachten Sie, dass es hierbei nur auf die Werte von f in einer beliebigen Umgebung von x_0 ankommt; daher braucht f nicht auf dem ganzen Raum M definiert zu sein.
Genau wie in der Analysis 1 erhalten wir:

Lemma 1.54 (Stetigkeit der Komposition stetiger Abbildungen) *Es seien* $(M_1, \mathcal{T}_1)$, $(M_2, \mathcal{T}_2)$, *und* $(M_3, \mathcal{T}_3)$ *topologische Räume und* $f : M_1 \to M_2$, $g : M_2 \to M_3$ *Abbildungen. Dann gilt:*

1. *Ist* $f : (M_1, \mathcal{T}_1) \to (M_2, \mathcal{T}_2)$ *stetig in* $x \in M_1$ *und ist* $g : (M_2, \mathcal{T}_2) \to (M_3, \mathcal{T}_3)$ *stetig in* $f(x)$, *so ist auch* $g \circ f : (M_1, \mathcal{T}_1) \to (M_3, \mathcal{T}_3)$ *stetig in* x.

2. *Sind* $f : (M_1, \mathcal{T}_1) \to (M_2, \mathcal{T}_2)$ *und* $g : (M_2, \mathcal{T}_2) \to (M_3, \mathcal{T}_3)$ *stetig, so ist auch* $g \circ f : (M_1, \mathcal{T}_1) \to (M_3, \mathcal{T}_3)$ *stetig.*

Beweis:

1. Ist U eine Umgebung von $g(f(x))$ in $(M_3, \mathcal{T}_3)$, so ist $g^{-1}[U]$ eine Umgebung von $f(x)$ in $(M_2, \mathcal{T}_2)$, da g stetig in $f(x)$ ist. Dann ist $(g \circ f)^{-1}[U] = f^{-1}[g^{-1}[U]]$ eine Umgebung von x in $(M_1, \mathcal{T}_1)$, da f stetig in x ist.

2. Dies folgt unmittelbar aus Teil 1. Alternativ folgt es direkt so: Ist U offen in $(M_3, \mathcal{T}_3)$, so ist $g^{-1}[U]$ offen in $(M_2, \mathcal{T}_2)$, da g stetig ist. Dann ist $(g \circ f)^{-1}[U] = f^{-1}[g^{-1}[U]]$ offen in $(M_1, \mathcal{T}_1)$, da f stetig ist.

$\square$

Ebenso wie in Analysis 1 erhalten wir:

Satz 1.55 (Metrische Charakterisierung der Stetigkeit) Es seien (M, d) und (M', d') halbmetrische Räume, $f : M \to M'$ eine Abbildung und $x \in M$. Dann sind äquivalent:

1. f ist stetig in x (bzgl. $\mathcal{T}_d$ und $\mathcal{T}_{d'}$).

2. $\forall \epsilon > 0 \; \exists \delta > 0 \; \forall y \in M : \; [d(x, y) < \delta \implies d'(f(x), f(y)) < \epsilon]$.

3. Für jede Folge $(x_n)_{n \in \mathbb{N}}$ in M mit $x_n \overset{n\to\infty}{\longrightarrow}_{\mathcal{T}_d} x$ folgt $f(x_n) \overset{n\to\infty}{\longrightarrow}_{\mathcal{T}_{d'}} f(x)$. (Hierzu sagt man auch: "f ist *folgenstetig* in x".)

Beweis: "1. $\Rightarrow$ 2.": Es sei f stetig in x. Weiter sei $\epsilon > 0$ gegeben. Weil $U := U_\epsilon^{d'}(f(x))$ eine offene Menge ist, die $f(x)$ enthält, ist U eine Umgebung von $f(x)$ bzgl. $\mathcal{T}_{d'}$. Wegen der Stetigkeit von f in x ist $f^{-1}[U]$ eine Umgebung von x bzgl. $\mathcal{T}_d$, besitzt also $U_\delta^d(x)$ für ein geeignetes $\delta > 0$ als Teilmenge. Für beliebige $y \in M$ mit $d(x, y) < \delta$, d.h. $y \in U_\delta^d(x) \subseteq f^{-1}[U]$, folgt $f(y) \in U$, also $d'(f(y), f(x)) < \epsilon$.

"2. $\Rightarrow$ 3.": Es gelte 2. und $x_n \overset{n\to\infty}{\longrightarrow}_{\mathcal{T}_d} x$. Weiter sei $\epsilon > 0$ gegeben. Mit 2. wählen wir ein $\delta > 0$ mit

$$\forall y \in M : \; [d(x, y) < \delta \implies d'(f(x), f(y)) < \epsilon].$$

Wegen $x_n \overset{n\to\infty}{\longrightarrow}_{\mathcal{T}_d} x$ folgt $d(x, x_n) < \delta$ und damit $d'(f(x), f(x_n)) < \epsilon$ für alle genügend großen $n \in \mathbb{N}$. Damit ist gezeigt: $f(x_n) \overset{n\to\infty}{\longrightarrow}_{\mathcal{T}_{d'}} f(x)$.

"3. $\Rightarrow$ 1.": Kontraposition. Angenommen, f ist unstetig in x. Dann finden wir eine Umgebung U von $f(x)$ in $(M', \mathcal{T}_{d'})$, deren Urbild $f^{-1}[U]$ *keine* Umgebung von x in $(M, \mathcal{T}_d)$ ist. Insbesondere gilt $U^d_{1/n}(x) \not\subseteq f^{-1}[U]$ für alle $n \in \mathbb{N}$. Wir wählen eine Folge $(x_n)_{n\in\mathbb{N}}$ aus, so dass $x_n \in U^d_{1/n}(x) \setminus f^{-1}[U]$ für alle $n \in \mathbb{N}$ gilt. Dann folgt einerseits $x_n \overset{n\to\infty}{\longrightarrow}_{\mathcal{T}_d} x$, da jede Umgebung von x eine $1/n$-Umgebung von X für geeignetes $n \in \mathbb{N}$ enthält. Andererseits gilt *nicht* $f(x_n) \overset{n\to\infty}{\longrightarrow}_{\mathcal{T}_{d'}} f(x)$, denn die Umgebung U von $f(x)$ enthält kein einziges Folgenglied $f(x_n)$, $n \in \mathbb{N}$. Unsere Annahme impliziert also das Gegenteil von 3.

$\square$

Beispiel 1.56 (Nichtäquivalenz von Stetigkeit und Folgenstetigkeit in allgemeinen topologischen Räumen) Für Abbildungen zwischen allgemeinen topologischen Räumen ist die Stetigkeit manchmal *nicht* äquivalent zur Folgenstetigkeit. Zur Illustration ein *Gegenbeispiel:* Es sei M eine überabzählbare Menge und $\mathcal{S}$ die Menge aller Teilmengen von $U \subseteq M$, für die gilt: $U = \emptyset$ oder $M \setminus U$ ist abzählbar. Weiter sei $\mathcal{T}$ die Potenzmenge von M. Sowohl $\mathcal{S}$ als auch $\mathcal{T}$ sind Topologien auf M. Eine Folge $(x_n)_{n\in\mathbb{N}}$ in M konvergiert genau dann gegen ein $x \in M$, wenn sie schließlich den Wert x annimmt, d.h. wenn für alle genügend großen $n \in \mathbb{N}$ gilt: $x_n = x$; dies gilt sowohl bezüglich $\mathcal{S}$ als auch bezüglich $\mathcal{T}$. Also ist die Identität $\mathrm{id} : (M, \mathcal{S}) \to (M, \mathcal{T})$, $\mathrm{id}(x) = x$, folgenstetig (in jedem Punkt). Andererseits ist $\mathrm{id} : (M, \mathcal{S}) \to (M, \mathcal{T})$ nicht stetig, da $\mathcal{T}$ keine Teilmenge von $\mathcal{S}$ ist.

Gleichmäßige Stetigkeit wird ebenfalls ganz analog wie in der Analysis 1 definiert:

Definition 1.57 (Gleichmäßige Stetigkeit) Eine Abbildung $f : M \to N$ zwischen zwei halbmetrischen Räumen (M, d) und (N, d') wird *gleichmäßig stetig* genannt, wenn gilt:

$$\forall \epsilon > 0\ \exists \delta > 0\ \forall x, y \in M : \ (d(x, y) < \delta \ \Rightarrow \ d'(f(x), f(y)) < \epsilon).$$

Gleichmäßig stetige Abbildungen sind offensichtlich stetig. Die Umkehrung davon gilt im Allgemeinen nicht, wie aus der Analysis 1 bekannt ist.

Beispiel 1.58 Man beachte, dass gleichmäßige Stetigkeit ein *metrischer* und *kein topologischer* Begriff ist. Bezeichnet z.B. $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ die Standardmetrik auf den reellen Zahlen und $d' : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $d'(x, y) = |\arctan x - \arctan y|$ die Metrik auf $\mathbb{R}$ analog zu Beispiel 1.2.2., so erzeugen d und d' die gleiche Topologie. Dennoch ist zwar die Identität $\mathrm{id} : (\mathbb{R}, d) \to (\mathbb{R}, d')$, $\mathrm{id}(x) = x$ gleichmäßig stetig, $\mathrm{id} : (\mathbb{R}, d') \to (\mathbb{R}, d)$ jedoch nicht.

Wir zeigen nun die (gleichmäßige) Stetigkeit einiger wichtiger Abbildungen. Der Grundkörper $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ wird dabei stets mit der Standardmetrik und der Standardtopologie $\mathcal{T}_{\mathbb{K}}$ versehen.

Lemma 1.59 (Gleichmäßige Stetigkeit der (Halb-)norm) *Ist* $(V, \|\cdot\|)$ *ein halbnormierter Raum, so ist* $\|\cdot\| : V \to \mathbb{R}$ *gleichmäßig stetig.*

Beweis: Es sei $\epsilon > 0$ gegeben. Wir setzen $\delta := \epsilon$. Dann gilt für beliebige $x, y \in V$ mit $\|x - y\| < \delta$:

$$\big|\|x\| - \|y\|\big| \leq \|x - y\| < \delta = \epsilon.$$

□

Lemma 1.60 (Gleichmäßige Stetigkeit der (Halb-)Metrik) *Es sei (M, d) ein halbmetrischer Raum. Weiter sei $M \times M$ mit der Produkthalbmetrik d' zu d aus Übung 1.11 versehen. Dann ist $d : M \times M \to \mathbb{R}$ gleichmäßig stetig.*

Beweis: Es sei $\epsilon > 0$ gegeben. Wir setzen $\delta := \epsilon/2$. Dann folgt für beliebige $(x_1, x_2), (y_1, y_2) \in M \times M$ mit

$$d'((x_1, x_2), (y_1, y_2)) < \delta$$

mit der Dreiecksungleichung:

$$d(x_1, x_2) \leq d(x_1, y_1) + d(y_1, x_2) \leq d(x_1, y_1) + d(y_1, y_2) + d(y_2, x_2)$$

und daher

$$d(x_1, x_2) - d(y_1, y_2) \leq d(x_1, y_1) + d(x_2, y_2) \leq 2d'((x_1, x_2), (y_1, y_2)).$$

Mit vertauschten Rollen von x und y erhalten wir ebenso

$$d(y_1, y_2) - d(x_1, x_2) \leq 2d'((y_1, y_2), (x_1, x_2)),$$

also zusammen

$$|d(x_1, x_2) - d(y_1, y_2)| \leq 2d'((x_1, x_2), (y_1, y_2)) < 2\delta = \epsilon.$$

□

Übung 1.61 (Stetigkeit kanonischer Projektionen und Inklusionen)
Es seien $(M_1, d_1), \ldots, (M_n, d_n)$ halbmetrische Räume, $n \in \mathbb{N}$. Weiter sei das kartesische Produkt $M = M_1 \times \ldots \times M_n$ mit der Produkthalbmetrik d versehen. Zeigen Sie:

1. Für alle $i \in \{1, \ldots, n\}$ ist die kanonische Projektion $f_i : M \to M_i$, $f(x_1, \ldots, x_n) = x_i$, gleichmäßig stetig.

2. Für alle $i \in \{1, \ldots, n\}$ und alle $(x_1, \ldots, x_n) \in M$ ist die Inklusionsabbildung $g_i : M_i \to M$, definiert durch $(g_i(y))_j = x_j$ für $j \in \{1, \ldots, n\}$ mit $j \neq i$ und $(g_i(y))_i = y$ eine Isometrie und daher gleichmäßig stetig.

Lemma 1.62 (Stetigkeit der Addition) *Es sei $(V, \|\cdot\|)$ ein halbnormierter Raum über $\mathbb{K}$. Wir versehen $V \times V$ mit der Produktmetrik*

$$d : (V \times V)^2 \to \mathbb{R}, \quad d((x, y), (u, v)) = \max\{\|x - u\|, \|y - v\|\}.$$

Dann ist die Addition $+ : V \times V \to V$ gleichmäßig stetig.

Beweis: Es sei $\epsilon > 0$ gegeben. Wir wählen $\delta := \epsilon/2$. Dann folgt für alle $x, y, u, v \in V$ mit $d((x, y), (u, v)) < \delta$:

$$\begin{aligned}\|(x+y)-(u+v)\| &= \|(x-u)+(y-v)\| \leq \|x-u\| + \|y-v\| \\ &\leq 2\max\{\|x-u\|, \|y-v\|\} = 2d((x,y),(u,v)) < 2\delta = \epsilon.\end{aligned}$$

□

Insbesondere ist die Addition $+ : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ gleichmäßig stetig.
Für lineare Abbildungen wird der Stetigkeitsbegriff besonders einfach:

Lemma 1.63 (Charakterisierung der Stetigkeit linearer Abbildungen) *Es seien $(V, \|\cdot\|_V))$, $(W, \|\cdot\|_W))$ zwei halbnormierte Räume über $\mathbb{K}$ und $f : V \to W$ eine lineare Abbildung. Dann sind folgende Aussagen äquivalent:*

1. *f ist gleichmäßig stetig,*
2. *f ist stetig,*
3. *f ist stetig in 0,*
4. *$\exists C > 0\ \forall x \in V:\ \|f(x)\|_W \leq C\|x\|_V$.*

Beweis: 1. ⇒ 2. ⇒ 3. ist trivial. Zu 3. ⇒ 4.: Ist f stetig in 0, so gibt es ein $\delta > 0$ mit

$$f[U_\delta^{\|\cdot\|_V}(0)] \subseteq U_1^{\|\cdot\|_W}(0)$$

Wir setzen $C := 1/\delta$. Gegeben $x \in V$, unterscheiden wir zwei Fälle: Falls $\|x\|_V = 0$, so gilt für alle $\lambda \in \mathbb{K}$ die Aussage $\|\lambda x\|_V = |\lambda|\|x\|_V = 0 < \delta$, also nach Voraussetzung $|\lambda|\|f(x)\|_W = \|\lambda f(x)\|_W = \|f(\lambda x)\|_W < 1$. Hieraus folgt $\|f(x)\|_W = 0 = C\|x\|_V$. Im Fall $\|x\|_V > 0$ setzen wir $y = \alpha\|x\|_V^{-1}x$ für ein beliebiges $\alpha \in]0, \delta[$ und erhalten $\|y\|_V = \alpha < \delta$ und daher $\alpha\|x\|_V^{-1}\|f(x)\|_W = \|f(y)\|_W < 1$, also $\alpha\|f(x)\|_W < \|x\|_V$. Weil $\alpha \in]0, \delta[$ beliebig war, erhalten wir $\delta\|f(x)\|_W \leq \|x\|_V$ und damit die Behauptung $\|f(x)\|_W \leq C\|x\|_V$.
Zu 4. ⇒ 1.: Es sei $C > 0$ wie in 4. fixiert. Gegeben $\epsilon > 0$, wählen wir $\delta = \epsilon/C > 0$. Dann folgt für beliebige $x, y \in V$ mit $\|x - y\|_V < \delta$:

$$\|f(x) - f(y)\|_W = \|f(x-y)\|_W \leq C\|x-y\|_V < C\delta = \epsilon.$$

Also ist f gleichmäßig stetig.

□

Übung 1.64 (Äquivalenz von Normen) Es seien $\|\cdot\|_A$ und $\|\cdot\|_B$ zwei Halbnormen auf dem gleichen $\mathbb{K}$-Vektorraum V. Zeigen Sie die Äquivalenz der folgenden drei Aussagen:

1. Die Identität $\mathrm{id} : (V, \|\cdot\|_A) \to (V, \|\cdot\|_B)$, $\mathrm{id}(x) = x$, ist stetig,
2. $\mathcal{T}_{\|\cdot\|_B} \subseteq \mathcal{T}_{\|\cdot\|_A}$,

3. $\exists C \geq 0 \, \forall x \in V: \ \|x\|_B \leq C\|x\|_A$.

Insbesondere erzeugen die beiden Halbnormen genau dann die gleiche Topologie, wenn gilt:

$$\exists c > 0, C > 0 \, \forall x \in V: \ c\|x\|_A \leq \|x\|_B \leq C\|x\|_A.$$

In diesem Fall heißen die beiden Halbnormen $\|\cdot\|_A$ und $\|\cdot\|_B$ *äquivalent*.

Übung 1.65 Zeigen Sie für $1 \leq p < q \leq \infty$, $n \in \mathbb{N}$, $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $x \in \mathbb{K}^n$:

$$n\|x\|_\infty \geq \|x\|_p \geq \|x\|_q \geq \|x\|_\infty.$$

Insbesondere sind alle Normen $\|\cdot\|_p$ auf $\mathbb{K}^n$, $1 \leq p \leq \infty$, zueinander äquivalent. Später werden wir sehen, dass dies sogar für *alle* Normen auf $\mathbb{K}^n$ gilt; siehe Satz 1.169.

Beispiel 1.66 (Integraloperatoren) Es seien $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, $a, b \in \mathbb{R}$ mit $a < b$ und $K : [a,b] \times [a,b] \to \mathbb{K}$ gleichmäßig stetig. (Später, in Übung 1.175, sehen wir mit einem Kompaktheitsargument, dass jede stetige Funktion auf $[a,b]^2$ auch gleichmäßig stetig ist.) Wir versehen den Raum $C([a,b], \mathbb{K})$ aller stetigen Funktionen $f : [a,b] \to \mathbb{K}$ mit der Supremumsnorm $\|f\|_\infty := \sup_{x\in[a,b]} |f(x)|$. Für $f \in C([a,b], \mathbb{K})$ definieren wir

$$L(f) : [a,b] \to \mathbb{K}, L(f)(x) = \int_a^b K(x,y)f(y)\,dy \text{ für } x \in [a,b].$$

Dann ist $L(f)$ wieder stetig. In der Tat: Ist $(x_n)_{n\in\mathbb{N}}$ eine Folge mit Werten in $[a,b]$ mit Limes $x \in [a,b]$, so konvergiert wegen der gleichmäßigen Stetigkeit von K und der Beschränktheit von f die Funktionenfolge $([a,b] \ni y \mapsto K(x_n,y)f(y))_{n\in\mathbb{N}}$ für $n \to \infty$ gleichmäßig gegen $[a,b] \ni y \mapsto K(x,y)f(y)$. Hieraus folgt

$$L(f)(x_n) = \int_a^b K(x_n,y)f(y)\,dy \overset{n\to\infty}{\longrightarrow} \int_a^b K(x,y)f(y)\,dy = L(f)(x)$$

Die Abbildung

$$L : C([a,b], \mathbb{K}) \to C([a,b], \mathbb{K})$$

ist also wohldefiniert. Sie ist auch gleichmäßig stetig bezüglich der Supremumsnorm $\|\cdot\|_\infty$. Dies folgt aus Lemma 1.63, denn L ist linear und es gilt

$$\begin{aligned} \|L(f)\|_\infty &= \sup_{x\in[a,b]} \left| \int_a^b K(x,y)f(y)\,dy \right| \leq \sup_{x\in[a,b]} \int_a^b |K(x,y)||f(y)|\,dy \\ &\leq \sup_{x\in[a,b]} \int_a^b |K(x,y)|\,dy\,\|f\|_\infty \end{aligned}$$

Beachten Sie, dass

$$\sup_{x\in[a,b]} \int_a^b |K(x,y)|\,dy \leq (b-a) \sup_{x,y\in[a,b]} |K(x,y)| < \infty$$

gilt, da aus der gleichmäßigen Stetigkeit von K und der Kompaktheit von $[a,b]$ auch die Beschränktheit von K folgt.

Übung 1.67 (Operatornorm) Es seien $(V, \|\cdot\|_V)$ und $(W, \|\cdot\|_W)$ zwei halbnormierte Räume über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und

$$\mathcal{B}(V,W) := \{L : (V, \|\cdot\|_V) \to (W, \|\cdot\|_W) | \ L \text{ ist stetig und linear}\}$$

Für $L \in \mathcal{B}(V,W)$ setzen wir

$$\|L\|_{V\to W} := \inf\{C \geq 0 | \ \forall x \in V : \ \|L(x)\|_W \leq C\|x\|_V\}.$$

Zeigen Sie, dass $\|\cdot\|_{V\to W} : \mathcal{B}(V,W) \to \mathbb{R}$ eine Halbnorm auf $\mathcal{B}(V,W)$ ist. Zeigen Sie weiter, dass $\|\cdot\|_{V\to W}$ eine Norm ist, wenn auch $\|\cdot\|_W$ eine Norm ist. Die Norm $\|\cdot\|_{V\to W} : \mathcal{B}(V,W) \to \mathbb{R}$ wird *Operatornorm* auf $\mathcal{B}(V,W)$ genannt. Zeigen Sie auch

$$\|L\|_{V\to W} = \sup\{\|L(x)\|_W | \ x \in V, \ \|x\|_V \leq 1\}.$$

Übung 1.68 Es sei $A = (a_{i,j})_{i=1,\ldots,n,\ j=1,\ldots,m} \in \mathbb{K}^{n\times m}$ eine $n \times m$-Matrix mit Einträgen in $\mathbb{K}$ und $L_A : \mathbb{K}^m \to \mathbb{K}^n$, $L_A(x) = Ax$ die zugehörige lineare Abbildung. Zeigen Sie

$$\|L_A\|_{\mathbb{K}^m\to\mathbb{K}^n} \leq \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m}|a_{i,j}|^2}$$

für die Operatornorm von $L_A : (\mathbb{K}^m, \|\cdot\|_2) \to (\mathbb{K}^n, \|\cdot\|_2)$. (*Hinweis:* Cauchy-Schwarz-Ungleichung.) Zeigen Sie mit einem Gegenbeispiel, dass

$$\|L_A\|_{\mathbb{K}^m\to\mathbb{K}^n} < \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m}|a_{i,j}|^2}$$

möglich ist.

Für $m, n \in \mathbb{N}$, seien Normen auf $\mathbb{K}^m$ und $\mathbb{K}^n$ und eine Matrix $A \in \mathbb{K}^{n\times m}$ gegeben. Die zugehörige Operatornorm der linearen Abbildung $L_A : \mathbb{K}^m \to \mathbb{K}^n$, $L_A(x) = Ax$ wird auch *Matrixnorm* von A genannt.

Übung 1.69 (Submultiplikativität der Operatornorm) *Lösen Sie eine der beiden folgenden Aufgaben:*

(a) Es seien $A \in \mathbb{K}^{n\times m}$, $B \in \mathbb{K}^{m\times l}$ mit $l, m, n \in \mathbb{N}$. Zeigen Sie für die Matrixnormen $\|\cdot\|$ zu beliebigen Normen auf $\mathbb{K}^l$, $\mathbb{K}^m$ und $\mathbb{K}^n$:

$$\|AB\| \leq \|A\|\|B\|.$$

(b) Es seien $(U, \|\cdot\|_U)$, $(V, \|\cdot\|_V)$ und $(W, \|\cdot\|_W)$ normierte Räume. Weiter seien $L : (U, \|\cdot\|_U) \to (V, \|\cdot\|_V)$ und $M : (V, \|\cdot\|_V) \to (W, \|\cdot\|_W)$ stetige lineare Abbildungen. Zeigen Sie, dass für die zugehörigen Operatornormen $\|\cdot\|_{U\to V}$, $\|\cdot\|_{V\to W}$ und $\|\cdot\|_{U\to W}$ gilt:

$$\|M \circ L\|_{U\to W} \leq \|M\|_{V\to W}\|L\|_{U\to V}.$$

Nun beschäftigen wir uns mit der Stetigkeit von bilinearen und sesquilinearen Abbildungen:

Lemma 1.70 (Charakterisierung der Stetigkeit von Produkten) *Es seien* $(V, \|\cdot\|_V)$, $(W, \|\cdot\|_W)$ *und* $(U, \|\cdot\|_U)$ *drei halbnormierte Räume über* $\mathbb{K}$ *und* $\langle \cdot, \cdot \rangle : V \times W \to U$ *eine Abbildung mit folgenden Eigenschaften:*

$$\begin{aligned}
&\forall x, y \in V \ \forall z \in W : \ \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle \,, \\
&\forall x \in V \ \forall y, z \in W : \ \langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle \,, \\
&\exists C \geq 0 \ \forall x \in V \ \forall y \in W : \| \langle x, y \rangle \|_U \leq C \|x\|_V \|y\|_W .
\end{aligned}$$

Wir versehen $V \times W$ *mit der Produktmetrik*

$$d : (V \times W)^2 \to \mathbb{R}, \quad d((x, y), (u, v)) = \max\{\|x - u\|_V, \|y - v\|_W\}.$$

Dann ist die Abbildung $\langle \cdot, \cdot \rangle : V \times W \to U$ *stetig. Für jedes* $r > 0$ *ist die Einschränkung* $\langle \cdot, \cdot \rangle : U_r^d(0, 0) \to U$ *sogar gleichmäßig stetig.*

Beweis: Es sei $C \geq 0$ wie in der dritten Voraussetzung gegeben. Weiter seien $x \in V$ und $y \in W$ gegeben. Wir nehmen $r > 0$ so groß, dass $(x, y) \in U_r^d(0, 0)$ gilt. Nun sei $\epsilon > 0$ gegeben. Wir wählen $\delta > 0$ so klein, dass $C\delta(2r + \delta) < \epsilon$ gilt. Dann folgen für beliebige $u \in V$ und $v \in W$ mit $d((x, y), (u, v)) < \delta$ die Ungleichungen $\|x - u\|_V < \delta$ und $\|y - v\|_W < \delta$, also

$$\begin{aligned}
\| \langle x, y \rangle - \langle u, v \rangle \|_U &\leq \| \langle x, y \rangle - \langle u, y \rangle \|_U + \| \langle u, y \rangle - \langle u, v \rangle \|_U \\
&= \| \langle x - u, y \rangle \|_U + \| \langle u, y - v \rangle \|_U \\
&\leq C\|x - u\|_V \|y\|_W + C\|u\|_V \|y - v\|_W \\
&\leq C\delta \|y\|_W + C\|u\|_V \delta \\
&\leq C\delta \|y\|_W + C(\|x\|_V + \|u - x\|_V)\delta \\
&\leq C\delta(\|y\|_W + \|x\|_V + \delta) \leq C\delta(2 \max\{\|y\|_W, \|x\|_V\} + \delta) \\
&= C\delta(2d((0, 0), (x, y)) + \delta) \\
&\leq C\delta(2r + \delta) < \epsilon.
\end{aligned}$$

Das beweist die behauptete Stetigkeit. Weil die Wahl von δ nur von r und von ϵ, nicht jedoch von $x, y \in U_r^d(0, 0)$ abhängt, zeigt es auch die behauptete gleichmäßige Stetigkeit auf $U_r^d(0, 0)$.

□

Wir nennen eine Teilmenge $N \subseteq M$ in einem halbmetrischen Raum (M, d) *beschränkt*, wenn die Einschränkung der Metrik d auf $N \times N$ beschränkt ist. Lemma 1.70 behauptet also die gleichmäßige Stetigkeit von $\langle \cdot, \cdot \rangle$ auf beschränkten Mengen.

Beispiel 1.71 Als Beispiele zu Lemma 1.70 erhalten wir die Stetigkeit und die gleichmäßige Stetigkeit auf beschränkten Mengen der folgenden Abbildungen:

1. der Multiplikation $\cdot : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$,

2. der Skalarmultiplikation $\cdot : \mathbb{K} \times V \to V$ zu einem halbnormierten Raum $(V, \|\cdot\|_V)$ über $\mathbb{K}$,

3. des Skalarprodukts $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$ zu einem Prähilbertraum $(V, \langle \cdot, \cdot \rangle)$ über $\mathbb{K}$ wegen der Cauchy-Schwarz-Ungleichung,

4. der Bilinearform $\langle \cdot, \cdot \rangle : \ell^p(I) \times \ell^q(I) \to \mathbb{K}$, $\langle f, g \rangle = \sum_{j \in I} f(j)g(j)$ für konjugierte $p, q \in [1, \infty]$ wegen der Hölder-Ungleichung,

5. des Matrixprodukts
$$\cdot : \mathbb{K}^{n \times m} \times \mathbb{K}^{m \times l} \to \mathbb{K}^{n \times l}$$
wegen Übung 1.69.

Übung 1.72 Es sei p eine Primzahl und $d_p : \mathbb{Z} \times \mathbb{Z} \to \mathbb{R}$ die p-adische Metrik aus Beispiel 1.2.3, sowie $v_p(k) \in \mathbb{N}_0 \cup \{\infty\}$ die Vielfachheit p in der Primfaktorzerlegung von $k \in \mathbb{Z}$. Weiter sei d die Produktmetrik zu d_p auf $\mathbb{Z} \times \mathbb{Z}$.

(a) Zeigen Sie für $k, l \in \mathbb{Z}$:
$$v_p(kl) = v_p(k) + v_p(l).$$
Sie dürfen dabei aus der elementaren Zahlentheorie für ganze Zahlen m, n als bekannt voraussetzen, dass mn nicht durch p teilbar ist, wenn m und n nicht durch p teilbar sind.

(b) Folgern Sie, dass die arithmetischen Operationen
$$+ : (\mathbb{Z} \times \mathbb{Z}, d) \to (\mathbb{Z}, d_p) \text{ und } \cdot : (\mathbb{Z} \times \mathbb{Z}, d) \to (\mathbb{Z}, d_p)$$
gleichmäßig stetig sind.

Übung 1.73 (Gleichmäßige Stetigkeit von Auswertungsabbildungen) Es sei I eine abzählbare Indexmenge, $j \in I$, $1 \leq p \leq \infty$ und $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Zeigen Sie, dass die Auswertungsabbildung $\delta_j : (\ell^p(I, \mathbb{K}), \|\cdot\|_p) \to (\mathbb{K}, |\cdot|)$, $\delta_j(f) := f(j)$, gleichmäßig stetig ist.

Übung 1.74 (Dualität zwischen ℓ^p und ℓ^q) Es seien $p, q \in]1, \infty[$ zueinander konjugiert, I eine abzählbare Indexmenge, $e_j = (\delta_{j,k})_{k \in I}$, $j \in I$, die kanonischen Einheitsvektoren in $\ell^p(I, \mathbb{K})$ mit $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, $f : \ell^p(I, \mathbb{K}) \to \mathbb{K}$ eine $\mathbb{K}$-lineare Abbildung und $y := (f(e_j))_{j \in I}$. Zeigen Sie die Äquivalenz der beiden folgenden Aussagen:

1. $f : (\ell^p(I, \mathbb{K}), \|\cdot\|_p) \to (\mathbb{K}, |\cdot|)$ ist stetig,

2. $y \in \ell^q(I, \mathbb{K})$ und $\forall x \in \ell^p(I, \mathbb{K}) : \; f(x) = \sum_{j \in I} x(j)y(j)$.

Übung 1.75 1. Es seien $(M, \mathcal{T}_M)$ ein topologischer Raum, $(N, \mathcal{T}_N)$ ein Hausdorffraum, $f, g : (M, \mathcal{T}_M) \to (N, \mathcal{T}_N)$ zwei stetige Abbildungen und $A \subseteq M$ eine dichte Teilmenge, so dass die Einschränkungen von f und von g auf A übereinstimmen. Zeigen Sie: $f = g$.

2. Es seien $(M, \mathcal{T}_M)$ ein topologischer Raum, $f, g : (M, \mathcal{T}_M) \to \mathbb{R}$ zwei stetige Abbildungen und $A \subseteq M$ eine dichte Teilmenge, so dass für alle $x \in A$ gilt: $f(x) \leq g(x)$. Zeigen Sie: $f \leq g$. *Hinweis:* Betrachten Sie das Urbild der abgeschlossenen Menge $]-\infty, 0] \subset \mathbb{R}$ unter der Abbildung $f - g$. Zeigen Sie, dass $f - g$ stetig ist.

1.6 Initial- und Finaltopologie

Lemma 1.76 (Initialtopologie einer Abbildung) Es seien X eine Menge, $(Y, \mathcal{S})$ ein topologischer Raum und $f : X \to Y$ eine Abbildung. Dann ist

$$\mathcal{T} := \{f^{-1}[U] |\, U \in \mathcal{S}\} \tag{10}$$

die *kleinste* Topologie auf X, bezüglich der $f : X \to (Y, \mathcal{S})$ stetig ist. Sie wird *Initialtopologie* auf X (bezüglich $f : X \to (Y, \mathcal{S})$) genannt. Die Abbildung $f : (X, \mathcal{T}) \to (Y, \mathcal{S})$ nennen wir dann *initial*.

Beweis:

- $\mathcal{T}$ ist eine Topologie auf X, denn $\emptyset = f^{-1}[\emptyset] \in \mathcal{T}$, $X = f^{-1}[Y] \in \mathcal{T}$, und $\mathcal{T}$ ist abgeschlossen unter endlicher Durchschnittsbildung und beliebiger Vereinigungsbildung, da $\mathcal{S}$ diese Eigenschaften hat und Urbildbildung mit endlicher Durchschnittsbildung und beliebiger Vereinigungsbildung vertauschbar ist.

- $f : (X, \mathcal{T}) \to (Y, \mathcal{S})$ ist nach Definition von $\mathcal{T}$ stetig.

- Ist $\mathcal{T}'$ eine weitere Topologie auf X, so dass $f : (X, \mathcal{T}) \to (Y, \mathcal{S})$ stetig ist, so gilt $\mathcal{T} \subseteq \mathcal{T}'$. In der Tat: Ist $V = f^{-1}[U] \in \mathcal{T}$ mit $U \in \mathcal{S}$, so folgt $V \in \mathcal{T}'$, da $f : (X, \mathcal{T}) \to (Y, \mathcal{S})$ stetig ist.

□

Beispiel 1.77 (Teilraumtopologie) Ist $(Y, \mathcal{S})$ ein topologischer Raum, $X \subseteq Y$ eine Teilmenge und $f : X \to Y$, $f(x) = x$, die Inklusionsabbildung, so gilt $f^{-1}[U] = U \cap X$ für alle $U \subseteq Y$. Die Initialtopologie zu $f : X \to (Y, \mathcal{S})$ wird also durch

$$\mathcal{T} = \{U \cap X |\, U \in \mathcal{S}\}$$

gegeben. Wir nennen diese Topologie auch die *Teilraumtopologie* oder *Relativtopologie* auf X (bezüglich $(Y, \mathcal{S})$). Die Elemente von $\mathcal{T}$ nennen wir *relativ zu X offen* oder kurz *in X offen* (bzgl. $\mathcal{S}$). Die in der Relativtopologie abgeschlossenen Mengen nennen wir *relativ zu X abgeschlossen* oder *in X abgeschlossen*.

Für Teilmengen X von $Y = \mathbb{R}$ mit der Standardtopologie ist Ihnen dieser Begriff aus der Analysis 1 bekannt.

Übung 1.78 (Transitivität der Dichtheit) Es sei $(X, \mathcal{T}_X)$ ein topologischer Raum, $(Y, \mathcal{T}_Y)$ ein dichter Teilraum von $(X, \mathcal{T}_X)$ und $(Z, \mathcal{T}_Z)$ ein dichter Teilraum von $(Y, \mathcal{T}_Y)$. Zeigen Sie, dass $(Z, \mathcal{T}_Z)$ ein dichter Teilraum von $(X, \mathcal{T}_X)$ ist.

Übung 1.79 (Initial- und Teilraumtopologie bei halbmetrischen Räumen) Zeigen Sie: Ist (Y, d) ein halbmetrischer Raum, X eine Menge und $f : X \to Y$ eine Abbildung, so ist $d' : X \times X \to \mathbb{R}$, $d'(x, y) = d(f(x), f(y))$ eine Halbmetrik auf X, die die Initialtopologie zu $f : X \to (Y, \mathcal{T}_d)$ erzeugt. Insbesondere wird für $X \subseteq Y$ die Teilraumtopologie auf X durch die Einschränkung von $d : Y \times Y \to \mathbb{R}$ auf $X \times X$ erzeugt.

Betrachten wir nun statt *einer* Abbildung $f : X \to (Y, \mathcal{S})$ eine *Familie* $(f_i : X \to (Y_i, \mathcal{S}_i))_{i \in I}$ von Abbildungen in topologische Räume $(Y_i, \mathcal{S}_i)$. Im allgemeinen ist das analog zu (10) gebildete Mengensystem

$$\mathcal{M} = \{f_i^{-1}[U] |\ i \in I, U \in \mathcal{S}_i\}$$

keine Topologie.

Lemma/Definition 1.80 (von einem Mengensystem erzeugte Topologie) Zu jedem Mengensystem $\mathcal{M} \subseteq \mathcal{P}(X)$ über einer Menge X gibt es eine kleinste Topologie auf X, die $\mathcal{M}$ umfasst, nämlich das Mengensystem $\mathrm{Top}_X(\mathcal{M})$ aller beliebigen Vereinigungen

$$\bigcup_{j \in J} U_j$$

(mit beliebiger Indexmenge J) von endlichen Durchschnitten der Gestalt

$$U_j = \bigcap_{k=1}^{n_j} V_{j,k}, \quad (n_j \in \mathbb{N}, V_{j,k} \in \mathcal{M} \cup \{X\}).$$

Die Topologie $\mathrm{Top}_X(\mathcal{M})$ wird die vom Mengensystem $\mathcal{M}$ erzeugte Topologie auf X genannt.

Beweis: Offensichtlich gilt $\mathcal{M} \cup \{X\} \subseteq \mathrm{Top}_X(\mathcal{M})$. Da jede Topologie auf X abgeschlossen unter endlicher Durchschnittsbildung und beliebiger Vereinigungsbildung ist und X enthält, gilt $\mathrm{Top}_X(\mathcal{M}) \subseteq \mathcal{T}$ für jede Topologie $\mathcal{T}$ auf X mit $\mathcal{M} \subseteq \mathcal{T}$.
Weiter ist $\mathrm{Top}_X(\mathcal{M})$ eine Topologie auf X: In der Tat ist $\emptyset = \bigcup_{j \in \emptyset} U_j \in \mathrm{Top}_X(\mathcal{M})$ und $X \in \mathrm{Top}_X(\mathcal{M})$. Sind nun

$$U = \bigcup_{j \in J} U_j, U' = \bigcup_{j' \in J'} U'_{j'} \in \mathrm{Top}_X(\mathcal{M})$$

mit

$$U_j = \bigcap_{k=1}^{n_j} V_{j,k}, \quad (n_j \in \mathbb{N}, V_{j,k} \in \mathcal{M} \cup \{X\}),$$

$$U'_{j'} = \bigcap_{k=1}^{n'_{j'}} V'_{j',k}, \quad (n'_{j'} \in \mathbb{N}, V'_{j',k} \in \mathcal{M} \cup \{X\})$$

so folgt auch

$$U \cap U' = \bigcup_{(j,j') \in J \times J'} U_j \cap U'_{j'} \in \mathrm{Top}_X(\mathcal{M}),$$

so dass $\mathrm{Top}_X(\mathcal{M})$ abgeschlossen unter endlicher Durchschnittsbildung ist. Die Abgeschlossenheit von $\mathrm{Top}_X(\mathcal{M})$ unter beliebiger Vereingungsbildung ist klar aus der Definition.

$\square$

Übung 1.81 Es seien $(X', \mathcal{T}')$ und $(X, \mathcal{T})$ topologische Räume, $\mathcal{M} \subseteq \mathcal{P}(X)$ ein Mengensystem mit $\mathrm{Top}_X(\mathcal{M}) = \mathcal{T}$ und $g : X' \to X$ eine Abbildung. Zeigen Sie, dass $g : (X', \mathcal{T}') \to (X, \mathcal{T})$ genau dann stetig ist, wenn für alle $U \in \mathcal{M}$ gilt: $g^{-1}[U] \in \mathcal{T}'$. Verwenden Sie dazu, dass Urbildbildung mit beliebigen Vereinigungsbildungen und endlichen Durchschnittsbildungen vertauscht.

Definition 1.82 (Initialtopologie einer Familie von Abbildungen) Gegeben sei eine *Familie* $(f_i : X \to (Y_i, \mathcal{S}_i))_{i \in I}$ von Abbildungen von einer Menge X in topologische Räume $(Y_i, \mathcal{S}_i)$. Die Topologie

$$\mathcal{T} := \mathrm{Top}_X(\{f_i^{-1}[U] |\ i \in I, U \in \mathcal{S}_i\})$$

wird *Initialtopologie* der Familie $(f_i : X \to (Y_i, \mathcal{S}_i))_{i \in I}$ genannt. Die Familie $(f_i : (X, \mathcal{T}) \to (Y_i, \mathcal{S}_i))_{i \in I}$ heißt dann *initial*.

Die Initialtopologie ist also die *kleinste* Topologie auf X, bezüglich der alle Abbildungen $(f_i : X \to (Y_i, \mathcal{S}_i))_{i \in I}$ stetig sind.

Übung 1.83 (Initialtopologie bei halbmetrischen Räumen)

1. Es seien (Y_i, d_i), $i = 1, \ldots, n$, halbmetrische Räume, $(X, \mathcal{T})$ ein topologischer Raum und $(f_i : (X, \mathcal{T}) \to (Y_i, \mathcal{T}_{d_i}))_{i=1,\ldots,n}$ initial. Zeigen Sie, dass $\mathcal{T}$ von der Halbmetrik

   $$d(x, y) = \max_{i=1\ldots,n} d_i(f_i(x), f_i(y)), \quad x, y \in X$$

   erzeugt wird. Insbesondere ist die Topologie zur Produktmetrik auf $Y_1 \times \ldots \times Y_n$ (siehe Übung 1.11) die Initialtopologie zu den kanonischen Projektionen.

2. Es seien (Y_i, d_i), $i \in \mathbb{N}$, halbmetrische Räume, $(X, \mathcal{T})$ ein topologischer Raum und $(f_i : (X, \mathcal{T}) \to (Y_i, \mathcal{T}_{d_i}))_{i=1,\ldots,n}$ initial. Zeigen Sie, dass $\mathcal{T}$ von der Halbmetrik

   $$d(x, y) = \sum_{i \in \mathbb{N}} \min\{2^{-i}, d_i(f_i(x), f_i(y))\}, \quad x, y \in X$$

   erzeugt wird.

Definition 1.84 Es seien $(Y_i, \mathcal{S}_i)_{i \in I}$ topologische Räume und

$$X = \prod_{i \in I} Y_i$$

deren kartesisches Produkt, also die Menge aller Familien $(y_i)_{i \in I}$ mit $y_i \in Y_i$ für alle $i \in I$. Die Initialtopologie zu den kanonischen Projektionen $(\pi_j : X \to Y_i)_{j \in I}$, $\pi_j((y_i)_{i \in I}) = y_j$, wird *Produkttopologie* genannt.

Lemma 1.85 (Stetigkeit von Abbildungen in eine initiale Quelle) *Gegeben sei eine initiale Familie $(f_i : (X, \mathcal{T}) \to (Y_i, \mathcal{S}_i))_{i \in I}$ von einem topologischen Raum $(X, \mathcal{T})$ in topologische Räume $(Y_i, \mathcal{S}_i)$. Außerdem sei ein weiterer topologischer Raum $(X', \mathcal{T}')$ und eine Abbildung $g : X' \to X$ gegeben. Dann ist $g : (X', \mathcal{T}') \to (X, \mathcal{T})$ genau dann stetig, wenn alle Abbildungen $f_i \circ g : (X', \mathcal{T}') \to (Y_i, \mathcal{S}_i)$, $i \in I$ stetig sind.*

Beweis: Ist $g : (X', \mathcal{T}') \to (X, \mathcal{T})$ stetig, so sind auch alle $f_i \circ g : (X', \mathcal{T}') \to (Y_i, \mathcal{S}_i)$, $i \in I$, stetig, weil alle $f_i : (X, \mathcal{T}) \to (Y_i, \mathcal{S}_i)$ stetig sind.

Es sei umgekehrt $f_i \circ g : (X', \mathcal{T}') \to (Y_i, \mathcal{S}_i)$ für alle $i \in I$, stetig. Dann gilt für alle $i \in I$ und $U \in \mathcal{S}_i$: $g^{-1}[f_i^{-1}[U]] \in \mathcal{T}'$. Weil das System der Mengen $f_i^{-1}[U]$ mit $i \in I$, $U \in \mathcal{S}_i$ die Initialtopologie $\mathcal{T}$ erzeugt, folgt die Stetigkeit von g wegen Übung 1.81.

□

Wir besprechen nun ein Gegenstück zur Initialtopologie, bei dem gewissermaßen alle Pfeilrichtungen und die Rollen von "groß" und "klein" umgedreht werden:

Lemma/Definition 1.86 (Finaltopologie) Es seien Y eine Menge, $(X_i, \mathcal{T}_i)_{i \in I}$ eine Familie von topologischen Räumen und $(f_i : X_i \to Y)_{i \in I}$ eine Familie von Abbildungen in Y. Dann ist

$$\mathcal{T} := \{U \subseteq Y \mid \forall i \in I : f_i^{-1}[U] \in \mathcal{T}_i\}$$

die *größte* Topologie auf Y, bezüglich der alle $f_i : (X_i, \mathcal{T}_i) \to Y$ stetig sind. Sie wird *Finaltopologie* auf X (bezüglich der gegebenen Daten) genannt. Die Familie von Abbildungen $(f_i : (X_i, \mathcal{T}_i) \to (Y, \mathcal{T}))_{i \in I}$ nennen wir dann *final*.

Beweis:

- $\mathcal{T}$ ist eine Topologie: $\emptyset \in \mathcal{T}$ gilt wegen $f_i^{-1}[\emptyset] = \emptyset \in \mathcal{T}_i$ für alle $i \in I$. $Y \in \mathcal{T}$ folgt ebenso wegen $f_i^{-1}[Y] = X_i \in \mathcal{T}_i$ für alle $i \in I$. Sind nun $U, V \in \mathcal{T}$, so folgt für alle $i \in I$: $f_i^{-1}[U], f_i^{-1}[V] \in \mathcal{T}_i$, also

  $$f_i^{-1}[U \cap V] = f_i^{-1}[U] \cap f_i^{-1}[V] \in \mathcal{T}_i$$

  und damit $U \cap V \in \mathcal{T}$. Ist schließlich $(U_j)_{j \in J}$ eine Familie von Mengen in $\mathcal{T}$, so folgt für alle $i \in I$ und $j \in J$: $f_i^{-1}[U_j] \in \mathcal{T}_i$, also

  $$f_i^{-1}\left[\bigcup_{j \in J} U_j\right] = \bigcup_{j \in J} f_i^{-1}[U_j] \in \mathcal{T}_i, \quad (i \in I)$$

  und daher $\bigcup_{j \in J} U_j \in \mathcal{T}$.

- Alle $f_i : (X_i, \mathcal{T}_i) \to (Y, \mathcal{T})$, $i \in I$, sind stetig. Dies folgt unmittelbar aus der Definition von $\mathcal{T}$.

- Ist $\mathcal{S}$ eine Topologie auf Y, bezüglich der alle $f_i : (X_i, \mathcal{T}_i) \to (Y, \mathcal{S})$, $i \in I$, stetig sind, so folgt $\mathcal{S} \subseteq \mathcal{T}$. In der Tat: Ist $U \in \mathcal{S}$, so folgt für alle $i \in I$: $f_i^{-1}[U] \in \mathcal{T}_i$, da $f_i : (X_i, \mathcal{T}_i) \to (Y, \mathcal{S})$ stetig ist, also $U \in \mathcal{T}$ nach Definition von $\mathcal{T}$.

$\square$

Übung 1.87 (Stetigkeit von Abbildungen von einer finalen Senke) Beweisen Sie folgendes Gegenstück zu Satz 1.85:
Gegeben sei eine finale Familie $(f_i : (X_i, \mathcal{T}_i) \to (Y, \mathcal{S}))_{i \in I}$ von topologischen Räumen $(X_i, \mathcal{T}_i)$ in einen topologischen Raum $(Y, \mathcal{S})$. Außerdem sei ein weiterer topologischer Raum $(Y', \mathcal{S}')$ und eine Abbildung $g : Y \to Y'$ gegeben. Dann ist $g : (Y, \mathcal{S}) \to (Y', \mathcal{S}')$ genau dann stetig, wenn alle Abbildungen $g \circ f_i : (X_i, \mathcal{T}_i) \to (Y', \mathcal{S}')$, $i \in I$ stetig sind.

Definition 1.88 (Quotiententopologie) Es sei $(Y, \mathcal{S})$ ein topologischer Raum, $\sim$ eine Äquivalenzrelation auf Y, $X = Y/{\sim}$ der Quotientenraum, und $f : Y \to Y/{\sim}$, $f(x) = [x]_\sim$ die kanonische Abbildung. Die Finaltopologie $\mathcal{T}$ zur Abbildung $f : (Y, \mathcal{S}) \to \mathcal{Y}/{\sim}$ wird *Quotiententopologie* zu Y und $\sim$ genannt.

Beispiel 1.89 *Es sei (M, d) ein halbmetrischer Raum, $\sim$ die Relation “Abstand 0”, und d' die auf dem Quotientenraum $M/{\sim}$ induzierte Metrik aus aus Lemma 1.9. Dann erzeugt d' die Quotiententopologie auf $M/{\sim}$.*

Beispiel 1.90 Es sei $S^1 := \{z \in \mathbb{C} |; |z| = 1\}$ der Einheitskreis und $g : \mathbb{R} \to S^1$, $g(x) = e^{ix}$. Dann ist g final. In der Tat: Ist $U \subseteq S^1$ eine Menge, so ist U genau dann offen in S^1, wenn $g^{-1}[U]$ offen in $\mathbb{R}$ ist. Die Richtung “$\Rightarrow$” folgt dabei aus der Stetigkeit von g, und die Richtung “$\Leftarrow$” folgt so: Ist $g^{-1}[U]$ offen in $\mathbb{R}$, so ist $\mathbb{R} \setminus g^{-1}[U]$ abgeschlossen in $\mathbb{R}$. Also ist $K := [0, 2\pi] \setminus g^{-1}[U] = [0, 2\pi] \cap g^{-1}[S^1 \setminus U]$ abgeschlossen und beschränkt, also kompakt. Weil g stetig ist, ist auch $g[K]$ kompakt, also abgeschlossen. Weil $g[[0, 2\pi]] = S^1$, folgt $g[K] = g[[0, 2\pi] \cap g^{-1}[S^1 \setminus U]] = S^1 \setminus U$. Also ist U offen in S^1.

Nun sei $\mathbb{R}/2\pi\mathbb{Z}$ der Quotientenraum von $\mathbb{R}$ modulo der Relation $\sim$ auf $\mathbb{R}$ mit $x \sim y :\Leftrightarrow x - y \in 2\pi\mathbb{Z}$, versehen mit der Quotiententopologie. Es sei $f : \mathbb{R} \to \mathbb{R}/2\pi\mathbb{Z}$, $f(x) = x + 2\pi\mathbb{Z}$ die kanonische Abbildung und $h : \mathbb{R}/2\pi\mathbb{Z} \to S^1$, $h(x + 2\pi\mathbb{Z}) := g(x)$ die von g induzierte Abbildung. Offensichtlich ist h wohldefiniert, bijektiv, und es gilt $g = h \circ f$. Nach Übung 1.87 ist h stetig, da g stetig und f final ist. Umgekehrt folgt wegen $f = h^{-1} \circ g$ nach Übung 1.87 auch: h^{-1} ist stetig, da f stetig und g final ist. Also ist h ein Homöomorphismus im Sinne der folgenden Definition:

Definition 1.91 Eine bijektive, stetige Abbildung zwischen zwei topologischen Räumen, deren Umkehrung ebenfalls stetig ist, wird *Homöomorphismus* genannt. Zwei Räume heißen *homöomorph*, wenn es einen Homöomorphismus zwischen ihnen gibt.

Grob gesagt ist “Homöomorphie” der topologische Isomorphiebegriff.

Lemma 1.92 (Finaltopologie unter Inklusionen) Es sei $(X, \mathcal{T})$ ein topologischer Raum und $(U_i)_{i\in I}$ eine Überdeckung von X, also $\bigcup_{i\in I} U_i = X$. Wir versehen jedes U_i mit der Teilraumtopologie $\mathcal{T}_i$. Dann ist die Familie der Inklusionsabbildungen $(\iota_i : (U_i, \mathcal{T}_i) \to (X, \mathcal{T}))_{i\in I}$, $\iota_i(x) = x$, final, wenn mindestens eine der folgenden drei Bedingungen erfüllt ist:

1. Alle U_i, $i \in I$, sind offen;
2. I ist endlich und alle U_i, $i \in I$, sind abgeschlossen.
3. Alle U_i, $i \in I$, sind abgeschlossen und $(U_i)_{i\in I}$ ist lokal endlich im folgenden Sinn: Für alle $x \in X$ gibt es eine offene Umgebung $V \in \mathcal{T}$ von x, für die $\{i \in I |\ U_i \cap V \neq \emptyset\}$ endlich ist.

Beweis: Weil die ι_i stetig sind, ist nur noch unter den Voraussetzungen 1., 2. oder 3. zu zeigen: Für alle $U \subseteq X$, für die $U \cap U_i \in \mathcal{T}_i$ für alle $i \in I$ gilt, folgt $U \in \mathcal{T}$. Es sei so ein U gegeben.

- Es gelte die Voraussetzung 1. Wir können für jedes $i \in I$ wegen $U \cap U_i \in \mathcal{T}_i$ nach Definition der Teilraumtopologie ein $V_i \in \mathcal{T}$ mit $V_i \cap U_i = U \cap U_i$ wählen. Dann wissen wir $V_i \cap U_i \in \mathcal{T}$ wegen $U_i, V_i \in \mathcal{T}$, und daher
$$U = U \cap \bigcup_{i\in I} U_i = \bigcup_{i\in I} (U \cap U_i) = \bigcup_{i\in I} (V_i \cap U_i) \in \mathcal{T}$$
- Voraussetzung 2. impliziert Voraussetzung 3. und wird deshalb dort mit behandelt.
- Es gelte die Voraussetzung 3., und es sei $x \in U$ gegeben. Es genügt zu zeigen, dass U eine Umgebung von x bzgl. $\mathcal{T}$ ist. Nach Voraussetzung 3. wählen wir eine offene Umgebung $W \in \mathcal{T}$ von x, so dass $E = \{i \in I |\ U_i \cap W \neq \emptyset\}$ endlich ist. Wir betrachten das Komplement $U^c = X \setminus U$ von U. Für jedes $i \in I$ gilt: Weil $U \cap U_i$ in U_i offen ist, ist $U^c \cap U_i$ in U_i abgeschlossen, also auch abgeschlossen in X, da U_i in X abgeschlossen ist. Also ist $W \cap U^c \cap U_i$ abgeschlossen in W, und daher die endliche Vereinigung $\bigcup_{i\in E} W \cap U^c \cap U_i$ abgeschlossen in W. Nach Definition von E gilt jedoch
$$\bigcup_{i\in E} W \cap U^c \cap U_i = \bigcup_{i\in I} W \cap U^c \cap U_i = W \cap U^c \cap \bigcup_{i\in I} U_i = W \cap U^c \cap X = W \cap U^c.$$
Damit ist gezeigt: $W \cap U^c$ ist abgeschlossen in W. Folglich ist $W \cap U$ offen in W, also offen in X, da W offen in X ist. Die Menge U ist also eine Umgebung von x in X.

$\square$

Übung 1.93 (Quotientenbildung in $\mathbb{C} \cup \{\infty\}$) Zeigen Sie, dass die Quotientenabbildung $q : \mathbb{C} \times (\mathbb{C} \setminus \{0\}) \to \mathbb{C}$, $q(x, y) = x/y$ eine Fortsetzung zu einer stetigen Abbildung $Q : (\mathbb{C} \cup \{\infty\})^2 \setminus \{(0, 0), (\infty, \infty)\} \to \mathbb{C} \cup \{\infty\}$, jedoch keine Fortsetzung zu einer in $(0, 0)$ oder in (∞, ∞) stetigen Funktion mit Werten in $\mathbb{C} \cup \{\infty\}$ besitzt. Dabei werden Teilmengen von $(\mathbb{C} \cup \{\infty\})^2$ mit der Teilraumtopologie zur Produkttopologie der Standardtopologie auf $\mathbb{C} \cup \{\infty\}$ versehen.

1.7 Cauchyfolgen und Vollständigkeit

Cauchyfolgen über halbmetrischen Räumen sind eine naheliegende Verallgemeinerung von Cauchyfolgen über $\mathbb{R}$ oder $\mathbb{C}$, wie Sie sie aus der Analysis 1 kennen:

Definition 1.94 (Cauchyfolgen) Es sei (M, d) ein halbmetrischer Raum.

1. Eine Folge $(a_n)_{n\in\mathbb{N}}$ mit Werten in M heißt *Cauchyfolge*, bzgl. d, wenn gilt:

$$\forall \epsilon > 0\ \exists n \in \mathbb{N}\ \forall k \geq n\ \forall l \geq n :\ d(a_k, a_l) < \epsilon.$$

2. Der Raum (M, d) wird *vollständig* genannt, wenn jede Cauchyfolge $(a_n)_{n\in\mathbb{N}}$ mit Werten in M in (M, d) konvergiert.

Eine Teilmenge $N \subseteq M$ heißt vollständig, wenn sie bezüglich der Einschränkung von d auf $N \times N$ vollständig ist. Anders gesagt: N ist vollständig, wenn jede Cauchyfolge mit Werten in N einen Grenzwert in N besitzt.

Übung 1.95 Zeigen Sie, dass jede konvergente Folge in einem halbmetrischen Raum eine Cauchyfolge ist.

Übung 1.96 Zeigen Sie, dass eine Folge $(a_n)_{n\in\mathbb{N}}$ mit Werten in M genau dann eine Cauchyfolge ist, wenn gilt:

$$\limsup_{n\to\infty} \sup_{m:\, m\geq n} d(a_n, a_m) = 0.$$

Bemerkung 1.97 Im Gegensatz zur Konvergenz ist die Eigenschaft, Cauchyfolge zu sein, eine *metrische* Eigenschaft, keine topologische Eigenschaft. Zum Beispiel erzeugen die Standardmetrik d auf $\mathbb{R}$ und die Metrik $d' : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $d'(x, y) = |\arctan x - \arctan y|$, zwar die gleiche Topologie, doch $(n)_{n\in\mathbb{N}}$ ist zwar eine Cauchyfolge bezüglich d', nicht jedoch bezüglich d.

Übung 1.98 (Vererbung der Cauchyfolgeneigenschaft) Zeigen Sie: Sind (M, d) und (N, d') zwei halbmetrische Räume, $f : M \to N$ eine gleichmäßig stetige Abbildung und $(a_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in (M, d), so ist $(f(a_n))_{n\in\mathbb{N}}$ eine Cauchyfolge in (N, d').

Übung 1.99 (Cauchyfolgen mit konvergenten Teilfolgen) Es sei $(a_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in einem halbmetrischen Raum (M, d), die eine gegen ein $x \in M$ konvergente Teilfolge besitzt. Zeigen Sie, dass auch $(a_n)_{n\in\mathbb{N}}$ gegen x konvergiert.

Aus der Analysis 1 wissen Sie, dass $\mathbb{R}$ und $\mathbb{C}$, versehen mit der Standardmetrik, vollständig sind. Vollständigkeit vererbt sich auf Produkthalbmetriken:

Lemma 1.100 (Vererbung der Vollständigkeit auf Produktmetriken) *Sind* (M_i, d_i), $i = 1, \ldots, n$, *vollständige halbmetrische Räume,* $M = M_1 \times \ldots \times M_n$ *und* $d : M \times M \to \mathbb{R}$ *die Produktmetrik dazu, so ist auch* (M, d) *vollständig.*

Beweis: Es sei $(a_k)_{k\in\mathbb{N}}$ eine Cauchyfolge mit Werten in M, $a_k = (a_{k,1}, \ldots, a_{k,n})$. Dann ist für alle $i = 1, \ldots, n$ die Folge $(a_{k,i})_{k\in\mathbb{N}}$ eine Cauchyfolge mit Werten in M_i. Dies folgt aus Übung 1.98: Bezeichnet nämlich $f_i : M \to M_i$ die i-te kanonische Projektion, $i = 1, \ldots, n$, so ist f_i gleichmäßig stetig, und es gilt $f_i(a_k) = a_{k,i}$ für alle $k \in \mathbb{N}$. Wegen der Vollständigkeit des Raums (M_i, d_i) konvergiert die Folge $(a_{k,i})_{k\in\mathbb{N}}$ für $k \to \infty$ gegen ein $b_i \in M_i$. Insbesondere gilt $d(b_i, a_{k,i}) \to 0$ für $k \to \infty$. Setzen wir $b := (b_1, \ldots, b_n)$, so folgt

$$d(b, a_k) = \max_{i=1,\ldots,n} d(b_i, a_{k,i}) \overset{k\to\infty}{\longrightarrow} 0.$$

Also konvergiert die Folge $(a_k)_{k\in\mathbb{N}}$ gegen b.

□

Beispiel 1.101 (Vollständigkeit von $\mathbb{R}^n$) *Die Räume* $\mathbb{K}^n$ *mit* $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ *und* $n \in \mathbb{N}$ *sind bezüglich der von der Norm* $\|\cdot\|_\infty$ *erzeugten Metrik vollständig. Dies folgt mit der Analysis 1 bekannten Vollständigkeit von* $\mathbb{K}$ *aus dem vorhergehenden Lemma als der Spezialfall* $M_i = \mathbb{K}$, $i = 1, \ldots, n$.

Wir zeigen später (Satz 1.169), dass *alle* Normen auf $\mathbb{R}^n$ äquivalent im Sinne von Übung 1.64 sind. Es folgt dann, dass $\mathbb{R}^n$ vollständig bezüglich jeder Norm ist. (Das Gleiche gilt auch für $\mathbb{C}^n$).

Wir besprechen nun eine Variante von Lemma 1.100 für kartesische Potenzen. Erinnern Sie sich dazu an den Beschränktheitsbegriff:

Definition 1.102 (Beschränktheit)

1. Eine Teilmenge $N \subseteq M$ eines halbmetrischen Raums (M, d) wird *beschränkt* genannt, wenn es ein $r \in \mathbb{R}^+$ gibt, so dass für alle $x, y \in N$ gilt: $d(x, y) \leq r$.

2. Es sei I eine Menge. Eine Abbildung $f : I \to M$ heißt beschränkt, wenn ihr Wertebereich $f[I]$ beschränkt ist. Mit M_b^I bezeichnen wir die Menge aller beschränkten Abbildungen $f : I \to M$.

Für $N \neq \emptyset$ kann man die Beschränktheit auch äquivalent durch $\exists x \in N \; \exists r > 0 : N \subseteq U_r^d(x)$ charakterisieren.

Übung 1.103 (sup-Metrik) *Zeigen Sie: Ist* $I \neq \emptyset$ *eine Indexmenge und* (M, d) *ein halbmetrischer Raum, so wird durch* $d_\infty : M_b^I \times M_b^I \to \mathbb{R}$, $d_\infty(f, g) = \sup_{i\in I} d(f_i, g_i)$ *eine Halbmetrik auf* M_b^I *gegeben. Ist* d *sogar eine Metrik, so ist auch* d_∞ *eine Metrik.*

Lemma 1.104 (Vollständigkeit beschränkter kartesischer Potenzen) *Ist (M, d) ein vollständiger halbmetrischer Raum und I eine Menge, so ist (M_b^I, d_∞) ebenfalls vollständig.*

Beweis: Es sei $(f_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in (M_b^I, d_∞). Es sei $j \in I$ gegeben. Die Auswertungsabbildung $\delta_j : M_b^I \to M$, $\delta_j(f) := f(j)$ ist wegen

$$\forall f, g \in M_b^I : \ d(\delta_j(f), \delta_j(g)) = d(f(j), g(j)) \leq d_\infty(f, g)$$

gleichmäßig stetig. Aus der Übung 1.98 folgt damit, dass $(f_n(j))_{n\in\mathbb{N}}$ eine Cauchyfolge in M ist. Weil (M, d) vollständig ist, konvergiert diese Folge gegen ein $a(j) \in M$. Wählen wir für jedes $j \in I$ ein solches $a(j)$ aus, so erhalten wir ein $a \in M^I$. Gegeben $\epsilon > 0$, nehmen wir mit der Cauchyfolgeneigenschaft von $(f_n)_{n\in\mathbb{N}}$ ein $n(\epsilon) \in \mathbb{N}$, so dass für alle $k, l \in \mathbb{N}$ mit $k, l \geq n(\epsilon)$ gilt: $d_\infty(f_k, f_l) \leq \epsilon$, also $d(f_k(j), f_l(j)) \leq \epsilon$ für alle $j \in I$. Gegeben $k \geq n(\epsilon)$, erhalten wir wegen $f_l(j) \xrightarrow{l\to\infty}_{\mathcal{T}_d} a(j)$ und der Stetigkeit der Halbmetrik die Aussage $d(f_k(j), a(j)) = \lim_{l\to\infty} d(f_k(j), f_l(j)) \leq \epsilon$ für alle $j \in I$. Hieraus folgt einerseits die Beschränktheit von a, also $a \in M_b^I$, denn

$$\begin{aligned}
&\sup_{i,j\in I} d(a(i), a(j)) \leq \sup_{i,j\in I} d(a(i), f_{n(\epsilon)}(i)) + d(f_{n(\epsilon)}(i), f_{n(\epsilon)}(j)) + d(f_{n(\epsilon)}(j), a(j)) \\
&\leq \epsilon + \sup_{i,j\in I} d(f_{n(\epsilon)}(i), f_{n(\epsilon)}(j)) + \epsilon < \infty
\end{aligned}$$

wegen $f_{n(\epsilon)} \in M_b^I$. Andererseits folgt $d_\infty(f_k, a) \leq \epsilon$ für $k \geq n(\epsilon)$, also $f_n \xrightarrow{n\to\infty}_{\mathcal{T}_{d_\infty}} a$. Damit haben wir gezeigt: Jede Cauchyfolge $(f_n)_{n\in N}$ in (M_b^I, d_∞) konvergiert in (M_b^I, d_∞).

Beispiel 1.105 (Vollständigkeit von ℓ^∞) *Aus dem Lemma folgt: Ist I eine abzählbare Indexmenge und $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, so ist $(\ell^\infty(I, \mathbb{K}), \|\cdot\|_\infty)$ vollständig.*[5] *Es ist nämlich $\ell^\infty(I, \mathbb{K}) = \mathbb{K}_b^I$ und d_∞ die von $\|\cdot\|_\infty$ induzierte Metrik.*

Für viele Anwendungen ist der Raum M_b^I zu groß. Allerdings vererbt sich die Vollständigkeit auf abgeschlossene Teilräume, wie das folgende Lemma zeigt:

Lemma 1.106 (Vollständigkeit und Abgeschlossenheit) Jede *abgeschlossene* Teilmenge $N \subseteq M$ eines vollständigen halbmetrischen Raums (M, d) ist vollständig. Ist d sogar eine Metrik, so gilt auch umgekehrt: Jede vollständige Teilmenge $N \subseteq M$ ist abgeschlossen.

Beweis: Ist (M, d) vollständig, $N \subseteq M$ abgeschlossen und $(a_n)_{n\in\mathbb{N}}$ eine Cauchyfolge mit Werten in N, so konvergiert die Folge gegen ein $x \in M$. Nach Lemma 1.49 ist x ein Berührpunkt von N, also $x \in N$ wegen der Abgeschlossenheit von N. Die Menge N ist also vollständig.

[5] Wir nennen einen halbnormierten Raum vollständig, wenn der davon induzierte halbmetrische Raum vollständig ist.

Es sei umgekehrt N vollständig und d eine Metrik, sowie $x \in \overline{N}$ ein Berührpunkt von N. Nach Lemma 1.49 gibt es eine Folge $(a_n)_{n\in\mathbb{N}}$ mit Werten in N, die gegen x konvergiert. Insbesondere ist diese Folge eine Cauchyfolge. Da N vollständig ist, konvergiert sie auch gegen ein $y \in N$. Weil d eine Metrik ist, sind Grenzwerte von Folgen eindeutig, so dass $x = y$ folgt. Es folgt $x \in N$. Also ist N abgeschlossen.

□

Wir nehmen nun an, dass die Indexmenge I selbst mit einer Topologie $\mathcal{T}$ versehen ist. Wir untersuchen also einen topologischen Raum $(I, \mathcal{T})$ und einen halbmetrischen Raum (M, d). Es bezeichne $C_b((I, \mathcal{T}), (M, d))$ (oder kurz $C_b(I, M)$) die Menge aller *stetigen, beschränkten*[6] Abbildungen $f : I \to M$.

Satz 1.107 (Vollständigkeit des Raums stetiger beschränkter Funktionen in der sup-Metrik) *Ist $(I, \mathcal{T})$ ein topologischer Raum und (M, d) ein halbmetrischer Raum, so ist der Teilraum $C_b((I, \mathcal{T}), (M, d))$ von (M_b^I, d_∞) abgeschlossen. Ist (M, d) vollständig, so ist auch $C_b((I, \mathcal{T}), (M, d))$ bezüglich d_∞ vollständig.*

Beweis: Es sei $(f_n)_{n\in\mathbb{N}}$ eine Folge mit Werten in $C_b((I, \mathcal{T}), (M, d))$, die bzgl. d_∞ gegen ein $a \in M_b^I$ konvergiert. Zu zeigen ist die Stetigkeit von a, also $a \in C_b((I, \mathcal{T}), (M, d))$. Hierzu seien $j \in I$ und $\epsilon > 0$ gegeben. Wir müssen eine Umgebung U von j in $(I, \mathcal{T})$ finden, so dass für alle $i \in U$ gilt: $d(a(j), a(i)) < \epsilon$. Weil $(f_n)_{n\in\mathbb{N}}$ gegen a konvergiert, finden wir ein $n \in \mathbb{N}$ mit $d_\infty(f_n, a) < \epsilon/3$. Insbesondere gilt $d(f_n(i), a(i)) < \epsilon/3$ für alle $i \in I$. Weil $f_n : I \to M$ stetig ist, finden wir eine Umgebung U von j, so dass für alle $i \in U$ gilt: $d(f_n(j), f_n(i)) < \epsilon/3$. Es folgt für diese i:

$$d(a(j), a(i)) \leq d(a(j), f_n(j)) + d(f_n(j), f_n(i)) + d(f_n(i), a(i)) < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

Damit ist die Abgeschlossenheit von $C_b((I, \mathcal{T}), (M, d))$ gezeigt. Ist nun (M, d) sogar vollständig, so folgt die Vollständigkeit von $C_b((I, \mathcal{T}), (M, d))$ bezüglich d_∞ hieraus mit Lemma 1.104 und Lemma 1.106.

□

Beispiel 1.108 *Der Raum $C_b(\mathbb{R}, \mathbb{R})$ aller beschränkten, stetigen Funktionen $f : \mathbb{R} \to \mathbb{R}$ ist bezüglich der Supremumsmetrik $d_\infty(f, g) = \sup_{x\in\mathbb{R}} |f(x) - g(x)|$ vollständig.*

Übung 1.109 (Unvollständigkeit von $C([a, b])$ bezüglich $\|\cdot\|_p$) Gegeben sei $1 < p < \infty$. Zeigen Sie, dass der Raum $(C([-1, 1], \mathbb{R}), \|\cdot\|_p)$ aus Übung 1.21 nicht vollständig ist. Betrachten Sie dazu die Folge $(f_n)_{n\in\mathbb{N}}$ mit $f_n : [-1, 1] \to \mathbb{R}$, $f_n(x) = \max\{0, \min\{1, nx\}\}$.

Satz 1.110 (ℓ^p ist vollständig) *Es sei I eine abzählbare Indexmenge, $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, und $1 \leq p \leq \infty$ Dann ist $\ell^p(I, \mathbb{K})$ bezüglich $\|\cdot\|_p$ vollständig.*

[6] "C" für "continuous", "b" für "beschränkt" oder "bounded"

Beweis: Für $p = \infty$ ist uns das schon aus Beispiel 1.105 bekannt. Wir dürfen also $p < \infty$ annehmen. Es sei also eine Cauchyfolge $(f_n)_{n\in\mathbb{N}}$ in $\ell^p(I)$ gegeben. Für jedes $j \in I$ gilt: Weil die Auswertungsabbildung $\delta_j : \ell^p(I) \to \mathbb{R}$, $f \mapsto f(j)$, nach Übung 1.73 gleichmäßig stetig ist, ist die Folge $(f_n(j))_{n\in\mathbb{N}}$ nach Übung 1.98 eine Cauchyfolge in $\mathbb{K}$, also konvergent gegen ein $a(j) \in \mathbb{K}$. Wir setzen $a = (a(j))_{j\in I}$. Zu zeigen ist nun $a \in \ell^p(I)$ und $\|f_n - a\|_p \to 0$ für $n \to \infty$. Hierzu verwenden wir das aus den Übungen zur Analysis 1 bekannte

Lemma von Fatou für Reihen: *Für alle* $(x_{n,j})_{n\in\mathbb{N},j\in I} \in [0,\infty]^{\mathbb{N}\times I}$ *gilt*

$$\sum_{j\in I} \liminf_{n\to\infty} x_{n,j} \le \liminf_{n\to\infty} \sum_{j\in I} x_{n,j}.$$

(Im Falle konvergenter Folgen kann man natürlich "lim" statt "lim inf" schreiben.) Es folgt damit einerseits

$$\|a\|_p^p = \sum_{j\in I} |a(j)|^p = \sum_{j\in I} \lim_{n\to\infty} |f_n(j)|^p \le \liminf_{n\to\infty} \sum_{j\in I} |f_n(j)|^p = \liminf_{n\to\infty} \|f_n\|_p^p. \tag{11}$$

Weil die Normabbildung $\|\cdot\|_p : (\ell^p(I), \|\cdot\|_p) \to (\mathbb{R}, |\cdot|)$ nach Lemma 1.59 gleichmäßig stetig ist, wird die Cauchyfolge $(f_n)_{n\in\mathbb{N}}$ nach Übung 1.98 auf eine *Cauchyfolge* $(\|f_n\|_p)_{n\in\mathbb{N}}$ in $\mathbb{R}$ abgebildet. Diese ist insbesondere in $\mathbb{R}$ konvergent. Das bedeutet $\liminf_{n\to\infty} \|f_n\|_p = \lim_{n\to\infty} \|f_n\|_p < \infty$, und wir erhalten zusammen mit (11) die Abschätzung $\|a\|_p < \infty$, also $a \in \ell^p(I)$.
Andererseits folgt aus dem Lemma von Fatou auch für alle $n \in \mathbb{N}$:

$$\begin{aligned}\|f_n - a\|_p^p &= \sum_{j\in I} |f_n(j) - a(j)|^p = \sum_{j\in I} |f_n(j) - \lim_{m\to\infty} f_m(j)|^p = \sum_{j\in I} \lim_{m\to\infty} |f_n(j) - f_m(j)|^p \\ &\le \liminf_{m\to\infty} \sum_{j\in I} |f_n(j) - f_m(j)|^p = \liminf_{m\to\infty} \|f_n - f_m\|_p^p \overset{n\to\infty}{\longrightarrow} 0,\end{aligned}$$

wobei wir im letzten Schritt nochmal verwendet haben, dass $(f_n)_{n\in\mathbb{N}}$ eine Cauchyfolge ist. Die Folge $(f_n)_{n\in\mathbb{N}}$ konvergiert also bezüglich $\|\cdot\|_p$ gegen $a \in \ell^p(I)$.

□

1.8 Vervollständigung

Definition 1.111 Es sei (M, d) ein halbmetrischer Raum. Ein Tripel $(\hat{M}, \hat{d}, i)$, bestehend aus einer Menge $\hat{M}$, einer Metrik $\hat{d}$ auf $\hat{M}$ und einer Abbildung $i : M \to \hat{M}$, wird *Vervollständigung* von (M, d) genannt, wenn gilt:

1. $i : (M, d) \to (\hat{M}, \hat{d})$ ist eine Isometrie.
2. Der Wertebereich $i[M]$ ist dicht in $(\hat{M}, \hat{d})$.
3. Der metrische Raum $(\hat{M}, \hat{d})$ ist vollständig.

Beispiel 1.112 ($\mathbb{R}$ als Vervollständigung von $\mathbb{Q}$) *Das Tripel $(\mathbb{R}, d_\mathbb{R}, i)$ mit der Standardmetrik $d_\mathbb{R}$ auf $\mathbb{R}$ und der Inklusionsabbildung $i : \mathbb{Q} \to \mathbb{R}$, $i(r) = r$, ist eine Vervollständigung des Raums $(\mathbb{Q}, d_\mathbb{Q})$ der rationalen Zahlen, versehen mit dem Standardabstand $d_\mathbb{Q}(r, s) = |r - s|$ für $r, s \in \mathbb{Q}$.*

Übung 1.113 (Dichtheit von Quadern mit dichten Seiten) Es seien (M_k, d_k), $k = 1, \ldots, n$ mit $n \in \mathbb{N}$, halbmetrische Räume mit dem Produktraum $M = M_1 \times \ldots \times M_n$, versehen mit der Produktmetrik d. Weiter sei $A_k \subseteq M_k$ für $k = 1, \ldots, n$. Zeigen Sie:

$$\overline{A_1} \times \ldots \times \overline{A_n} = \overline{A_1 \times \ldots \times A_n}.$$

Folgern Sie: Ist für alle $k = 1, \ldots, n$ die Menge A_k dicht in (M_k, d_k), so ist auch $A_1 \times \ldots \times A_n$ dicht in M.

Übung 1.114 (Vervollständigung von endlichen kartesischen Produkten) Es seien (M_k, d_k), $k = 1, \ldots, n$ mit $n \in \mathbb{N}$, halbmetrische Räume mit dem Produktraum $M = M_1 \times \ldots \times M_n$, versehen mit der Produktmetrik d. Für jedes $k = 1, \ldots, n$ sei $(\hat{M}_k, \hat{d}_k, i_k)$ eine Vervollständigung von M_k. Es seien $\hat{M} := \hat{M}_1 \times \ldots \times \hat{M}_n$ und $\hat{d}$ die Produktmetrik auf $\hat{M}$ zu $d_1, \ldots, d_n$, sowie $i : M \to \hat{M}$, $i(x_1, \ldots, x_n) = (i_1(x_1), \ldots, i_n(x_n))$. Zeigen Sie, dass $(\hat{M}, \hat{d}, i)$ eine Vervollständigung von (M, d) ist.

Satz 1.115 (Fortsetzung gleichmäßig stetiger Abbildungen) *Es seien (M, d_M) ein halbmetrischer Raum, $i : (M, d_M) \to (\hat{M}, \hat{d}_M)$ eine Isometrie in einen metrischen Raum $(\hat{M}, \hat{d}_M)$ mit dichtem Wertebereich $i[M]$, $(\hat{N}, \hat{d}_N)$ ein vollständiger metrischer Raum und $f : M \to \hat{N}$ eine Abbildung.*

1. *Ist $f : (M, d_M) \to (\hat{N}, \hat{d}_N)$ eine gleichmäßig stetige Abbildung, so gibt es eine gleichmäßig stetige Abbildung $F : (\hat{M}, \hat{d}) \to (\hat{N}, \hat{d}_N)$ mit $f = F \circ i$. Sie ist die einzige stetige Abbildung $F : (\hat{M}, \hat{d}) \to (\hat{N}, \hat{d}_N)$ mit $f = F \circ i$.*

2. *Sind die Einschränkungen von f auf beliebige beschränkte Mengen $A \subseteq M$ gleichmäßig stetig, so gibt es eine eindeutige stetige Abbildung $F : (\hat{M}, \hat{d}) \to (\hat{N}, \hat{d}_N)$ mit $f = F \circ i$. Eingeschränkt auf beliebige beschränkte Teilmengen von $\hat{M}$ ist sie sogar gleichmäßig stetig.*

Insbesondere ist dieser Satz anwendbar, wenn $(\hat{M}, \hat{d}_M, i)$ eine Vervollständigung von (M, d) ist.

Beweis: 1. Wir definieren zunächst $F : \hat{M} \to \hat{N}$ unter der Annahme, dass f gleichmäßig stetig ist. Hierzu sei $y \in \hat{M}$. Weil $i[M]$ dicht in $\hat{M}$ ist, können wir eine Folge $(x_n)_{n\in\mathbb{N}}$ in M mit $\lim^{\hat{d}_M}_{n\to\infty} i(x_n) = y$ wählen. Insbesondere ist $(i(x_n))_{n\in\mathbb{N}}$ eine Cauchyfolge in $(\hat{M}, \hat{d}_M)$ und daher $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge in (M, d_M), da i eine Isometrie ist. Die gleichmäßige Stetigkeit von $f : (M, d_M) \to (\hat{N}, \hat{d}_N)$ und Übung (1.98) implizieren dann, dass $(f(x_n))_{n\in\mathbb{N}}$ eine Cauchyfolge in $(\hat{N}, \hat{d}_N)$ ist. Weil $(\hat{N}, \hat{d}_N)$ vollständig ist, konvergiert sie gegen ein $z \in \hat{N}$. Wir setzen $F(y) := z$, müssen aber noch zeigen, dass dies wohldefiniert ist: Ist $(x'_n)_{n\in\mathbb{N}}$ eine weitere Folge in M mit $\lim^{\hat{d}_M}_{n\to\infty} i(x'_n) = y$ und $z' = \lim^{\hat{d}_N}_{n\to\infty} f(x'_n)$, so folgt

$d_M(x_n, x'_n) = \hat{d}_M(i(x_n), i(x'_n)) \to 0$ für $n \to \infty$, also wegen der Stetigkeit der Metrik $\hat{d}_N$ und der gleichmäßigen Stetigkeit von f:

$$\hat{d}_N(z, z') = \lim_{n\to\infty} \hat{d}_N(f(x_n), f(x'_n)) = 0.$$

Da $\hat{d}_N$ eine Metrik ist, folgt $z = z'$ und damit die Wohldefiniertheit von F.
Nun zeigen wir, dass das die so definierte Abbildung $F : \hat{M} \to \hat{N}$ gleichmäßig stetig ist. Wegen der gleichmäßigen Stetigkeit von f gibt es eine Abbildung $\delta : \mathbb{R}^+ \to \mathbb{R}^+$, so dass gilt:

$$\forall \epsilon > 0\ \forall x, y \in M :\ d_M(x, y) < \delta(\epsilon) \Rightarrow \hat{d}_N(f(x), f(y)) \leq \epsilon.$$

Wir zeigen, dass dann auch gilt:

$$\forall \epsilon > 0\ \forall x, y \in \hat{M} :\ \hat{d}_M(x, y) < \delta(\epsilon) \Rightarrow \hat{d}_N(F(x), F(y)) \leq \epsilon.$$

Hierzu seien $\epsilon > 0$ und $x, y \in \hat{M}$ mit $\hat{d}_M(x, y) < \delta(\epsilon)$ gegeben. Weil $i[M]$ dicht in $\hat{M}$ ist, können wir Folgen $(x_n)_{n\in\mathbb{N}_0}$ und $(y_n)_{n\in\mathbb{N}_0}$ mit Werten in M wählen, für die $i(x_n) \to x$ und $i(y_n) \to y$ für $n \to \infty$ bzgl. $\hat{d}_M$ gilt. Weil i isometrisch ist, folgt mit Stetigkeit der Metrik $\hat{d}_M$:

$$\lim_{n\to\infty} d_M(x_n, y_n) = \lim_{n\to\infty} \hat{d}_M(i(x_n), i(y_n)) = \hat{d}_M(x, y) < \delta(\epsilon)$$

also $d_M(x_n, y_n) < \delta(\epsilon)$ für alle genügend großen $n \in \mathbb{N}_0$. Für diese n folgt

$$\hat{d}_N(f(x_n), f(y_n)) \leq \epsilon,$$

also

$$\hat{d}_N(F(x), F(y)) = \lim_{n\to\infty} \hat{d}_N(f(x_n), f(y_n)) \leq \epsilon,$$

wobei wir $f(x_n) \to F(x)$, $f(y_n) \to F(y)$ und die Stetigkeit der Metrik d_N verwendet haben.
Nun zeigen wir $f = F \circ i$. Hierzu sei $x \in M$ gegeben. Wir wählen die konstante Folge $(x_n = x)_{n\in\mathbb{N}_0}$, für die offensichtlich $i(x_n) \to i(x)$ in $(\hat{M}, \hat{d}_M)$ gilt. Nach Definition von F folgt $F(i(x)) = \lim_{n\to\infty} f(x_n) = f(x)$, also $f = F \circ i$, wie behauptet.
Als nächstes zeigen wir: Ist $G : (\hat{M}, \hat{d}_M) \to (\hat{N}, \hat{d}_N)$ eine weitere stetige Abbildung mit $G \circ i = f$, so gilt $G = F$. Hierzu sei $x \in \hat{M}$ gegeben. Wir wählen wieder eine Folge $(x_n)_{n\in\mathbb{N}_0}$ in M, für die $i(x_n)$ für $n \to \infty$ in $(\hat{M}, \hat{d}_M)$ gegen x konvergiert. Wegen der Stetigkeit von G konvergiert dann $f(x_n) = G(i(x_n))$ für $n \to \infty$ in $(\hat{N}, \hat{d}_N)$ gegen $G(x)$. Weil $f(x_n)$ nach Definition von F gegen $F(x)$ konvergiert, folgt $G(x) = F(x)$ wegen der Eindeutigkeit des Grenzwerts in metrischen Räumen. Das zeigt $G = F$. Man beachte, dass dieser Nachweis von $F = G$ nur die Stetigkeit von F und G, nicht jedoch gleichmäßige Stetigkeit benötigt.

2. Nun sei $f : (M, d_M) \to (\hat{N}, \hat{d}_N)$ eine Abbildung, deren Einschränkungen auf beliebige beschränkte Mengen $A \subseteq M$ gleichmäßig stetig sind. Der Fall $M = \emptyset$ ist trivial; deshalb nehmen wir $M \neq \emptyset$ an.
Zu jeder beschränkten Menge $B \subseteq \hat{M}$ gibt es eine beschränkte Menge $A \subseteq M$, so dass $B \subseteq \overline{i[A]}$ gilt. In der Tat: Gegeben solch ein B und ein $x \in M$, finden wir ein $r > 0$

mit $B \subseteq U_r^{\hat{d}_M}(i(x))$. Setzen wir $A := U_r^{d_M}(x)$. Insbesondere ist A in (M, d_M) beschränkt. Nun sei $b \in B$ gegeben. Zu zeigen ist nun $b \in \overline{i[A]}$. Weil $i[M]$ dicht in $\hat{M}$ ist, finden wir eine Folge $(b_n)_{n\in\mathbb{N}}$ in M mit $\lim_{n\to\infty}^{\hat{d}_M} i(b_n) = b$. Wegen $d_M(x, b_n) = \hat{d}_M(i(x), i(b_n)) \to \hat{d}_M(i(x), b) < r$ für $n \to \infty$ gilt $d_M(x, b_n) < r$, also $b_n \in A$ für alle genügend großen n. Wegen $\lim_{n\to\infty}^{\hat{d}_M} i(b_n) = b$ folgt $b \in \overline{i[A]}$ nach Lemma 1.49.
Gegeben $n \in \mathbb{N}$, setzen wir nun $B_n := U_n^{\hat{d}_M}(i(x))$ und $A_n := U_n^{d_M}(x)$, so folgt $B_n \subseteq \overline{i[A_n]}$ nach dem eben Gezeigten. Da f eingeschränkt auf A_n gleichmäßig stetig ist und $i[A_n]$ dicht in $\overline{i[A_n]}$ ist, gibt es eine eindeutig bestimmte gleichmäßig stetige Abbildung $F_n : \overline{i[A_n]} \to \hat{N}$ mit $F_n(i(x)) = f(x)$ für $x \in A_n$. Weil auch $F_{n+1}(i(x)) = f(x)$ für $x \in A_n$ gilt und auch F_{n+1} auf $\overline{i[A_n]} \subseteq \overline{i[A_{n+1}]}$ gleichmäßig stetig ist, folgt, dass $F_{n+1} : \overline{i[A_{n+1}]} \to \hat{N}$ eine Fortsetzung von $F_n : \overline{i[A_n]} \to \hat{N}$ ist.
Wir haben daher eine gemeinsame Fortsetzung $F : \hat{M} \to \hat{N}$ *aller* $F_n : \overline{i[A_n]} \to \hat{N}$, wobei wir $\hat{M} \supseteq \bigcup_{n\in\mathbb{N}} \overline{i[A_n]} \supseteq \bigcup_{n\in\mathbb{N}} B_n \supseteq \hat{M}$ verwenden. Insbesondere gilt $F \circ i(x) = f(x)$ für alle $x \in M$. F ist stetig, da es eingeschränkt auf jede der Mengen in der offenen Überdeckung $(B_n)_{n\in\mathbb{N}}$ stetig ist.
Die Eindeutigkeit von F wurde schon oben gezeigt.

□

Korollar 1.116 (Hochheben von Abbildungen auf die Vervollständigung) *Es seien (M, d_M) und (N, d_N) halbmetrische Räume mit Vervollständigungen $(\hat{M}, \hat{d}_M, i_M)$ und $(\hat{N}, \hat{d}_N, i_N)$. Weiter sei $f : M \to N$ eine Abbildung.*

1. *Ist $f : (M, d_M) \to (N, d_N)$ gleichmäßig stetig, so gibt es eine eindeutig bestimmte gleichmäßig stetige Abbildung $F : (\hat{M}, \hat{d}_M) \to (\hat{N}, \hat{d}_N)$ mit $F \circ i_M = i_N \circ f$.*

2. *Ist $f : (M, d_M) \to (N, d_N)$ eingeschränkt auf beliebige beschränkte Mengen gleichmäßig stetig, so gibt es eine eindeutig bestimmte stetige Abbildung $F : (\hat{M}, \hat{d}_M) \to (\hat{N}, \hat{d}_N)$ mit $F \circ i_M = i_N \circ f$. Die Abbildung F ist eingeschränkt auf beliebige beschränkte Mengen gleichmäßig stetig.*

Beweis: Dies folgt unmittelbar aus dem Satz 1.115, angewandt auf $i_N \circ f$.

□

Korollar 1.117 (Eindeutigkeit der Vervollständigung) *Sind $(\hat{M}_1, \hat{d}_1, i_1)$ und $(\hat{M}_2, \hat{d}_2, i_2)$ zwei Vervollständigungen des gleichen halbmetrischen Raums (M, d), so gibt es eine eindeutige isometrische Bijektion $j : (\hat{M}_1, \hat{d}_1) \to (\hat{M}_2, \hat{d}_2)$ mit $j \circ i_1 = i_2$.*

Beweis: Weil $i_2 : (M, d) \to (\hat{M}_2, \hat{d}_2)$ isometrisch, also auch gleichmäßig stetig ist, und weil $(\hat{M}_1, \hat{d}_1, i_1)$ eine Vervollständigung von (M, d) ist, gibt es nach dem Satz 1.115 eine eindeutig bestimmte stetige Abbildung $j : (\hat{M}_1, \hat{d}_1) \to (\hat{M}_2, \hat{d}_2)$ mit $j \circ i_1 = i_2$. Ebenso – mit vertauschten Rollen von “1” und “2” – gibt es eine eindeutig bestimmte stetige Abbildung $k : (\hat{M}_2, \hat{d}_2) \to (\hat{M}_1, \hat{d}_1)$ mit $k \circ i_2 = i_1$. Es folgt $k \circ j \circ i_1 = k \circ i_2 = i_1 = \mathrm{id}_{\hat{M}_1} \circ i_1$. Es folgt $k \circ j = \mathrm{id}_{\hat{M}_1}$, weil $i_1[M]$ dicht im metrischen Raum $(\hat{M}_1, \hat{d}_1)$ ist. Mit vertauschten Rollen von “1” und “2” folgt ebenso $j \circ k = \mathrm{id}_{\hat{M}_2}$. Die Abbildungen j und k sind also invers zueinander, also Bijektionen. Da i_1 und i_2 isometrisch sind, folgt für alle $x, y \in M$:

$$\hat{d}_2(j(i_1(x)), j(i_1(y))) = \hat{d}_2(i_2(x), i_2(y)) = d(x, y) = \hat{d}_1(i_1(x), i_1(y)).$$

Die Einschränkung von j auf $i_1[M]$ ist also isometrisch.
Wir betrachten nun die Abbildung $f : \hat{M}_1 \times \hat{M}_1 \to \mathbb{R}$, $f(x, y) = \hat{d}_2(j(x), j(y))$. Da j und d_2 stetig sind, ist auch f stetig, und auf der dichten Teilmenge $i_1[M] \times i_1[M]$ von $\hat{M}_1 \times \hat{M}_1$ (siehe Übung 1.113) stimmt f mit $\hat{d}_1$ überein. Mit Übung 1.75 folgt $f = \hat{d}_1$. Also ist j eine Isometrie.

□

Wir beschäftigen uns jetzt mit der Existenz von Vervollständigungen.
Die Hauptarbeit der Konstruktion steckt in folgendem Lemma:

Lemma 1.118 (Raum der Cauchyfolgen) *Es sei (M, d) ein halbmetrischer Raum und CF_M die Menge aller Cauchyfolgen auf (M, d).*

1. *Durch*

$$d_{\mathrm{CF}} : \mathrm{CF}_M \times \mathrm{CF}_M \to \mathbb{R}, \quad d_{\mathrm{CF}}((x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}) := \lim_{n\to\infty} d(x_n, y_n)$$

   wird eine Halbmetrik auf CF_M definiert.

2. *Der halbmetrische Raum $(\mathrm{CF}_M, d_{\mathrm{CF}})$ ist vollständig.*

3. *Die Abbildung $i : (M, d) \to (\mathrm{CF}_M, d_{\mathrm{CF}})$, $i(x) = (x)_{n\in\mathbb{N}}$, die jedem $x \in M$ die konstante Cauchyfolge mit dem Wert x zuordnet, ist eine Isometrie.*

4. *Der Wertebereich $i[M]$ ist dicht in $(\mathrm{CF}_M, d_{\mathrm{CF}})$.*

Beweis:

1. Sind $x = (x_n)_{n\in\mathbb{N}}$ und $y = (y_n)_{n\in\mathbb{N}}$ Cauchyfolgen in (M, d), so ist $((x_n, y_n))_{n\in\mathbb{N}}$ eine Cauchyfolge im Raum $M \times M$ bezüglich der Produktmetrik, also $(d(x_n, y_n))_{n\in\mathbb{N}}$ eine Cauchyfolge in $\mathbb{R}$, da d gleichmäßig stetig ist. Also existiert $\lim_{n\to\infty} d(x_n, y_n)$, so dass d_{CF} wohldefiniert ist. Weiter gilt $d_{\mathrm{CF}}(x, y) = d_{\mathrm{CF}}(y, x) \geq 0$, weil $d(x_n, y_n) = d(y_n, x_n) \geq 0$ für alle $n \in \mathbb{N}$ gilt. Ist $z = (z_n)_{n\in\mathbb{N}}$ eine weitere Cauchyfolge in (M, d), so folgt die Dreiecksungleichung

$$d_{\mathrm{CF}}(x, z) = \lim_{n\to\infty} d(x_n, z_n) \leq \lim_{n\to\infty} d(x_n, y_n) + d(y_n, z_n) = d_{\mathrm{CF}}(x, y) + d_{\mathrm{CF}}(y, z).$$

   Damit ist gezeigt, dass $d_{\mathrm{CF}}(x, y)$ eine Halbmetrik ist.

2. Der folgende Beweis der Vollständigkeit von $(\mathrm{CF}_M, d_{\mathrm{CF}})$ ist der komplizierteste Teil des Lemmas:

   Es sei $(x_j)_{j\in\mathbb{N}}$ eine Cauchyfolge in $(\mathrm{CF}_M, d_{\mathrm{CF}})$. Wir schreiben $x_j = (x_{j,n})_{n\in\mathbb{N}}$ für jedes $j \in \mathbb{N}$. Zu zeigen ist nun, dass die Cauchyfolge $(x_j)_{j\in\mathbb{N}}$ in $(\mathrm{CF}_M, d_{\mathrm{CF}})$ gegen ein $y \in \mathrm{CF}_M$ konvergiert.

   Weil $(x_j)_{j\in\mathbb{N}}$ eine Cauchyfolge in $(\mathrm{CF}_M, d_{\mathrm{CF}})$ ist, können wir eine Funktion $J : \mathbb{R}^+ \to \mathbb{N}$ wählen, so dass

   $$\forall \epsilon > 0 \, \forall j, j' \geq J(\epsilon) : d_{\mathrm{CF}}(x_j, x_{j'}) < \epsilon. \tag{12}$$

   Insbesondere gilt

   $$\forall \epsilon > 0 \, \forall \delta, \delta' \in \,]0, \epsilon] : \; d_{\mathrm{CF}}(x_{J(\delta)}, x_{J(\delta')}) < \epsilon. \tag{13}$$

   In der Tat: Es seien $\epsilon > 0$ und $\delta, \delta' \in \,]0, \epsilon]$ gegeben. Dann gilt nach der Wahl von J im Fall $J(\delta) \leq J(\delta')$:

   $$d_{\mathrm{CF}}(x_{J(\delta)}, x_{J(\delta')}) < \delta \leq \epsilon$$

   und im Fall $J(\delta) > J(\delta')$:

   $$d_{\mathrm{CF}}(x_{J(\delta)}, x_{J(\delta')}) < \delta' \leq \epsilon.$$

   Da $x_j = (x_{j,n})_{n\in\mathbb{N}}$ für jedes $j \in \mathbb{N}$ eine Cauchyfolge in (M, d) ist, können wir eine Funktion $N : \mathbb{N} \times \mathbb{R}^+ \to \mathbb{N}$ wählen, so dass gilt:

   $$\forall j \in \mathbb{N} \, \forall \epsilon > 0 \, \forall n, n' \geq N(j, \epsilon) : \; d(x_{j,n}, x_{j,n'}) < \epsilon. \tag{14}$$

   Da für alle $j, j' \in \mathbb{N}$ gilt:

   $$\lim_{m\to\infty} d(x_{j,m}, x_{j',m}) = d_{\mathrm{CF}}(x_j, x_{j'}),$$

   können wir eine Funktion $K : \mathbb{N} \times \mathbb{N} \times \mathbb{R}^+ \to \mathbb{N}$ wählen, so dass gilt:

   $$\forall j, j' \in \mathbb{N} \, \forall \epsilon > 0 \, \forall m \geq K(j, j', \epsilon) : \; |d(x_{j,m}, x_{j',m}) - d_{\mathrm{CF}}(x_j, x_{j'})| < \epsilon. \tag{15}$$

   Weiter sei $(\epsilon_n)_{n\in\mathbb{N}}$ eine gegen 0 konvergente Folge positiver Zahlen, z.B. $\epsilon_n = 1/n$.

   Wir definieren nun die Folge $y = (y_n)_{n\in\mathbb{N}}$ mit Werten in M durch folgenden "*Diagonalfolgenansatz*":

   $$y_n := x_{J(\epsilon_n), N(J(\epsilon_n), \epsilon_n)}. \tag{16}$$

   Wir zeigen nun, dass y eine Cauchyfolge in (M, d) ist, d.h. dass gilt:

   $$\forall \epsilon > 0 \, \exists n_0 \in \mathbb{N} \, \forall n, n' \geq n_0 : \; d(y_n, y_{n'}) < \epsilon.$$

Hierzu sei $\epsilon > 0$ gegeben. Wir wählen $n_0 \in \mathbb{N}$ so groß, dass gilt:

$$\forall n \geq n_0 : \epsilon_n < \frac{\epsilon}{4}.$$

Nun seien $n, n' \geq n_0$ gegeben. Insbesondere gilt hierfür $\epsilon_n < \frac{\epsilon}{4}$ und $\epsilon_{n'} < \frac{\epsilon}{4}$. Wir kürzen ab: $j := J(\epsilon_n)$, $j' := J(\epsilon_{n'})$ und setzen

$$m := \max\{N(j, \epsilon_n), N(j', \epsilon_{n'}), K(j, j', \frac{\epsilon}{4})\}.$$

Dann gilt:

$$d(y_n, y_{n'}) \leq d(y_n, x_{j,m}) + d(x_{j,m}, x_{j',m}) + d(x_{j',m}, y_{n'}).$$

Wir schätzen die drei Summanden auf der rechten Seite einzeln ab: Für den ersten Summanden erhalten wir wegen (14) und der Wahl von m:

$$d(y_n, x_{j,m}) = d(x_{j,N(j,\epsilon_n)}, x_{j,m}) < \epsilon_n \leq \frac{\epsilon}{4}.$$

Das gleiche Argument liefert für den dritten Summanden

$$d(y_{n'}, x_{j',m}) = d(x_{j',N(j',\epsilon_{n'})}, x_{j',m}) \leq \epsilon_{n'} \leq \frac{\epsilon}{4}.$$

Den zweiten Summanden schätzen wir mit (15) unter Verwendung von $m \geq K(j, j', \frac{\epsilon}{4})$ ab:

$$d(x_{j,m}, x_{j',m}) \leq d_{\mathrm{CF}}(x_j, x_{j'}) + |d(x_{j,m}, x_{j',m}) - d_{\mathrm{CF}}(x_j, x_{j'})| < d_{\mathrm{CF}}(x_j, x_{j'}) + \frac{\epsilon}{4}$$

Weiter erhalten wir mit (13), angewandt auf $\epsilon/4$ statt ϵ und ϵ_n, $\epsilon_{n'}$ für δ, δ':

$$d_{\mathrm{CF}}(x_j, x_{j'}) = d_{\mathrm{CF}}(x_{J(\epsilon_n)}, x_{J(\epsilon_{n'})}) < \frac{\epsilon}{4}.$$

Zusammengefasst:

$$d(y_n, y_{n'}) < \frac{\epsilon}{4} + \frac{\epsilon}{4} + \frac{\epsilon}{4} + \frac{\epsilon}{4} = \epsilon.$$

Damit ist $y \in \mathrm{CF}_M$ gezeigt.

Wir zeigen nun, dass die Folge $(x_j)_{j \in \mathbb{N}}$ im Raum $(\mathrm{CF}_M, d_{\mathrm{CF}})$ gegen y konvergiert. Zu zeigen ist also

$$\forall \epsilon > 0 \, \exists j \in \mathbb{N} \, \forall l \geq j : d_{\mathrm{CF}}(x_l, y) < \epsilon.$$

Hierzu sei $\epsilon > 0$ gegeben. Wir setzen $j = J(\epsilon/6)$. Nun sei $l \in \mathbb{N}$ mit $l \geq j$ gegeben. Insbesondere gilt wegen (12):

$$d_{\mathrm{CF}}(x_l, x_j) < \frac{\epsilon}{6}.$$

Wir erhalten

$$d_{\mathrm{CF}}(x_l, y) \leq d_{\mathrm{CF}}(x_l, x_j) + d_{\mathrm{CF}}(x_j, y) < \frac{\epsilon}{6} + d_{\mathrm{CF}}(x_j, y). \tag{17}$$

Weil $\lim_{n\to\infty} d(x_{j,n}, y_n) = d_{\mathrm{CF}}(x_j, y)$, finden wir ein $n_0 \in \mathbb{N}$, so dass gilt:

$$\forall n \geq n_0 : \ |d(x_{j,n}, y_n) - d_{\mathrm{CF}}(x_j, y)| < \frac{\epsilon}{6}.$$

Wir wählen $n \in \mathbb{N}$ so groß, dass gilt:

(a) $n \geq n_0$,

(b) $n \geq N(j, \frac{\epsilon}{6})$,

(c) $\epsilon_n < \frac{\epsilon}{6}$.

Dann gilt

$$d_{\mathrm{CF}}(x_j, y) \leq |d(x_{j,n}, y_n) - d_{\mathrm{CF}}(x_j, y)| + d(x_{j,n}, y_n) < \frac{\epsilon}{6} + d(x_{j,n}, y_n),$$

also in (17) eingesetzt:

$$d_{\mathrm{CF}}(x_l, y) < \frac{\epsilon}{6} + \frac{\epsilon}{6} + d(x_{j,n}, y_n) = \frac{\epsilon}{3} + d(x_{j,n}, y_n).$$

Wir setzen $j' = J(\epsilon_n)$, $n' = N(j', \epsilon_n)$ und

$$m = \max\{N(j, \frac{\epsilon}{6}), n', K(j, j', \frac{\epsilon}{6})\}.$$

Insbesondere gilt $y_n = x_{j',n'}$ nach der Definition (16) von y. Wir erhalten

$$d(x_{j,n}, y_n) = d(x_{j,n}, x_{j',n'}) \leq d(x_{j,n}, x_{j,m}) + d(x_{j,m}, x_{j',m}) + d(x_{j',m}, x_{j',n'}).$$

Schätzen wir die drei Summanden auf der rechten Seite einzeln ab:

$$d(x_{j,n}, x_{j,m}) < \frac{\epsilon}{6}$$

gilt wegen (14) und $n, m \geq N(j, \frac{\epsilon}{6})$; siehe (b). Ebenso gilt

$$d(x_{j',m}, x_{j',n'}) < \epsilon_n < \frac{\epsilon}{6}$$

wegen $m, n' \geq N(j', \epsilon_n)$, (14) und (c). Schließlich gilt

$$d(x_{j,m}, x_{j',m}) \leq |d(x_{j,m}, x_{j',m}) - d_{\mathrm{CF}}(x_j, x_{j'})| + d_{\mathrm{CF}}(x_j, x_{j'}).$$

Wir schätzen auch hier die Summanden auf der rechten Seite einzeln ab: Aus (15), angewandt auf $\epsilon/6$ statt ϵ, folgt

$$|d(x_{j,m}, x_{j',m}) - d_{\mathrm{CF}}(x_j, x_{j'})| < \frac{\epsilon}{6}$$

wegen $m \geq K(j, j', \epsilon/6)$. Wegen (13), angewandt auf $\epsilon/6$ statt ϵ und $\epsilon/6, \epsilon_n \in]0, \epsilon/6]$ statt δ, δ', folgt

$$d_{\mathrm{CF}}(x_j, x_{j'}) = d_{\mathrm{CF}}(x_{J(\epsilon/6)}, x_{J(\epsilon_n)}) < \frac{\epsilon}{6}.$$

Fassen wir alle Abschätzungen zusammen:

$$d(x_{j,m}, x_{j',m}) < \frac{\epsilon}{6} + \frac{\epsilon}{6} = \frac{\epsilon}{3},$$
$$d(x_{j,n}, y_n) \leq d(x_{j,n}, x_{j,m}) + d(x_{j,m}, x_{j',m}) + d(x_{j',m}, x_{j',n'}) < \frac{\epsilon}{6} + \frac{\epsilon}{3} + \frac{\epsilon}{6} = \frac{2}{3}\epsilon,$$
$$d_{\mathrm{CF}}(x_l, y) < \frac{\epsilon}{3} + d(x_{j,n}, y_n) < \frac{\epsilon}{3} + \frac{2}{3}\epsilon = \epsilon.$$

Damit ist bewiesen, dass die Folge $(x_j)_{j\in\mathbb{N}}$ im Raum $(\mathrm{CF}_M, d_{\mathrm{CF}})$ gegen y konvergiert. Also ist $(\mathrm{CF}_M, d_{\mathrm{CF}})$ vollständig.

3. Es seien $x, y \in M$ und $i(x), i(y) \in \mathrm{CF}_M$ die konstanten Folgen mit Wert x bzw. y. Dann gilt:
$$d_{\mathrm{CF}}(i(x), i(y)) = \lim_{n\to\infty} d(x, y) = d(x, y).$$
Also ist $i : (M, d) \to (\mathrm{CF}_M, d_{\mathrm{CF}})$ eine Isometrie.

4. Um zu zeigen, dass $i[M]$ dicht in $(\mathrm{CF}_M, d_{\mathrm{CF}})$ ist, sei $x = (x_n)_{n\in\mathbb{N}} \in \mathrm{CF}_M$ gegeben. Wir zeigen, dass die Folge $(i(x_n))_{n\in\mathbb{N}}$ in $(\mathrm{CF}_M, d_{\mathrm{CF}})$ gegen x konvergiert. In der Tat:
$$d_{\mathrm{CF}}(i(x_n), x) = \lim_{m\to\infty} d(x_n, x_m) \leq \sup_{m:\, m\geq n} d(x_n, x_m) \xrightarrow{n\to\infty} 0$$
folgt mit Übung 1.96, weil $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge ist.

□

Im Allgemeinen ist jedoch d_{CF} jedoch *keine* Metrik, selbst dann nicht, wenn d eine Metrik ist. Daher ist $(\mathrm{CF}_M, d_{\mathrm{CF}}, i)$ noch nicht ganz die gesuchte Vervollständigung. Sind zum Beispiel x und y zwei konvergente Folgen (M, d), die gegen den gleichen Grenzwert konvergieren, so gilt $d_{\mathrm{CF}}(x, y) = 0$, obwohl x und y i.A. verschieden sein können. Um aus $(\mathrm{CF}_M, d_{\mathrm{CF}}, i)$ eine Vervollständigung von (M, d) zu gewinnen, verkleben wir noch Punkte mit Abstand 0:

Satz 1.119 (Existenz der Vervollständigung) *Ist (M, d) ein halbmetrischer Raum, so gibt es eine Vervollständigung $(\hat{M}, \hat{d}, \hat{i})$ von (M, d).*

Beweis: Es sei $(\hat{M}, \hat{d})$ der nach Lemma 1.9 durch Verkleben von Punkten mit Abstand 0 aus dem Raum $(\mathrm{CF}_M, d_{\mathrm{CF}})$ gewonnene metrische Raum und $\iota : (\mathrm{CF}_M, d_{\mathrm{CF}}) \to (\hat{M}, \hat{d})$ die kanonische Abbildung (“Verklebeabbildung”). Die Abbildung ι ist surjektiv und isometrisch. Insbesondere ist eine Folge $(x_n)_{n\in\mathbb{N}}$ in $(\mathrm{CF}_M, d_{\mathrm{CF}})$ genau dann eine Cauchyfolge, wenn ihr Bild $(\iota(x_n))$ eine Cauchyfolge in $(\hat{M}, \hat{d})$ ist, und genau dann konvergent in $(\mathrm{CF}_M, d_{\mathrm{CF}})$, wenn $(\iota(x_n))_{n\in\mathbb{N}}$ in $(\hat{M}, \hat{d})$ konvergiert. Damit erbt der Raum $(\hat{M}, \hat{d})$ die Vollständigkeitseigenschaft von $(\mathrm{CF}_M, d_{\mathrm{CF}})$. Nun sei $i : (M, d) \to (\mathrm{CF}_M, d_{\mathrm{CF}})$ die Isometrie aus Lemma 1.118. Wir setzen $\hat{i} := \iota \circ i : (M, d) \to (\hat{M}, \hat{d})$. Dann ist auch $\hat{i}$ eine Isometrie, und das Bild $\hat{i}[M]$ ist dicht in $(\hat{M}, \hat{d})$, weil $i[M]$ dicht in $(\mathrm{CF}_M, d_{\mathrm{CF}})$ und ι surjektiv ist. Also ist $(\hat{M}, \hat{d}, \hat{i})$ die gesuchte Vervollständigung.

$\square$

Weil die Vervollständigung bis auf eine eindeutige bijektive Isometrie eindeutig ist, sprechen wir ab nun von "der" Vervollständigung eines Raumes (M, d) statt von "einer" Vervollständigung. Trotzdem ist es in manchen Anwendungen nützlich, eine "konkretere" Realisierung der Vervollständigung als die "abstrakte" Konstruktion mit Äquivalenzklassen von Cauchyfolgen zur Verfügung zu haben.
Der Fall, dass d selbst eine Metrik auf M ist, ist besonders wichtig. Hier ist $\hat{i} : (M, d) \to (\hat{M}, \hat{d})$ eine isometrische Injektion. In diesem Fall identifizieren wir ab jetzt jedes $x \in M$ mit $\hat{i}(x) \in \hat{M}$ und fassen somit M als eine Teilmenge von $\hat{M}$ und $\hat{d}$ als eine Fortsetzung von d auf.

Beispiel 1.120 (Konstruktion von $\mathbb{R}$) Als Beispiel erhalten wir die Menge der reellen Zahlen $(\mathbb{R}, d_\mathbb{R})$ versehen mit der Standardmetrik als Vervollständigung $(\hat{\mathbb{Q}}, \hat{d}_\mathbb{Q})$ der Menge der rationalen Zahlen $(\mathbb{Q}, d_\mathbb{Q})$ mit der Standardmetrik. Allerdings enhält diese Konstruktion der reellen Zahlen als Äquivalenzklassen von rationalen Cauchyfolgen noch einen kleinen Zirkelschluss, weil wir bei der Konstruktion der Vervollständigung an einigen Stellen schon reelle Zahlen verwendet haben, z.B. in der Definition der Cauchyeigenschaft einer Folge $(a_n)_{n\in\mathbb{N}} \in \mathbb{Q}^\mathbb{N}$:

$$\forall \epsilon \in \mathbb{R}^+ \; \exists n \in \mathbb{N} \; \forall k, l \geq n : |a_k - a_l| < \epsilon.$$

Dieser Zirkelschluss lässt sich aber leicht auflösen, indem man die Verwendung von reellen Zahlen vermeidet, bevor sie konstruiert sind, zum Beispiel unter Verwendung der äquivalenten Definition

$$\forall \epsilon \in \mathbb{Q}^+ \; \exists n \in \mathbb{N} \; \forall k, l \geq n : |a_k - a_l| < \epsilon.$$

Die arithmetischen Operationen $+, \cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ erhält man durch Hochheben der entsprechenden Operationen $+, \cdot : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$ nach Korollar 1.116. Man beachte hierbei, dass $+$ und $\cdot$ auf $\mathbb{Q}$ gleichmäßig stetig auf beschränkten Mengen sind. Die arithmetischen Gesetze (Kommutativgesetze, Assoziativgesetze, Distributivgesetz etc.) vererben sich nach Übung 1.75 mit Dichtheits- und Stetigkeitsargumenten von $\mathbb{Q}$ auf $\mathbb{R}$. Ebenso wird die Ordnungsrelation $\leq \subseteq \mathbb{Q} \times \mathbb{Q}$ durch Abschlussbildung in $\mathbb{R} \times \mathbb{R}$ von $\mathbb{Q}$ auf $\mathbb{R}$ hochgehoben. Wir verzichten hier auf die Ausführung der Details.

Beispiel 1.121 (Vervollständigung (halb-)normierter Räume) Ist $(V, \|\cdot\|)$ ein normierter Raum über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, so wird die Vervollständigung $\hat{V}$ von V wieder zu einem normierten Raum $(\hat{V}, \|\cdot\|)$, indem man die Vektorraumoperationen $+ : V \times V$, $\cdot : \mathbb{K} \times V \to V$ und die Norm $\|\cdot\| : V \to \mathbb{R}$ mittels Korollar 1.116 zu entsprechenden Operationen $+ : \hat{V} \times \hat{V}$, $\cdot : \mathbb{K} \times \hat{V} \to \hat{V}$ und $\|\cdot\| : \hat{V} \to \mathbb{R}$ hochhebt. Dazu verwendet man die gleichmäßige Stetigkeit dieser Operationen auf beschränkten Mengen; siehe Beispiel 1.71. Die Regeln für die Vektorraumoperationen (Kommutativ- und Assoziativgesetz für $+$ etc.) und für die Norm (siehe Definition 1.3) vererben sich mit Dichtheits- und Stetigkeitsargumenten wie in Übung 1.75 auf die Vervollständigung.

Beispiel Dreiecksungleichung: $D := \{(x,y) \in \hat{V} \times \hat{V} | \ \|x\| + \|y\| - \|x+y\| \geq 0\}$ ist abgeschlossen, denn es ist das Urbild der abgeschlossenen Menge $[0, \infty[$ unter der stetigen Abbildung $\hat{V} \times \hat{V} \ni (x,y) \mapsto \|x\| + \|y\| - \|x+y\| \in \mathbb{R}$. Weiter enthält D die dichte Teilmenge $V \times V$, da die Dreiecksungleichung in $(V, \|\cdot\|)$ gilt. Also ist $D = \hat{V} \times \hat{V}$, d.h. die Dreiecksungleichung gilt in $(\hat{V}, \|\cdot\|)$.
Die Vervollständigung eines normierten Raums liefert also einen vollständigen normierten Raum.
Ist $(V, \langle\cdot,\cdot\rangle)$ sogar ein Prähilbertraum, so hat auch das Skalarprodukt eine eindeutige Fortsetzung zu einem Skalarprodukt auf der Vervollständigung. Die dabei benötigte gleichmäßige Stetigkeit des Skalarprodukts auf beschränkten Mengen gilt nämlich nach Beispiel 1.71.
Die Vervollständigung eines Prähilbertraums liefert also einen vollständigen Prähilbertraum.

Definition 1.122 (Banachraum, Hilbertraum) Ein vollständiger normierter Raum wird *Banachraum* genannt. Ein vollständiger Prähilbertraum heißt *Hilbertraum*.

Die Vervollständigung eines normierten Raums bzw. eines Prähilbertraums liefert also einen Banachraum bzw. einen Hilbertraum.

Beispiel 1.123 ($\mathbb{R}^n$ als Banachraum; die Räume $L^p([a,b])$) Für $n \in \mathbb{N}$ und $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ ist $\mathbb{K}^n$ bezüglich jeder Norm darauf ein Banachraum. Ebenso ist für $1 \leq p \leq \infty$ und jede abzählbare Menge I der Raum $\ell^p(I, \mathbb{K})$ ein Banachraum bezüglich $\|\cdot\|_p$; die Vollständigkeit gilt nach dem Satz 1.110. Für $p = 2$ wird $\ell^2(I, \mathbb{K})$ und $\mathbb{K}^n$ mit dem Standardskalarprodukt sogar zu einem Hilbertraum. Für $1 \leq p < \infty$ und $a < b$ ist $(C([a,b], \mathbb{K}), \|\cdot\|_p)$ nach Übung 1.109 jedoch unvollständig. Vervollständigung davon liefert einen Banachraum, der $(L^p([a,b], \mathbb{K}), \|\cdot\|_p)$ genannt wird. Im Fall $p = 2$ erhält man aus $(C([a,b], \mathbb{K}), \langle\cdot,\cdot\rangle)$ durch Vervollständigung sogar einen Hilbertraum $(L^2([a,b], \mathbb{K}), \langle\cdot,\cdot\rangle)$. In der Analysis 3 werden wir eine "konkrete" Realisierung von $L^p([a,b], \mathbb{K})$ als Funktionenraum mit Verkleben besprechen. Insbesondere ist $L^p([a,b], \mathbb{K})$ größer als der Raum der Riemann-integrierbaren Funktionen $\mathcal{R}([a,b], \mathbb{K})$.

Übung 1.124 (Approximation einer Sprungfunktion durch stetige Funktionen) Es sei $a > 0$ und $f_n : [-a,a] \to \mathbb{R}$, $f_n(x) = \min\{1, \max\{0, n(a/2 - |x|)\}\}$ für $n \in \mathbb{N}$. Zeigen Sie, dass die Folge $(f_n)_{n\in\mathbb{N}}$ in $(L^2([-a,a], \mathbb{K}), \|\cdot\|_2)$ gegen ein $f \in L^2([-a,a], \mathbb{K})$ konvergiert. Zeigen Sie $\|f\|_2 = \sqrt{a}$.

Übung 1.125 Es seien $1 \leq p < \infty$, $\alpha > 0$, und $f_n : [0,1] \to \mathbb{R}$, $f_n(x) = \min\{n, x^{-\alpha}\}$ für $n \in \mathbb{N}$, wobei wir formal $0^{-\alpha} := +\infty$ setzen. Bestimmen Sie, für welche Werte von p und α die Folge $(f_n)_{n\in\mathbb{N}}$ in $(L^p([0,1], \mathbb{K}), \|\cdot\|_p)$ konvergiert.

Übung 1.126 Zeigen Sie für $a < b$, dass die Menge $M = \{f \in C([a,b], \mathbb{K}) | \ f(a) = f(b)\}$ dicht in $(C([a,b], \mathbb{K}), \|\cdot\|_2)$ und in $(L^2([a,b], \mathbb{K}), \|\cdot\|_2)$ ist.

Beispiel 1.127 (p-adische ganze Zahlen) Es sei p eine Primzahl und d_p die p-adische Metrik auf $\mathbb{Z}$ aus Beispiel 1.2.3. Die Vervollständigung $(\mathbb{Z}_p, \hat{d}_p)$ von $(\mathbb{Z}, d_p)$ wird *Menge der p-adischen ganzen Zahlen* genannt. Nach Übung 1.72 sind die arithmetischen Operationen $+, \cdot : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ bezüglich d_p gleichmäßig stetig, können also nach Korollar 1.116 zu gleichmäßig stetigen Operationen $+, \cdot : \mathbb{Z}_p \times \mathbb{Z}_p \to \mathbb{Z}_p$. hochgehoben werden. Mit dem Stetigkeits- und Dichtheitsargument aus Beispiel 1.75 vererben sich auch die Assoziativgesetze, Kommutativgesetze, das Distributivgesetz und Eigenschaften von 0 und 1 von $(\mathbb{Z}, +, \cdot)$ nach $(\mathbb{Z}_p, +, \cdot)$. Damit wird $(\mathbb{Z}_p, +, \cdot)$ zu einem kommutativen Ring mit 1.

Übung 1.128 (p-adische Zahlendarstellung) Es sei p eine Primzahl und $(a_n)_{n\in\mathbb{N}_0}$ eine Folge mit Werten in $\{0, 1, \ldots, p-1\}$. Zeigen Sie, dass die Reihe

$$\left(\sum_{k=0}^{n} a_n p^n \right)_{n\in\mathbb{N}}$$

in $(\mathbb{Z}_p, \hat{d}_p)$ konvergiert.

Übung 1.129 ($-1/4$ als Element von $\mathbb{Z}_5$) Wir setzen

$$x = \sum_{k=0}^{\infty} 5^n,$$

wobei die Konvergenz der Reihe in $(\mathbb{Z}_5, \hat{d}_5)$ gemeint ist. Zeigen Sie: $4x = -1$. In diesem Sinn ist $-1/4 \in \mathbb{Z}_5$.

1.9 Der Banachsche Fixpunktsatz

Übung 1.130 Geben Sie eine beliebige Zahl x in Ihren Taschenrechner ein und drücken Sie immer wieder (im Bogenmaß-Modus) die Cosinus-Taste. Wiederholen Sie das Experiment mit anderen Zahlen. Beschreiben Sie, was Sie beobachten.

Definition 1.131 (Fixpunkt) Es sei M eine Menge und $f : M \to M$ eine Abbildung. Ein Punkt $x \in M$ heißt *Fixpunkt* der Abbildung f, wenn $f(x) = x$ gilt.

Wir beschäftigen uns in diesem Abschnitt mit einem hinreichenden Kriterium, dass eine Abbildung $f : M \to M$ einen eindeutigen Fixpunkt besitzt.
Hierzu definieren wir:

Definition 1.132 (Kontraktion) Es sei (M, d) ein halbmetrischer Raum. Eine Abbildung $f : M \to M$ heißt eine *Kontraktion*, wenn es eine reelle Zahl $K < 1$ gibt, *Kontraktionskonstante* genannt, so dass für alle $x, y \in M$ gilt:

$$d(f(x), f(y)) \leq K d(x, y).$$

Achtung, häufige Fehlerquelle: Die Kontraktionskonstante K muss unabhängig von x und y sein, also *gleichmäßig* in x und y gewählt werden. Um zu zeigen, dass $f : M \to M$ eine Kontraktion ist, ist es *nicht* hinreichend, für alle $x, y \in M$ mit $x \neq y$ die Aussage $d(f(x), f(y)) < d(x, y)$ zu zeigen. Zum Beispiel gilt diese Aussage zwar für die Cosinusfunktion auf $\mathbb{R}$ bezüglich der Standardmetrik (Übung: warum?); dennoch ist $\cos : \mathbb{R} \to \mathbb{R}$ *keine* Kontraktion. (Warum nicht?)

Beispiel 1.133 Die Cosinusfunktion bildet $[-1, 1]$ in $[-1, 1]$ ab, und für $K := \sin 1 < 1$ gilt: $\cos : [-1, 1] \to [-1, 1]$ ist eine Kontraktion. In der Tat gibt es nach dem Mittelwertsatz der Differentialrechnung für alle $x, y \in [-1, 1]$ ein ξ zwischen x und y mit $|\cos x - \cos y| \leq |\sin \xi||x - y| \leq K|x - y|$.

Der folgende Satz ist das **wichtigste Werkzeug der Mathematik zur Lösung von Gleichungen und Gleichungssystemen**, nicht nur über $\mathbb{R}$, sondern z.B. auch in $\mathbb{R}^n$ und in Funktionenräumen. Auf ihm beruhen nicht nur viele theoretische Existenzbeweise, sondern auch zahlreiche numerische Verfahren.

Satz 1.134 (Banachscher Fixpunktsatz) *Ist $f : M \to M$ eine Kontraktion auf einem nichtleeren vollständigen metrischen Raum (M, d), so besitzt f genau einen Fixpunkt x^*.*

Ist $x \in M$ ein beliebiger Punkt und definiert man rekursiv:

$$x_0 := x, \quad x_{n+1} = f(x_n) \quad \text{für } n \in \mathbb{N}_0,$$

so gilt $x_n \overset{n\to\infty}{\longrightarrow} x^*$. Bezeichnet $K < 1$ eine Kontraktionskonstante von f, so gilt für alle $n \in \mathbb{N}_0$:

a-priori-Abschätzung:
$$d(x_n, x^*) \leq \frac{K^n}{1-K} d(x_0, x_1) \tag{18}$$

a-posteriori-Abschätzung:
$$d(x_{n+1}, x^*) \leq \frac{K}{1-K} d(x_n, x_{n+1}) \tag{19}$$

Beweis: Wir zeigen zunächst, dass es höchstens einen Fixpunkt von f geben kann. Hierzu seien $x^* \in M$ und $y* \in M$ mit $f(x^*) = x^*$ und $f(y^*) = y^*$ gegeben. Dann folgt

$$d(x^*, y^*) = d(f(x^*), f(y^*)) \leq K d(x^*, y^*)$$

also

$$(1 - K) d(x^*, y^*) \leq 0.$$

Wegen $K < 1$ folgt $d(x^*, y^*) \leq 0$, also $x^* = y^*$.

Nun seien $x \in M$ ein beliebiger Punkt und die Folge $(x_n)_{n\in\mathbb{N}_0}$ hierzu wie oben definiert. Grob gesagt beruht der Beweis auf einem Vergleich von Abständen zwischen Folgengliedern mit einer geometrischen Summe. Genauer gesagt zeigen wir nun:

$$\forall n, m \in \mathbb{N}_0 : d(x_n, x_{n+m}) \leq \frac{1 - K^m}{1 - K} d(x_n, x_{n+1}) \tag{20}$$

Hierzu sei $n \in \mathbb{N}_0$ gegeben. Wir zeigen nun die Ungleichung in (20) für alle $m \in \mathbb{N}_0$ durch vollständige Induktion über m.
Induktionsanfang, $m = 0$: Es gilt

$$d(x_n, x_{n+0}) = 0 = \frac{1 - K^0}{1 - K} d(x_n, x_{n+1}).$$

Induktionsvoraussetzung: Es sei $m \in \mathbb{N}_0$ gegeben und es gelte

$$d(x_n, x_{n+m}) \leq \frac{1 - K^m}{1 - K} d(x_n, x_{n+1}).$$

Induktionsschluss, $m \rightsquigarrow m + 1$: Aus der rekursiven Definition der Folge $(x_k)_{k \in \mathbb{N}_0}$, der Kontraktionseigenschaft von f und der Induktionsvoraussetzung erhalten wir

$$d(x_{n+1}, x_{n+m+1}) = d(f(x_n), f(x_{n+m})) \leq K d(x_n, x_{n+m}) \leq K \frac{1 - K^m}{1 - K} d(x_n, x_{n+1})$$

und damit die Induktionsbehauptung:

$$\begin{aligned}
d(x_n, x_{n+m+1}) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+m+1}) \\
&\leq d(x_n, x_{n+1}) + K \frac{1 - K^m}{1 - K} d(x_n, x_{n+1}) \\
&= \left(1 + K \frac{1 - K^m}{1 - K}\right) d(x_n, x_{n+1}) \\
&= \frac{1 - K + K - K^{m+1}}{1 - K} d(x_n, x_{n+1}) \\
&= \frac{1 - K^{m+1}}{1 - K} d(x_n, x_{n+1}).
\end{aligned}$$

Weiter zeigen wir

$$\forall n \in \mathbb{N}_0 : d(x_n, x_{n+1}) \leq K^n d(x_0, x_1) \tag{21}$$

durch vollständige Induktion.
Der *Induktionsanfang, $n = 0$,* ist trivial, denn es gilt

$$d(x_0, x_1) = K^0 d(x_0, x_1).$$

Induktionsvoraussetzung: Es sei $n \in \mathbb{N}_0$ gegeben und es gelte

$$d(x_n, x_{n+1}) \leq K^n d(x_0, x_1).$$

Induktionsschluss, $n \rightsquigarrow n + 1$: Es folgt die Induktionsbehauptung so:

$$d(x_{n+1}, x_{n+2}) = d(f(x_n), f(x_{n+1})) \leq K d(x_n, x_{n+1}) \leq K^{n+1} d(x_0, x_1),$$

wobei wir beim ersten "$\leq$" die Kontraktionseigenschaft und beim zweiten "$\leq$" die Induktionsvoraussetzung verwendet haben.
Die Kombination der Ungleichungen (20) und (21) liefert für alle $n, m \in \mathbb{N}_0$:

$$d(x_n, x_{n+m}) \leq \frac{1-K^m}{1-K} d(x_n, x_{n+1}) \leq K^n \frac{1-K^m}{1-K} d(x_0, x_1) \leq \frac{K^n}{1-K} d(x_0, x_1)$$

wobei wir im letzten Schritt $K < 1$ verwendet haben.
Mit der Übung 1.96 folgt, dass $(x_n)_{n \in \mathbb{N}_0}$ eine Cauchyfolge ist, denn

$$\limsup_{n\to\infty} \sup_{m \in \mathbb{N}_0} d(x_n, x_{n+m}) \leq \limsup_{n\to\infty} \frac{K^n}{1-K} d(x_0, x_1) = 0,$$

wobei wir im letzten Schritt verwendet haben, dass aus $0 \leq K < 1$ folgt: $\lim_{n\to\infty} K^n = 0$.
Wegen der Vollständigkeit von (M, d) konvergiert x_n für $n \to \infty$ gegen ein $x^* \in M$.
Wir zeigen nun, dass x^* ein Fixpunkt von f ist. Als Kontraktion ist f gleichmäßig stetig, also auch stetig. Es folgt:

$$x^* = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} f(x_n) = f(x^*)$$

also ist x^* ein Fixpunkt von f.
Schließlich gilt wegen $\lim_{n\to\infty} x_n = x^*$ und der Stetigkeit der Metrik für alle $n \in \mathbb{N}_0$ die Abschätzung:

$$d(x_n, x^*) = \lim_{m\to\infty} d(x_n, x_{n+m}) \leq \lim_{m\to\infty} \frac{1-K^m}{1-K} d(x_n, x_{n+1}) = \frac{d(x_n, x_{n+1})}{1-K},$$

wobei wir (20) verwendet haben. Die a-priori Abschätzung (18) folgt hieraus mit (21) so:

$$d(x_n, x^*) \leq \frac{d(x_n, x_{n+1})}{1-K} \leq \frac{K^n}{1-K} d(x_0, x_1),$$

und die a-posteriori Abschätzung (19) so:

$$d(x_{n+1}, x^*) = d(f(x_n), f(x^*)) \leq K d(x_n, x^*) \leq \frac{K}{1-K} d(x_n, x_{n+1}).$$

$\square$

Beispiel 1.135 (Fortsetzung von Beispiel 1.133) Für alle $y \in \mathbb{R}$ ist $x = \cos y \in [-1, 1]$. Nun ist $[-1, 1] \subset \mathbb{R}$ bzgl. der Standardmetrik nach Lemma 1.106 vollständig, denn es ist eine abgeschlossene Teilmenge des vollständigen Raums $\mathbb{R}$. Da $\cos : [-1, 1] \to [-1, 1]$ eine Kontraktion ist, besitzt diese Funktion einen eindeutigen Fixpunkt x^*, und Iteration des Cosinus bei beliebigem Startpunkt $y \in \mathbb{R}$ liefert eine gegen x^* konvergente Folge.

Beispiel 1.136 (Kontrahierende Abbildungen auf Teilmengen von $\mathbb{R}$, Heronverfahren) Ist allgemeiner $M \subseteq \mathbb{R}$ ein abgeschlossenes Intervall und $f : M \to M$ eine stetige, im Inneren M° von M differenzierbare Abbildung mit $K := \sup_{x \in M^\circ} |f'(x)| < 1$,

so ist nach dem Mittelwertsatz der Differentialrechnung die Abbildung f eine Kontraktion bezüglich der Standardmetrik. Da M bezüglich der Standardmetrik vollständig ist, ist der Banachsche Fixpunktsatz anwendbar. Insbesondere besitzt f also einen eindeutigen Fixpunkt.
Als ein Beispiel betrachten wir das Heronverfahren zur numerischen Berechnung von Quadratwurzeln $\sqrt{a}$, $a > 0$. Es wird durch die Rekursionsvorschrift

$$x_{n+1} = \frac{1}{2}\left(\frac{a}{x_n} + x_n\right), \quad n \in \mathbb{N}_0$$

mit einem Startwert $x_0 > 0$ definiert.
Die Abbildung

$$f : \mathbb{R}^+ \to \mathbb{R}^+, \quad f(x) = \frac{1}{2}\left(\frac{a}{x} + x\right)$$

hat den Wertebereich $f[\mathbb{R}^+] = [\sqrt{a}, \infty[$ und die Ableitung

$$f'(x) = \frac{1}{2}\left(1 - \frac{a}{x^2}\right).$$

Eingeschränkt auf $f[\mathbb{R}^+] = [\sqrt{a}, \infty[$ nimmt f' Werte in $[0, 1/2[$ an. Also ist die Einschränkung $f : [\sqrt{a}, \infty[\to [\sqrt{a}, \infty[$ eine Kontraktion mit der Kontraktionskonstante $K = \frac{1}{2}$. Das Heronverfahren konvergiert also für alle Startwerte in $f[\mathbb{R}^+]$ und daher (mit einem Schritt zu Beginn mehr) auch für alle Startwerte in $\mathbb{R}^+$. Es konvergiert mit Startwerten x_0 nahe bei $\sqrt{a}$ sogar wesentlich schneller, als es die Schranken aus dem Banachschen Fixpunktsatz vermuten lassen. Das liegt daran, dass $f'(\sqrt{a}) = 0$ gilt.

Übung 1.137 Berechnen Sie mit Taschenrechnergenauigkeit die ersten drei Näherungen x_1, x_2, x_3 an $\sqrt{2}$ nach dem Heronverfahren mit der Startnäherung $x_0 = 2$. Berechnen Sie die a-priori-Schranke und die a-posteriori-Schranke für den Fehler $|x_n - \sqrt{2}|$, $n = 1, 2, 3$, die sich aus dem Banachschen Fixpunktsatz mit der Kontraktionskonstante $K = \frac{1}{2}$ ergibt. Vergleichen Sie (mit Taschenrechnergenauigkeit) diese Fehlerschranken mit dem tatsächlichen Wert des Fehlers $|x_n - \sqrt{2}|$.

Beispiel 1.138 (Lösung einer Integralgleichung) Gegeben seien $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, zwei reelle Zahlen $a < b$, eine stetige Funktion $g : [a, b] \to \mathbb{K}$ und eine gleichmäßig stetige Funktion $K : [a, b] \times [a, b] \to \mathbb{K}$ mit

$$k := \sup_{x \in [a,b]} \int_a^b |K(x, y)| \, dy < 1.$$

Wir betrachten die folgende Integralgleichung für eine unbekannte stetige Funktion $f : [a, b] \to \mathbb{K}$:

$$f(x) = \int_a^b K(x, y) f(y) \, dy + g(x) \text{ für } x \in [a, b]. \tag{22}$$

Es sei

$$L : (C([a,b],\mathbb{K}), \|\cdot\|_\infty) \to (C([a,b],\mathbb{K}), \|\cdot\|_\infty)$$
$$L(f)(x) = \int_a^b K(x,y)f(y)\,dy \text{ für } f \in C([a,b],\mathbb{K}),\ x \in [a,b]$$

der Integraloperator aus Beispiel 1.66. Dort wurde gezeigt:

$$\forall f \in C([a,b],\mathbb{K}) : \|L(f)\|_\infty \le k\|f\|_\infty.$$

Es folgt wegen $k < 1$: Die Abbildung

$$\Phi : C([a,b],\mathbb{K}) \to C([a,b],\mathbb{K}), \quad \Phi(f) = L(f) + g$$

ist eine Kontraktion, da für $f, \tilde{f} \in C([a,b],\mathbb{K})$ gilt:

$$\|\Phi(f) - \Phi(\tilde{f})\|_\infty = \|L(f) - L(\tilde{f})\|_\infty = \|L(f - \tilde{f})\|_\infty \le k\|f - \tilde{f}\|_\infty.$$

Da $(C([a,b],\mathbb{K}), \|\cdot\|_\infty)$ nach Satz 1.107 vollständig ist,[7] folgt aus dem Banachschen Fixpunktsatz die Existenz einer eindeutigen Lösung $f \in C([a,b],\mathbb{K})$ der Integralgleichung (22), geschrieben als Fixpunktgleichung $\Phi(f) = f$.

Beispiel 1.139 (Die imaginäre Einheit in $\mathbb{Z}_5$) Wir betrachten die Gleichung $x^2+1 = 0$ in der Vervollständigung $(\mathbb{Z}_5, \hat{d}_5)$ von $\mathbb{Z}$ bezüglich der 5-adischen Metrik d_5.
Schreiben wir die Gleichung $x^2 + 1 = 0$ in Fixpunktform $f(x) = x$ mit der stetigen Abbildung $f : \mathbb{Z}_5 \to \mathbb{Z}_5$, $f(x) = x^2 + x + 1$. Es sei $\overline{M} = 2 + 5\mathbb{Z}_5$ der Abschluss in $(\mathbb{Z}_5, \hat{d}_5)$ der Menge

$$M = \{x \in \mathbb{Z} |\ d_5(x,2) \le \frac{1}{5}\} = 2 + 5\mathbb{Z}.$$

Insbesondere ist $\overline{M}$ bezüglich $\hat{d}_5$ vollständig.
Es gilt $f[M] \subseteq M$, denn für $x \in M$ folgt $x^2 + x + 1 \in 2^2 + 2 + 1 + 5\mathbb{Z} = 7 + 5\mathbb{Z} = 2 + 5\mathbb{Z}$.
Wegen der Stetigkeit von f folgt $f[\overline{M}] \subseteq \overline{M}$.
Nun gilt für alle $x, y \in M$: $x + y + 1 \in 2 + 2 + 1 + 5\mathbb{Z} = 5\mathbb{Z}$, also $v_5(x + y + 1) \ge 1$. Es folgt

$$\begin{aligned} v_5(f(x) - f(y)) &= v_5((x^2 + x + 1) - (y^2 + y + 1)) = v_5((x - y)(x + y + 1)) \\ &= v_5(x - y) + v_5(x + y + 1) \ge v_5(x - y) + 1 \end{aligned}$$

und daher

$$d_5(f(x), f(y)) \le \frac{1}{5} d_5(x,y).$$

Wegen der Stetigkeit der Metrik und der Abbildung f folgt auch für alle $x, y \in \overline{M}$

$$\hat{d}_5(f(x), f(y)) \le \frac{1}{5} \hat{d}_5(x,y).$$

Insbesondere ist $f : (\overline{M}, \hat{d}_5) \to (\overline{M}, \hat{d}_5)$, $f(x) = x^2 + x + 1$, eine Kontraktion. Nach dem Banachschen Fixpunktsatz folgt: Es gibt ein eindeutig bestimmtes $i \in \overline{M}$ mit $f(i) = i$, also mit $i^2 = -1$.

[7] Man beachte $C([a,b],\mathbb{K}) = C_b([a,b],\mathbb{K})$.

Übung 1.140 (Approximation an i in $\mathbb{Z}_5$) Berechnen Sie ein $x \in \{0, 1, \ldots, 124\}$ mit $\hat{d}_5(x, i) \leq 1/125$, indem Sie die Rekursion aus dem Banachschen Fixpunktsatz auf die Abbildung $f : \overline{M} \to \overline{M}$ aus Beispiel 1.139 mit dem Startwert $x_0 = 2$ anwenden. Stellen Sie x und $x^2 + 1$ in 5-adischer Darstellung $\sum_{k=0}^{n} a_k 5^k$ mit $a_k \in \{0, 1, 2, 3, 4\}$ dar.

1.10 Existenz- und Eindeutigkeitssatz für Anfangswertprobleme

Als eine Anwendung des Banachschen Fixpunktsatzes beweisen wir einen Existenz- und Eindeutigkeitssatz für Anfangswertprobleme zu Differentialgleichungssystemen. Dabei besprechen wir nur eine einfache Variante; Verallgemeinerungen werden in der Vorlesung "Gewöhnliche Differentialgleichungen" behandelt.

Beispiel 1.141 (Schwingungsgleichung) Aus der Analysis 1 wissen Sie: Für gegebene $b_1, b_2 \in \mathbb{R}$ besitzt das Schwingungsgleichungssystem

$$\begin{aligned} y_1'(x) &= y_2(x), \\ y_2'(x) &= -y_1(x) \end{aligned}$$

mit der Anfangsbedingung

$$\begin{aligned} y_1(0) &= b_1, \\ y_2(0) &= b_2 \end{aligned}$$

genau eine Lösung $y_1 : \mathbb{R} \to \mathbb{R}$, $y_2 : \mathbb{R} \to \mathbb{R}$, nämlich

$$\begin{aligned} y_1(x) &= b_1 \cos x + b_2 \sin x, \\ y_2(x) &= -b_1 \sin x + b_2 \cos x. \end{aligned}$$

Hier ist eine viel allgemeinere Fragestellung. Wie immer sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

Definition 1.142 (Anfangswertproblem) Es sei $I \subseteq \mathbb{R}$ ein Intervall. Ein Gleichungssystem für unbekannte Funktionen $y_1, \ldots, y_n : I \to \mathbb{K}$ der Gestalt

$$y_j'(x) = f_j(x, y_1(x), \ldots, y_n(x)) \text{ für } j = 1, \ldots, n$$

mit gegebenen Funktionen $f_j : I \times U \to \mathbb{K}$, $j = 1, \ldots, n$ mit $U \subseteq \mathbb{K}^n$ offen, wird (gewöhnliches) *Differentialgleichungssystem 1. Ordnung* genannt. Gegeben $a \in I$ und ein Vektor $b = (b_1, \ldots, b_n) \in U$, nennen wir eine Bedingung der Gestalt

$$y_j(a) = b_j \text{ für } j = 1, \ldots, n$$

eine *Anfangsbedingung*. Die Frage nach den Lösungen des Differentialgleichungssystems, die eine gegebene Anfangsbedingung erfüllen, nennen wir ein *Anfangswertproblem*.

Im Fall, dass ein Randpunkt a von I zum Intervall I gehört, ist die Ableitung bei a natürlich als einseitige Ableitung zu verstehen.

Folgende Kurznotation ist praktisch: Wir fassen $y_1, \ldots, y_n : I \to \mathbb{K}$ zu einer vektorwertigen Abbildung $y = (y_1, \ldots, y_n) : I \to \mathbb{K}^n$, $y(x) = (y_1(x), \ldots, y_n(x))$ zusammen. Wir kürzen ab: $y' = (y_1', \ldots, y_n')$ und

$$\int_a^b y(x)\, dx = \left(\int_a^b y_j(x)\, dx \right)_{j=1,\ldots,n}$$

Ebenso fassen wir die f_j zu einer Abbildung $f = (f_1, \ldots, f_n) : I \times U \to \mathbb{K}^n$, $f(x, z) = (f_1(x, z), \ldots, f_n(x, z))$ für $(x, z) \in I \times U$ zusammen. Damit können wir das Anfangswertproblem kurz als

$$y'(x) = f(x, y(x)) \text{ für } x \in I, \quad y(a) = b$$

schreiben. Zur Vereinfachung beschränken wir uns ab nun auf den wichtigsten Fall $U = \mathbb{K}^n$; allgemeinere Fälle werden in der Vorlesung "Gewöhnliche Differentialgleichungen" behandelt.
Im Zusammenhang mit Matrixnotationen ist es zweckmäßig, sich y und f als *Spaltenvektoren*

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad f = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix},$$

statt als Zeilenvektoren vorzustellen. Zum Beispiel erhält damit das Schwingungsgleichungssystem die Gestalt

$$y'(x) = Ay(x), \quad y(0) = b$$

mit

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Allgemeiner nennt man ein Differentialgleichungssystem der Gestalt

$$y'(x) = Ay(x) \quad x \in \mathbb{R}$$

mit einer gegebenen Matrix $A \in \mathbb{K}^{n \times n}$ und einer gesuchten vektorwertigen Funktion $y : I \to \mathbb{K}^n$ ein *homogenes lineares Differentialgleichungssystem mit konstanten Koeffizienten.*
Fügt man noch eine gegebene Funktion $g : I \to \mathbb{K}^n$ in der Form

$$y'(x) = Ay(x) + g(x) \quad x \in \mathbb{R}$$

hinzu, so spricht man von einem *inhomogenen linearen Differentialgleichungssystem mit konstanten Koeffizienten.* Zum Beispiel führen Schwingungen eines Massepunkts unter dem Einfluss zusätzlicher äußerer Störungen zu solchen inhomogenen linearen Differentialgleichungssystemen.
Der Schlüssel zur allgemeinen Lösung von Anfangswertproblemen liegt in folgender Übersetzung in eine Integralgleichung.

Lemma 1.143 (Volterra-Integralgleichung) *Gegeben seien $I \subseteq \mathbb{R}$ ein Intervall, $a \in I$, $n \in \mathbb{N}$, $b \in \mathbb{K}^n$ und zwei stetige Funktionen $f : I \times \mathbb{K}^n \to \mathbb{K}^n$ und $y : I \to \mathbb{K}^n$. Die Funktion y ist genau dann (komponentenweise) differenzierbar und löst das Anfangswertproblem*

$$y'(x) = f(x, y(x)) \text{ für alle } x \in I, \quad y(a) = b,$$

wenn sie die folgende Integralgleichung löst:

$$y(x) = b + \int_a^x f(t, y(t))\, dt \text{ für alle } x \in I.$$

Diese Integralgleichung wird Volterra-Integralgleichung zu dem Anfangswertproblem genannt.

Beweis: Die Aussage des Lemma folgt unmittelbar aus dem Hauptsatz der Differential- und Integralrechnung.

□

Wir lösen nun das Anfangswertproblem mit dem Banachschen Fixpunktsatz, indem wir die Volterra-Integralgleichung als eine Fixpunktgleichung für die unbekannte Funktion y auffassen.

Satz 1.144 (Satz von Picard-Lindelöf) *Gegeben seien ein kompaktes Intervall $I \subset \mathbb{R}$, $a \in I$, $n \in \mathbb{N}$, $b \in \mathbb{K}^n$ und eine stetige Funktion $f : I \times \mathbb{K}^n \to \mathbb{K}^n$. Weiter existiere eine "Lipschitzkonstante" $L \geq 0$, so dass f Lipschitz-stetig im 2. Argument mit der Konstanten L im folgenden Sinn ist:*

$$\forall x \in I \; \forall y, z \in \mathbb{K}^n : \; \|f(x, y) - f(x, z)\| \leq L\|y - z\|$$

mit einer beliebigen Norm $\|\cdot\|$ auf $\mathbb{K}^n$. Dann besitzt das Anfangswertproblem

$$y'(x) = f(x, y(x)) \text{ für alle } x \in I, \quad y(a) = b$$

eine eindeutige Lösung $y \in C(I, \mathbb{K}^n)$.

Beweis: Wir fixieren ein $M > L$ und und arbeiten im normierten Raum $(C(I, \mathbb{K}^n), \|\cdot\|_{M,a})$ mit der Norm

$$\|y\|_{M,a} := \sup_{x \in I} e^{-M|x-a|} \|y(x)\|.$$

Wegen der Kompaktheit von I ist für jedes stetige $y : I \to \mathbb{K}^n$ dieses Supremum endlich. Insbesondere ist $C(I, \mathbb{K}^n) = C_b(I, \mathbb{K}^n)$.
Die Volterra-Integralgleichung zum gegebenen Anfangswertproblem ist genau die Fixpunktgleichung $\Phi(y) = y$ zur Abildung

$$\begin{aligned} &\Phi : C_b(I, \mathbb{K}^n) \to C_b(I, \mathbb{K}^n), \\ &\qquad \Phi(y)(x) = b + \int_a^x f(t, y(t))\, dt \text{ für } y \in C_b(I, \mathbb{K}^n), x \in I. \end{aligned}$$

Man beachte, dass $\Phi(y)$ für alle $y \in C_b(I, \mathbb{K}^n)$ sogar differenzierbar, also erst recht stetig ist. Für $y, \tilde{y} \in C_b(I, \mathbb{K}^n)$ und $t \in I$ erhalten wir:

$$\|y(t) - \tilde{y}(t)\| \leq e^{M|t-a|} \|y - \tilde{y}\|_{M,a},$$

also mit der Integralversion der Dreiecksungleichung aus Übung 1.8 für $x \in I$:[8]

$$\begin{aligned}
&\|\Phi(y)(x) - \Phi(\tilde{y})(x)\| = \left\| \int_a^x f(t, y(t))\, dt - \int_a^x f(t, \tilde{y}(t))\, dt \right\| \\
&= \left\| \int_a^x [f(t, y(t)) - f(t, \tilde{y}(t))]\, dt \right\| \leq \left| \int_a^x \|f(t, y(t)) - f(t, \tilde{y}(t))\|\, dt \right| \\
&\leq \left| \int_a^x L\|y(t) - \tilde{y}(t)\|\, dt \right| \leq \left| \int_a^x L e^{M|t-a|} \|y - \tilde{y}\|_{M,a}\, dt \right| \\
&= L\|y - \tilde{y}\|_{M,a} \left| \int_a^x e^{M|t-a|}\, dt \right| = L\|y - \tilde{y}\|_{M,a} \frac{e^{M|x-a|} - 1}{M} \\
&\leq K\|y - \tilde{y}\|_{M,a} e^{M|x-a|},
\end{aligned}$$

wobei wir $K := L/M \in]0, 1[$ gesetzt haben.
Dividieren wir durch $e^{M|x-a|}$ und bilden wir das Supremum über $x \in I$, erhalten wir

$$\|\Phi(y) - \Phi(\tilde{y})\|_{M,a} \leq K\|y - \tilde{y}\|_{M,a}.$$

Die Abbildung Φ ist also eine Kontraktion mit Kontraktionskonstante $K = L/M$ auf dem Raum $(C(I, \mathbb{K}^n), \|\cdot\|_{M,a})$.
Um den Banachschen Fixpunktsatz auf die Abbildung Φ anwenden zu können, müssen wir uns noch von der Vollständigkeit des Raums $(C(I, \mathbb{K}^n), \|\cdot\|_{M,a})$ überzeugen. Hierzu vergleichen wir die "gewichtete" Norm $\|\cdot\|_{M,a}$ mit der ungewichteten Norm $\|y\|_\infty := \sup_{x \in I} \|y(x)\|$ mit Hilfe der Multiplikationsabbildung mit $e^{M|\cdot - a|}$, gegeben durch

$$\begin{aligned}
w : (C_b(I, \mathbb{K}^n), \|\cdot\|_\infty) &\to (C_b(I, \mathbb{K}^n), \|\cdot\|_{M,a}), \\
w(y)(x) &= e^{M|x-a|} y(x) \text{ für } y \in C_b(I, \mathbb{K}^n), x \in I.
\end{aligned}$$

w ist offensichtlich eine isometrische Bijektion bezüglich der angegebenen Normen. Weil $(C_b(I, \mathbb{K}^n), \|\cdot\|_\infty)$ nach Satz 1.107 vollständig ist (denn $(\mathbb{K}^n, \|\cdot\|)$ ist vollständig), folgt, dass auch $(C_b(I, \mathbb{K}^n), \|\cdot\|_{M,a})$ vollständig ist.
Damit liefert der Banachsche Fixpunktsatz, dass die Abbildung Φ genau einen Fixpunkt $y \in C_b(I, \mathbb{K}^n)$ besitzt. Die Volterra-Integralgleichung $\Phi(y) = y$ und damit das gegebene Anfangswertproblem besitzen also eine eindeutige Lösung y.

□

[8] Zur Vereinfachung kann man sich $x \geq a$ vorstellen, da in diesem Fall die Absolutbeträge weggelassen werden können.

Übung 1.145 (Fehlerabschätzung im Iterationsverfahren nach Picard-Lindelöf) Mit den Bezeichnungen von eben sei

$$(y^{(m)})_{m\in\mathbb{N}_0}$$

die wie im Banachschen Fixpunktsatz rekursiv definierte Folge mit der konstanten Startfunktion

$$y^{(0)}(x) = b \text{ für } x \in I$$

und dem Rekursionsschritt

$$y^{(m+1)} = \Phi(y^{(m)}) \text{ für } m \in \mathbb{N}_0.$$

Es sei $C = \sup_{x\in I} \|f(x,b)\|$. Zeigen Sie für alle $x \in I$ und $m \in \mathbb{N}_0$:

$$\|y^{(m+1)}(x) - y^{(m)}(x)\| \le CL^m \frac{|x-a|^{m+1}}{(m+1)!}.$$

Folgern Sie für den Fixpunkt y von Φ, $x \in I$ und $m \in \mathbb{N}_0$ im Fall $L > 0$:

$$\begin{aligned}\|y(x) - y^{(m)}(x)\| &\le C \sum_{k=m+1}^{\infty} L^{k-1} \frac{|x-a|^k}{k!} \\ &= \frac{C}{L}\left(e^{L|x-a|} - \sum_{k=0}^{m} \frac{(L|x-a|)^k}{k!}\right) = C\left|\int_a^x \frac{|L(x-t)|^m}{m!} e^{L|t-a|}\,dt\right|.\end{aligned}$$

Hinweis: Verwenden Sie die Lagrange-Darstellung des Taylor-Restglieds.

Beispiel 1.146 (Lineare homogene Differentialgleichungssysteme mit konstanten Koeffizienten) Es sei $A \in \mathbb{K}^{n\times n}$ eine $n \times n$-Matrix und $b \in \mathbb{K}^n$. Wir lösen das Anfangswertproblem

$$y'(x) = Ay(x), \quad y(0) = b \tag{23}$$

mit dem Iterationsverfahren aus dem Beweis des Satzes von Picard-Lindelöf:

$$\begin{aligned}y^{(0)}(x) &:= b, \\ y^{(m+1)}(x) &:= b + \int_0^x Ay^{(m)}(t)\,dt \text{ für } x \in \mathbb{R}, m \in \mathbb{N}_0.\end{aligned}$$

Wir zeigen induktiv für alle $x \in \mathbb{R}$ und $m \in \mathbb{N}_0$:

$$y^{(m)}(x) = \sum_{k=0}^{m} \frac{x^k}{k!} A^k b.$$

Das ist offensichtlich für $m = 0$. Als Induktionsvoraussetzung nehmen wir an, dass die Behauptung für ein gegebenes $m \in \mathbb{N}_0$ und alle $x \in \mathbb{R}$ gelte. Dann folgt für $x \in \mathbb{R}$:

$$
\begin{aligned}
y^{(m+1)}(x) &= b + \int_0^x Ay^{(m)}(t)\,dt = b + A\sum_{k=0}^{m}\int_0^x \frac{t^k}{k!}A^k b\,dt = b + \sum_{k=0}^{m}\int_0^x \frac{t^k}{k!}\,dt\,A^{k+1}b \\
&= b + \sum_{k=0}^{m}\frac{x^{k+1}}{(k+1)!}A^{k+1}b = b + \sum_{k=1}^{m+1}\frac{x^k}{k!}A^k b = \sum_{k=0}^{m+1}\frac{x^k}{k!}A^k b,
\end{aligned}
$$

wie behauptet.
Nun ist die Multiplikationsabbildung mit A, $L_A : \mathbb{K}^n \to \mathbb{K}^n$, $L_A(u) := Au$ Lipschitz-stetig, denn es gilt $\|L_A u - L_A v\| = \|A(u-v)\| \leq \|A\|\|u-v\|$ für $u, v \in \mathbb{K}^n$ mit der Operatornorm $\|A\|$ von A; vgl Übung 1.67. Der Beweis des Satzes von Picard-Lindelöf liefert, dass

$$
y : \mathbb{R} \to \mathbb{K}^n, \quad y(x) = \lim_{m\to\infty}\sum_{k=0}^{m}\frac{1}{k!}(xA)^k b = \sum_{k=0}^{\infty}\frac{1}{k!}(tA)^k b
$$

das obige Anfangswertproblem löst.

In Verallgemeinerung der Exponentialfunktion auf $\mathbb{C}$ definieren wir

Definition 1.147 (Matrix-Exponentialfunktion) *Für $A \in \mathbb{C}^{n\times n}$ sei*

$$
e^A = \exp(A) := \sum_{k=0}^{\infty}\frac{1}{k!}A^k,
$$

wobei die Konvergenz in der Standardtopologie auf $\mathbb{C}^{n\times n}$ gemeint ist.

Wegen $\|A^k\| \leq \|A\|^k$ (vgl. Übung 1.69) wird die Matrix-Exponentialreihe durch die gewöhnliche Exponentialreihe dominiert:

$$
\left\|\sum_{k=m}^{l}\frac{1}{k!}A^k\right\| \leq \sum_{k=m}^{l}\frac{\|A^k\|}{k!} \leq \sum_{k=m}^{l}\frac{\|A\|^k}{k!}
$$

für $m \leq l$, so dass die Konvergenz der Matrix-Exponentialreihe aus der Vollständigkeit von $(\mathbb{K}^{n\times n}, \|\cdot\|)$ und der Cauchyfolgeneigenschaft der gewöhnlichen Exponentialreihe folgt. Fassen wir zusammen:

Satz 1.148 (Lösung linearer homogener Differentialgleichungssysteme mit konstanten Koeffizienten) *Das Anfangswertproblems (23) besitzt die eindeutige Lösung*

$$
y(x) = e^{xA}b, \quad x \in \mathbb{R}.
$$

Wir haben nämlich diese eindeutige Lösung zunächst auf beliebigen kompakten Intervallen $I \ni 0$, also auch auf ganz $\mathbb{R}$.

Beispiel 1.149 (Drehungen in $\mathbb{R}^2$) Für das Schwingungsgleichungs-Anfangswertproblem

$$y'(x) = Ay(x), \quad y(0) = b$$

mit

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

errechnen wir

$$A^0 = \mathrm{Id}, \quad A^1 = A, \quad A^2 = -\mathrm{Id}, \quad A^3 = -A, \quad A^4 = \mathrm{Id}$$

und daher

$$A^{2k} = (-1)^k \mathrm{Id}, \quad A^{2k+1} = (-1)^k A$$

für alle $k \in \mathbb{N}_0$. Wir erhalten für $x \in \mathbb{R}$:

$$e^{xA} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k} \mathrm{Id} + \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1} A = (\cos x) \mathrm{Id} + (\sin x) A = \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix},$$

und daher die eindeutige Lösung

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

Natürlich ist uns diese Lösung längst bekannt.

Berechnung der Matrix-Exponentialfunktion. Wir besprechen nun, wie man e^A für Matrizen $A \in \mathbb{C}^{n \times n}$ berechnen kann. Der einfachste Fall ist der einer Diagonalmatrix $A = \mathrm{diag}(a_1, \ldots, a_n)$ mit Diagonaleinträgen $a_1, \ldots, a_n \in \mathbb{C}$. Hier erhalten wir $A^k = \mathrm{diag}(a_1^k, \ldots, a_n^k)$ für alle $k \in \mathbb{N}_0$, also

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} \mathrm{diag}(a_1^k, \ldots, a_n^k) = \mathrm{diag}(e^{a_1}, \ldots, e^{a_n}).$$

Ein weiterer einfacher Fall betrifft die nilpotente Matrix

$$N = \begin{pmatrix} 0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & 0 & \ldots & 0 \end{pmatrix} \in \mathbb{C}^{n \times n},$$

die Einsen in der ersten oberen Nebendiagonale und sonst Nullen als Einträge besitzt. Die Matrix N^k besitzt für $k \in \mathbb{N}_0$ mit $k < n$ Einsen in der k-ten oberen Nebendiagonale und Nullen sonst für $k < n$. Für $k \geq n$ ist $N^k = 0$. Es folgt für $x \in \mathbb{R}$:

$$e^{xN} = \sum_{k=0}^{n-1} \frac{x^k}{k!} N^k = \begin{pmatrix} 1 & x & x^2/2! & x^3/3! & \dots & x^{n-1}/(n-1)! \\ 0 & 1 & x & x^2/2! & \dots & x^{n-2}/(n-2)! \\ 0 & 0 & 1 & x & \dots & x^{n-3}/(n-3)! \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 & \ddots & x \\ 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}, \tag{24}$$

also $(e^{xN})_{ij} = 0$ für $i, j \in \{1, \dots n\}$ mit $i > j$ und $(e^{xN})_{ij} = x^{j-i}/(j-i)!$ für $i, j \in \{1, \dots n\}$ mit $i \leq j$.
Übersetzt in die Sprache von Anfangswertproblemen bedeutet das: Das Anfangswertproblem

$$y' = Ny, \quad y(0) = b$$

für die obige nilpotente Matrix N, also

$$\begin{aligned} y_i' &= y_{i+1}, \quad y_i(0) = b_i \text{ für } i = 1, \dots, n-1, \\ y_n' &= 0, \quad y_n(0) = b_n \end{aligned}$$

besitzt die eindeutige Lösung

$$y_i(x) = \sum_{j=i}^{n} \frac{x^{j-i}}{(j-i)!} b_j \text{ für } i = 0, \dots, n \text{ und } x \in \mathbb{R}.$$

Natürlich ist uns y_1 längst als Taylorpolynom bei gegebenen Ableitungen

$$\left. \frac{d^{i-1}}{dx^{i-1}} y_1(x) \right|_{x=0} = y_i(0) = b_i$$

und $y_i(x)$ für $i = 2, \dots, n$ als dessen $(i-1)$-te Ableitung bekannt.
Kompliziertere Fälle behandeln wir durch Zusammensetzen aus den einfachen Fällen. Dazu brauchen wir

Lemma 1.150 (Funktionalgleichung für die Matrix-Exponentialfunktion) *Sind* $A, B \in \mathbb{C}^{n \times n}$, $n \in \mathbb{N}$, *zwei Matrizen mit*

$$AB = BA,$$

so gilt

$$e^{A+B} = e^A e^B.$$

Achtung, Fehlerquelle:
Wenn $AB \neq BA$, gilt die Funktionalgleichung im allgemeinen *nicht*!

Beweis der Funktionalgleichung: Der Beweis verläuft ganz analog zum Beweis der Funktionalgleichung für die gewöhnliche Exponentialfunktion; daher beschränken wir uns hier auf eine Beweisskizze.
Wie im Beweis der binomischen Formel sieht man

$$(A+B)^n = \sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k},$$

wobei man hier die Vertauschbarkeit $AB = BA$ braucht. Es folgt:

$$\begin{aligned} e^{A+B} &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k} = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{1}{k!} A^k \frac{1}{(n-k)!} B^{n-k} \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} A^k \sum_{l=0}^{\infty} \frac{1}{l!} B^l = e^A e^B, \end{aligned}$$

wobei die Umordnung der Reihen nach dem großen Umordnungssatz (Version für summierbare Summanden), angewandt auf die einzelnen Matrixeinträge, zulässig ist, da wir die Majorante[9]

$$\begin{aligned} &\sum_{n=0}^{\infty} \sum_{k=0}^{n} \left\| \frac{1}{k!} A^k \frac{1}{(n-k)!} B^{n-k} \right\| \le \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{1}{k!} \|A\|^k \frac{1}{(n-k)!} \|B\|^{n-k} \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \|A\|^k \sum_{l=0}^{\infty} \frac{1}{l!} \|B\|^l = e^{\|A\|} e^{\|B\|} < \infty \end{aligned}$$

haben; hier ist die Umordnung der Reihen nach dem großen Umordnungssatz (Version für nichtnegative Summanden) erlaubt.

□

Als Beispiel berechnen wir e^{xJ}, wenn J ein "Jordan-Kästchen" der Gestalt

$$J = \lambda \operatorname{Id} + N = \begin{pmatrix} \lambda & 1 & 0 & 0 & \dots & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & 0 & \dots & \lambda \end{pmatrix} \in \mathbb{C}^{n \times n}$$

mit der nilpotenten Matrix N von oben und $\lambda \in \mathbb{C}$ ist. Wegen $\operatorname{Id} N = N = N \operatorname{Id}$ folgt

$$e^{xJ} = e^{x\lambda \operatorname{Id} + xN} = e^{x\lambda \operatorname{Id}} e^{xN} = e^{x\lambda} e^{xN},$$

[9] Wir verwenden hier eine Matrixnorm, die alle absolut genommenen Einträge der Matrix nach oben beschränkt, z.B. die Matrixnorm zur euklidischen Norm.

denn $e^{x\lambda \operatorname{Id}} = e^{x\lambda} \operatorname{Id}$. Hierbei wird e^{xN} durch Gleichung (24) gegeben.

Der Fall $\lambda = -1$, der zum Differentialgleichungssystem

$$
\begin{aligned}
y_i' &= y_{i+1} - y_i \text{ für } i = 1, \ldots, n-1 \\
y_n' &= -y_n
\end{aligned}
$$

gehört, besitzt eine schöne **anschauliche Interpretation**:

In einem Reaktionsgefäß befinden sich n chemische Substanzen, indiziert mit $i = 1, \ldots, n$. In einem (kleinen) Zeitintervall $[x, x+t]$, $t > 0$, wandelt sich ein Teil $y_{i+1}(x)t + o(t)$ für $t \to 0$ der $(i+1)$-ten Substanz in die i-te Substanz um, $i = 0, 1, \ldots, n$, wobei die 0-te Substanz ein nicht mehr betrachtetes Abbauprodukt ist:

$$
n \to (n-1) \to (n-2) \to \ldots \to 2 \to 1 \to 0
$$

Der Term y_{i+1} in $y_i' = y_{i+1} - y_i$ modelliert also den *zufließenden* Strom von der $(i+1)$-ten Substanz zur i-ten Substanz, während der Term $-y_i$ den *wegfließenden* Strom von der i-ten Substanz zur $(i-1)$-ten Substanz modelliert. Betrachten wir nun die Anfangsbedingung $y_n(0) = 1$, $y_i(0) = 0$ für alle anderen i, die besagt, dass zu Beginn nur eine Einheit der Ausgangssubstanz, aber noch keine Reaktionsprodukte vorhanden sind. Wir wenden nun Satz 1.148 mit $b = (0, \ldots, 0, 1)^t$ und $A = -\operatorname{Id} + N$ an und erhalten aus der letzten Spalte in (24) mal e^{-x}: Der Anteil $e^{-x}\frac{x^k}{k!}$ der Ausgangssubstanz hat zur Zeit x genau k Reaktionsschritte durchlaufen. Anders gesagt: Die Wahrscheinlichkeit, dass ein zur Zeit 0 vorgegebenes Molekül der Ausgangssubstanz zur Zeit $x > 0$ genau k Reaktionen hinter sich hat, beträgt $e^{-x}\frac{x^k}{k!}$. Die zugehörige zufällige Dynamik heißt *Poissonprozess* und spielt eine wichtige Rolle bei der Modellierung zufälliger Vorgänge.

Besitzt eine Matrix $A \in \mathbb{C}^{n \times n}$ Block-Diagonalgestalt

$$
A = \left(\begin{array}{c|c|c|c}
J_1 & 0 & \ldots & 0 \\
\hline
0 & J_2 & \ldots & 0 \\
\hline
\vdots & \vdots & \ddots & \vdots \\
\hline
0 & 0 & \ldots & J_j
\end{array}\right)
$$

mit Diagonal-Blöcken $J_i \in \mathbb{C}^{n_i \times n_i}$ für $i = 1, \ldots, j$, wobei $\sum_{i=1}^{j} n_i = n$, so gilt

$$
e^{xA} = \left(\begin{array}{c|c|c|c}
e^{xJ_1} & 0 & \ldots & 0 \\
\hline
0 & e^{xJ_2} & \ldots & 0 \\
\hline
\vdots & \vdots & \ddots & \vdots \\
\hline
0 & 0 & \ldots & e^{xJ_j}
\end{array}\right),
$$

wie mit Hilfe der Matrix-Exponentialreihe und

$$A^k = \left(\begin{array}{c|c|c|c} J_1^k & 0 & \dots & 0 \\ \hline 0 & J_2^k & \dots & 0 \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline 0 & 0 & \dots & J_j^k \end{array}\right) \quad \text{für } k \in \mathbb{N}_0$$

unmittelbar folgt. Insbesondere können wir damit e^{xA} berechnen, wenn A in Jordan-Normalform ist, d.h. wenn alle Diagonalblöcke J_i Jordan-Kästchen sind. Der Fall, dass $A = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ eine Diagonalmatrix ist, ist der einfachste (und wichtigste) Spezialfall: In diesem Fall sind alle Jordan-Kästchen 1×1-Matrizen (λ_i).
In der Linearen Algebra lernen Sie, dass jede Matrix $A \in \mathbb{C}^{n\times n}$ in Jordan-Normalform gebracht werden kann:

$$A = TBT^{-1}$$

mit einer invertierbaren Matrix $T \in \mathbb{C}^{n\times n}$ und einer Matrix $B \in \mathbb{C}^{n\times n}$ in Jordan-Normalform.
Im wichtigsten Spezialfall, dass A diagonalisierbar ist, ist B eine Diagonalmatrix mit den Eigenwerten $\lambda_1, \dots, \lambda_n$ von A in der Diagonalen und T eine Matrix mit einer Basis von $\mathbb{C}^n$ aus Eigenvektoren von A in den Spalten, aufgezählt in der zu $\lambda_1, \dots, \lambda_n$ passenden Reihenfolge.
Nun gilt allgemein:

Lemma 1.151 *Für alle Matrizen $B, T \in \mathbb{C}^{n\times n}$, $n \in \mathbb{N}$, so dass T invertierbar ist, gilt für $A = TBT^{-1}$:*

$$e^A = Te^BT^{-1}$$

Beweis: Aus

$$A^k = TB^kT^{-1} \text{ für } k \in \mathbb{N}_0$$

folgt zunächst

$$\sum_{k=0}^{n} \frac{1}{k!} A^k = T \left(\sum_{k=0}^{n} \frac{1}{k!} B^k \right) T^{-1}$$

und damit durch Limesbildung (mit der Stetigkeit der Matrixmultiplikation)

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k = T \left(\sum_{k=0}^{\infty} \frac{1}{k!} B^k \right) T^{-1} = TB^kT^{-1}.$$

Insbesondere können wir damit e^{xA} für $x \in \mathbb{R}$ berechnen, wenn eine Transformation $A = TBT^{-1}$ mit B in Jordan-Normalform bekannt ist.

Beispiel 1.152 (nochmal Schwingungsgleichung) Betrachten wir die Schwingungsgleichung

$$y'(x) = Ay(x), \quad y(0) = b$$

mit

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

unter diesem Aspekt: Die Matrix A besitzt die Eigenwerte i und $-i$ mit den zugehörigen Eigenvektoren $\binom{1}{i}$ und $\binom{1}{-i}$. Die aus den Eigenvektoren in den Spalten gebildete Matrix

$$T = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$$

besitzt die Inverse

$$T^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}.$$

Damit wird A wie folgt diagonalisiert:

$$A = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\frac{1}{2}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}$$

Wir erhalten für $x \in \mathbb{R}$ wieder die bekannte Drehmatrix:

$$\begin{aligned} e^{xA} &= \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}\begin{pmatrix} e^{ix} & 0 \\ 0 & e^{-ix} \end{pmatrix}\frac{1}{2}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} e^{ix}+e^{-ix} & ie^{-ix}-ie^{ix} \\ ie^{ix}-ie^{-ix} & e^{ix}+e^{-ix} \end{pmatrix} = \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix}. \end{aligned}$$

Übung 1.153 *Diese Aufgabe setzt Kenntnisse aus der Linearen Algebra zur Berechnung einer Diagonalisierung bzw. einer Jordan-Zerlegung voraus.*
Berechnen Sie die Lösung des Anfangswertproblems

$$y'(x) = Ay(x), \quad y(0) = b \in \mathbb{R}^n$$

für folgende Fälle:

a) $n = 2$, $A = \begin{pmatrix} -2 & 2 \\ -6 & 5 \end{pmatrix}$,

b) $n = 3$, $A = \begin{pmatrix} -1 & -2 & 4 \\ 0 & 2 & 0 \\ -1 & -1 & 3 \end{pmatrix}$.

Übung 1.154 Es seien $n \in \mathbb{N}$, $b \in \mathbb{C}^n$ und $A \in \mathbb{C}^{n\times n}$ eine hermitesche Matrix, d.h. $A^* = A$, wobei A^* das hermitesch Konjungierte von A, also das komplex Konjugierte der Transponierten von A bezeichnet. Weiter sei $y : \mathbb{R} \to \mathbb{C}^n$ die Lösung des Anfangswertproblems $y'(x) = iAy(x)$, $y(0) = b$. Zeigen Sie $\|y(x)\|_2 = \|b\|_2$ für alle $x \in \mathbb{R}$, indem Sie die Ableitung $\frac{d}{dx}\|y(x)\|_2^2$ berechnen.

Übung 1.155 (gedämpfte Schwingungen) Eine punktförmige Masse hängt an einer Feder (Federkonstante 1). Zusätzlich zur Federkraft, die die Masse zur Gleichgewichtslage zurücktreibt, wirkt noch eine bremsende Reibungskraft proportional zur Geschwindigkeit. Wir beschreiben dieses Modell durch die Differentialgleichung 2. Ordnung

$$y_1''(t) = -y_1(t) - \mu y_1'(t)$$

oder äquivalent durch das System

$$\begin{aligned} y_1'(t) &= y_2(t), \\ y_2'(t) &= -y_1(t) - \mu y_2(t) \end{aligned}$$

mit einem "Reibungskoeffizienten" $\mu \geq 0$. Berechnen Sie die Lösung dieses Differentialgleichungssystems für gegebene Anfangsbedingung $y_1(0) = b_1$, $y_2(0) = b_2$. Für welche Werte von μ ist die zum Differentalgleichungssystem gehörende Matrix diagonalisierbar? Für welchen Wert von μ braucht man ein 2×2-Jordankästchen?

1.11 Häufungspunkte

Im Folgenden sei $(X, \mathcal{T})$ ein topologischer Raum, $x \in X$ und $(a_n)_{n\in\mathbb{N}}$ eine Folge mit Werten in X. Wie in der Analysis 1 definieren wir:

Definition 1.156 (Häufungspunkte) *Der Punkt $x \in X$ wird Häufungspunkt einer Folge $(a_n)_{n\in\mathbb{N}}$ genannt, wenn für jede offene Umgebung U von x gilt:*

$$\forall m \in \mathbb{N}\ \exists n \geq m : \ a_n \in U,$$

d.h. wenn jede Umgebung von x unendlich viele Folgenglieder enthält.

Lemma 1.157 (Charakterisierung von Häufungspunkten) *Ist d eine Halbmetrik auf X mit $\mathcal{T} = \mathcal{T}_d$, so sind äquivalent:*

1. *Der Punkt x ist ein Häufungspunkt der Folge $(a_n)_{n\in\mathbb{N}}$.*

2. *Die Folge $(a_n)_{n\in\mathbb{N}}$ besitzt eine gegen x konvergente Teilfolge.*

Beweis: "1. $\Rightarrow$ 2." Wir wählen rekursiv für alle $k \in \mathbb{N}$ ein $n_k \in \mathbb{N}$ aus, so dass gilt:

- $n_k > n_l$ für alle $l \in \mathbb{N}$ mit $k > l$.

- $a_{n_k} \in U^d_{1/k}(x)$.

Eine solche Wahl ist möglich, da $U^d_{1/k}(x)$ unendlich viele der Folgenglieder enthält. Dann ist $(a_{n_k})_{k\in\mathbb{N}}$ offensichtlich eine gegen x konvergente Teilfolge von $(a_n)_{n\in\mathbb{N}}$.

"2. $\Rightarrow$ 1." Die Folge $(a_n)_{n\in\mathbb{N}}$ besitze eine gegen x konvergente Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$. Dann können wir zu jeder offenen Umgebung U von x ein $k_0 \in \mathbb{N}$ wählen, so dass für alle $k \geq k_0$ gilt: $a_{n_k} \in U$. Ist nun $m \in \mathbb{N}$ gegeben, so wählen wir $k \in \mathbb{N}$ so groß, dass $k \geq k_0$ und $n := n_k \geq m$ gilt. Es folgt $a_n \in U$. Der Punkt x ist also ein Häufungspunkt der Folge $(a_n)_{n\in\mathbb{N}}$.

□

Man beachte, dass der Beweis von “1. $\Rightarrow$ 2.” die Halbmetrik d verwendet, der Beweis von “2. $\Rightarrow$ 1.” jedoch nicht. Die Implikation “2. $\Rightarrow$ 1.” gilt daher auch in beliebigen topologischen Räumen. Die Aussage “1. $\Rightarrow$ 2.” ist allerdings in manchen topologischen Räumen falsch.

1.12 Kompaktheit

Erinnern Sie sich an den Kompaktheitsbegriff aus der Analysis 1:

Definition 1.158 (Kompaktheit und Folgenkompaktheit) Gegeben sei eine Teilmenge $K \subseteq X$ eines topologischen Raums $(X, \mathcal{T})$.

1. K heißt *kompakt* (in $(X, \mathcal{T})$), wenn jede Überdeckung $(U_i)_{i\in I}$ von X mit offenen Mengen $U_i \in \mathcal{T}$ eine endliche Teilüberdeckung $(U_i)_{i\in E}$, $E \subseteq I$ endlich, besitzt.

2. K heißt *folgenkompakt* (in $(X, \mathcal{T})$), wenn jede Folge $(x_n)_{n\in\mathbb{N}}$ mit Werten in K eine konvergente Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ mit einem Limes $a \in K$ besitzt.

Kompaktheit und Folgenkompaktheit in allgemeinen topologischen Räumen sind *nicht* äquivalent: Es gibt kompakte topologische Räume, die nicht folgenkompakt sind, und folgenkompakte topologische Räume, die nicht kompakt sind.
Aus der Analysis 1 ist bekannt: Eine Menge $K \subseteq \mathbb{R}$ reeller Zahlen ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist. Diese Charakterisierung lässt sich zwar auf $\mathbb{R}^n$ und auf $\mathbb{C}^n$ verallgemeinern, wie wir unten sehen werden, nicht jedoch auf allgemeine metrische Räume, wie das folgende Beispiel zeigt:

Übung 1.159 (Beispiel einer abgeschlossenen und beschränkten, aber nicht kompakten Menge) Es sei $K = \{f \in \ell^2(\mathbb{N}, \mathbb{R})|\ \|f\|_2 \le 1\}$ der abgeschlossene Einheitsball in $\ell^2(\mathbb{N})$. Zeigen Sie, dass K nicht kompakt ist. Zeigen Sie dazu, dass die offene Überdeckung $(U_{1/\sqrt{2}}(x))_{x\in K}$ von K mit den offenen Kugeln $U_{1/\sqrt{2}}(x) = \{y \in \ell^2(\mathbb{N})|\ \|y - x\|_2 < 1/\sqrt{2}\}$ keine endliche Teilüberdeckung besitzt, indem Sie zeigen, dass keine der Mengen $U_{1/\sqrt{2}}(x)$ mindestens zwei “kanonische Einheitsvektoren” e_n, $n \in \mathbb{N}$, enthält, wobei $e_n(j) = \delta_{n,j}$ (Kronecker-Delta).

Als einen “Ersatz” für die Beschränktheit führen wir den Begriff der *Totalbeschränktheit* ein:

Definition 1.160 (Totalbeschränktheit) Eine Teilmenge $N \subseteq M$ eines halbmetrischen Raums (M, d) heißt *totalbeschränkt*, wenn es für jedes $\epsilon > 0$ eine endliche Teilmenge $E \subseteq M$ gibt, so dass gilt:

$$N \subseteq \bigcup_{x\in E} U_\epsilon^d(x).$$

Anschaulich gesprochen besagt das, dass man die Menge N mit endlich vielen Kugeln mit beliebig kleinem Radius überdecken kann.

Übung 1.161 Zeigen Sie, dass der Begriff der Totalbeschränktheit sich nicht ändert, wenn man in der Definition statt "$E \subseteq M$" fordert: "$E \subseteq N$".

Übung 1.162 (Totalbeschränktheit von $\mathbb{Z}$ in der p-adischen Metrik) Es sei p eine Primzahl. Zeigen Sie, dass $\mathbb{Z}$ bezüglich der p-adischen Metrik aus Beispiel 1.2.3 totalbeschränkt ist.

Übung 1.163 (Totalbeschränktheit impliziert Beschränktheit) Zeigen Sie, dass jede totalbeschränkte Menge in einem halbmetrischen Raum auch beschränkt ist.

Über $\mathbb{R}^n$ gilt auch die Umkehrung hiervon:

Beispiel 1.164 (Beschränktheit in $\mathbb{R}^n$ impliziert Totalbeschränktheit) *In $(\mathbb{R}^n, \|\cdot\|_\infty)$ gilt: Jede beschränkte Teilmenge $N \subset \mathbb{R}^n$ in $(\mathbb{R}^n, \|\cdot\|_\infty)$, $n \in \mathbb{N}$, ist totalbeschränkt.*

Begründung: Wegen der Beschränktheit gibt es ein $m \in \mathbb{N}$ mit $N \subseteq [-m, m]^n$. Gegeben $\epsilon > 0$, setzen wir $E = ([-m, m] \cap \epsilon\mathbb{Z})^n$. Dann ist E endlich, und es gilt bezüglich der von $\|\cdot\|_\infty$ erzeugten Metrik d:

$$N \subseteq \bigcup_{x \in E} U_\epsilon^d(x).$$

Also ist N totalbeschränkt.

Übung 1.165 Zeigen Sie analog, dass jede beschränkte Teilmenge von $\mathbb{C}^n$ totalbeschränkt ist, wobei $\mathbb{C}^n$ mit der von $\|\cdot\|_\infty$ induzierten Metrik versehen wird.

Aus der Analysis 1 wissen Sie, dass folgende drei Aussagen für eine Teilmenge $K \subseteq \mathbb{R}$ äquivalent sind: 1. K ist kompakt; 2. K ist folgenkompakt; 3. K ist abgeschlossen und beschränkt. Zwar kann diese Aussage nicht wörtlich auf allgemeine halbmetrische Räume verallgemeinert werden, doch folgende Variante davon gilt allgemein:

Satz 1.166 (Metrische Charakterisierungen der Kompaktheit) *Ist $K \subseteq X$ in einem halbmetrischen Raum (X, d), so sind bezüglich $\mathcal{T}_d$ äquivalent:*

1. *K ist kompakt,*

2. *K ist folgenkompakt.*

3. *K ist vollständig und totalbeschränkt.*

Beweis: "1. $\Rightarrow$ 2.": Es sei K kompakt und $(x_n)_{n\in\mathbb{N}}$ eine Folge mit Werten in K. Wir zeigen indirekt, dass diese Folge mindestens einen Häufungspunkt in K besitzt. Angenommen, das wäre nicht der Fall. Dann können wir für jedes $x \in K$ eine offene Umgebung U_x auswählen, die nur endlich viele der Folgenglieder enthält. Wegen der Kompaktheit von K besitzt die offene Überdeckung $(U_x)_{x\in K}$ von K eine endliche Teilüberdeckung $(U_x)_{x\in E}$, $E \subseteq K$ endlich. Dann enthält auch die endliche Vereinigung $\bigcup_{x\in E} U_x$ nur endlich viele der Folgenglieder von $(x_n)_{n\in\mathbb{N}}$, im Widerspruch zu $x_n \in K \subseteq \bigcup_{x\in E} U_x$ für alle $n \in \mathbb{N}$.

Ist nun $x \in K$ ein Häufungspunkt von $(x_n)_{n\in\mathbb{N}}$, so gibt es eine gegen x konvergente Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ nach Lemma 1.157.

"2. $\Rightarrow$ 3." Es sei K folgenkompakt. Wir zeigen zunächst, das K vollständig ist. Hierzu sei $(a_n)_{n\in\mathbb{N}}$ eine Cauchyfolge mit Werten in K. Wegen der Folgenkompaktheit von K besitzt sie eine gegen ein $x \in K$ konvergente Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$. Aus der Übung 1.99 folgt, dass dann auch die Folge $(a_n)_{n\in\mathbb{N}}$ gegen x konvergiert. Also ist K vollständig.

Nun zeigen wir, dass K totalbeschränkt ist. Hierzu sei $\epsilon > 0$ gegeben. Wir gehen indirekt vor und nehmen daher an, dass für jede endliche Menge $E \subseteq K$ die Familie $(U_\epsilon(x))_{x\in E}$ die Menge K nicht überdeckt. Mit dieser Annahme können wir rekursiv eine Folge $(x_n)_{n\in\mathbb{N}}$ mit Werten in K wie folgt wählen: Ist $n \in \mathbb{N}$ und sind x_k für $k \in \mathbb{N}$ mit $k < n$ nach Rekursionsvoraussetzung schon gegeben, so überdeckt die endliche Familie $(U_\epsilon(x_k))_{k=1,\ldots,n-1}$ die Menge K nach unserer Annahme *nicht*. (Für $n = 1$ ist das natürlich die leere Familie.) Wir können also ein $x_n \in K$ mit

$$x_n \notin \bigcup_{k=1}^{n-1} U_\epsilon(x_k)$$

auswählen. Insbesondere gilt $d(x_k, x_n) \geq \epsilon$ für alle $k < n$. Die so rekursiv definierte Folge $(x_n)_{n\in\mathbb{N}}$ besitzt keine Cauchyfolge als Teilfolge, da je zwei verschiedene Folgenglieder mindestens den Abstand ϵ besitzen. Also besitzt sie erst recht keine konvergente Teilfolge, da jede konvergente Teilfolge auch eine Cauchyfolge wäre. Dies widerspricht der angenommenen Folgenkompaktheit von K.

"3. $\Rightarrow$ 1." Es sei K vollständig und totalbeschränkt und $(V_j)_{j\in I}$ eine offene Überdeckung von K. Wir gehen wieder indirekt vor. Nehmen wir also an, dass es keine endliche Teilüberdeckung zur Überdeckung $(V_j)_{j\in I}$ von K gibt. Wir konstruieren nun rekursiv eine Folge $(x_n)_{n\in\mathbb{N}}$ mit Werten in K, verwenden die Abkürzung

$$K_n := \{y \in K \,|\, \forall m \in \mathbb{N}, m < n : \; d(y, x_m) < 1/m\} = \begin{cases} \bigcup_{m=1}^{n-1} U^d_{1/m}(x_m) \cap K & \text{für } n > 1, \\ K & \text{für } n = 1 \end{cases}$$

und beweisen simultan durch Induktion für alle $n \in \mathbb{N}$ die folgende

Aussage $\Phi(n)$*:* Für jede endliche Teilmenge E von I überdeckt $(V_j)_{j\in E}$ die Menge K_n nicht, und es gilt $x_m \in K_m$ für alle $m \in \mathbb{N}$ mit $m < n$.

Man beachte, dass die Definition von K_n und die Aussage $\Phi(n)$ nur die x_m mit $m < n$ verwenden.

Induktionsanfang: Die Aussage $\Phi(1)$ ist nach unserer obigen Annahme richtig, denn $K_1 = K$.

Nun sei $n \in \mathbb{N}$ fixiert.

Rekursionsvoraussetzung und Induktionsvoraussetzung:
Die Folgenglieder x_m für $m \in \mathbb{N}$ mit $m < n$ seien schon gegeben, und es gelte die Aussage $\Phi(n)$.

Rekursionsschritt und Induktionsschritt:
Wir müssen x_n so definieren, dass auch die Aussage $\Phi(n+1)$ gilt. Als Teilmenge der totalbeschränkten Menge K ist K_n totalbeschränkt. Es gibt also eine endliche Teilmenge A_n von K_n, so dass $(U^d_{1/n}(x))_{x\in A_n}$ die Menge K_n überdeckt. Könnten wir für jedes $x \in A_n$ eine endliche Menge $E_x \subseteq I$ wählen, so dass $K_n \cap U^d_{1/n}(x) \subseteq \bigcup_{j\in E_x} V_j$, so folgte

$$K_n = K_n \cap \bigcup_{x\in A_n} U^d_{1/n}(x) = \bigcup_{x\in A_n} K_n \cap U^d_{1/n}(x) \subseteq \bigcup_{x\in A_n} \bigcup_{j\in E_x} V_j,$$

so dass K_n von den endlich vielen V_j mit $j \in \bigcup_{x\in A_n} E_x$ überdeckt würde, im Widerspruch zur in der Induktionsvoraussetzung angenommenen Aussage $\Phi(n)$. Wir können also ein $x_n \in A_n$ auswählen, so dass $K_{n+1} = K_n \cap U^d_{1/n}(x_n)$ *nicht* von endlich vielen V_j, $j \in I$, überdeckt wird. Wegen $A_n \subseteq K_n$ folgt $x_n \in K_n \subseteq K$. Mit dieser rekursiven Wahl von x_n ist also auch die Aussage $\Phi(n+1)$ erfüllt.

Dies beendet die rekursive Definition der Folge $(x_n)_{n\in\mathbb{N}}$ und den induktiven Beweis der Aussage $\Phi(n)$.

Weil $x_n \in K_n$ für alle $n \in \mathbb{N}$ gilt, wissen wir insbesondere $\forall n, k \in \mathbb{N}, k < n: \ d(x_n, x_k) < 1/k$, woraus folgt, dass $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge ist. Weil K nach Voraussetzung vollständig ist, konvergiert die Cauchyfolge $(x_n)_{n\in\mathbb{N}}$ in K gegen ein $z \in K$. Nun gibt es ein $i \in I$ mit $z \in V_i$, da $(V_j)_{j\in I}$ die Menge K überdeckt. Da V_i offen ist, gibt es ein $\epsilon > 0$ mit $U^d_\epsilon(z) \subseteq V_i$. Wählen wir nun $n \in \mathbb{N}$ so groß, dass $2/n < \epsilon$ gilt. Wegen der Stetigkeit der Metrik und wegen $x_m \xrightarrow{m\to\infty} z$ folgt $d(x_n, z) = \lim_{m\to\infty} d(x_n, x_m) \le 1/n$, da $d(x_n, x_m) < 1/n$ für alle $m \in \mathbb{N}$ mit $m > n$ gilt. Für alle $y \in K_{n+1}$ erhalten wir $d(y, x_n) < 1/n$ nach der Definition von K_{n+1}, also

$$d(y, z) \le d(y, x_n) + d(x_n, z) < \frac{1}{n} + \frac{1}{n} < \epsilon.$$

Es folgt $K_{n+1} \subseteq U^d_\epsilon(z) \subseteq V_i$. Die Menge K_{n+1} wird also schon von *einem einzigen* V_i überdeckt, im Widerspruch zur oben gezeigten Aussage $\Phi(n+1)$.

Damit ist indirekt gezeigt, dass K kompakt ist.

$\square$

Man beachte, dass alle Teile des eben dargestellten Beweises metrische Begriffe verwenden. Das Lemma lässt sich nicht auf allgemeine topologische Räume übertragen. In der Tat gibt es kompakte topologische Räume, die nicht folgenkompakt sind, und folgenkompakte topologische Räume, die nicht kompakt sind.

Korollar 1.167 (Charakterisierung der Kompaktheit in $\mathbb{R}^n$) *Es seien $n \in \mathbb{N}$ und $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Eine Teilmenge K von $\mathbb{K}^n$ ist genau dann kompakt, wenn sie abgeschlossen und beschränkt bezüglich der von $\|\cdot\|_\infty$ erzeugten Metrik d ist.*

Beweis: Weil $\mathbb{K}^n$ vollständig und d eine Metrik ist, ist die Abgeschlossenheit von K äquivalent zur Vollständigkeit nach Lemma 1.106. Nach den Übungen 1.163 und 1.165 und dem Beispiel 1.164 ist die Beschränktheit von K äquivalent zur Totalbeschränktheit. Damit folgt die Behauptung aus dem Satz 1.166.

$\square$

Beispiel 1.168 *Die "Einheitssphäre" $A = \{x \in \mathbb{K}^n | \; \|x\|_\infty = 1\}$ bezüglich $\|\cdot\|_\infty$ ist kompakt in $(\mathbb{R}^n, \|\cdot\|_\infty)$, denn sie ist abgeschlossen als Urbild der abgeschlossenen Menge $\{1\}$ unter der Normabbildung, die stetig ist. Auch ist sie wegen $A \subseteq [-1,1]^n$ beschränkt. Also ist A kompakt.*

Erinnern Sie sich an folgende wichtige Eigenschaften stetiger Funktionen auf kompakten Mengen aus der Analysis 1:

- Bilder kompakter Mengen unter stetigen Abbildungen sind kompakt.
- Stetige Funktionen mit Werten in $\mathbb{R}$ auf nichtleeren, kompakten Mengen nehmen ihr Minimum und ihr Maximum als Wert an.

Erinnern Sie sich, dass zwei Normen auf dem gleichen $\mathbb{K}$-Vektorraum äquivalent genannt werden, wenn sie die gleiche Topologie erzeugen. Das ist genau dann der Fall, wenn die eine Norm durch konstante Vielfache der anderen Norm nach unten und nach oben abgeschätzt werden kann; siehe Übung 1.64.

Satz 1.169 (Äquivalenz aller Normen auf $\mathbb{R}^n$) *Für $n \in \mathbb{N}$ und $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ gilt: Alle Normen auf $\mathbb{K}^n$ sind äquivalent.*

Beweis Es sei $\|\cdot\| : \mathbb{K}^n \to \mathbb{R}$ eine Norm. Wir zeigen die Äquivalenz der Normen $\|\cdot\|$ und $\|\cdot\|_\infty$. Mit Lemma 1.7 folgt für $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$:

$$\|x\| = \left\| \sum_{j=1}^n x_j e_j \right\| \leq \|x\|_\infty \sum_{j=1}^n \|e_j\|,$$

wobei e_j den j-ten kanonischen Einheitsvektor in $\mathbb{K}^n$ bezeichnet. Aus Übung 1.64 folgt damit: Die Identität $\mathrm{id} : (\mathbb{K}, \|\cdot\|_\infty) \to (\mathbb{K}^n, \|\cdot\|)$, $\mathrm{id}(x) = x$, ist stetig. Mit der Stetigkeit der Normabbildung $\|\cdot\| : (\mathbb{K}^n, \|\cdot\|) \to (\mathbb{R}, |\cdot|)$ nach Lemma 1.59 folgt durch Komposition: Die Normabbildung $\|\cdot\| : (\mathbb{K}^n, \|\cdot\|_\infty) \to (\mathbb{R}, |\cdot|)$. ist stetig. Es bezeichne $A = \{x \in \mathbb{K}^n | \; \|x\|_\infty = 1\}$ die "Einheitssphäre" bzgl. $\|\cdot\|_\infty$. Wegen $0 \notin A$ gilt $\|x\| > 0$ für alle $x \in A$. Weil die "Einheitssphäre" $A = \{x \in \mathbb{K}^n | \; \|x\|_\infty = 1\}$ nichtleer wegen $e_1 \in A$ und kompakt nach Beispiel 1.168 ist, nimmt die Norm $\|\cdot\|$ auf A ein Minimum $m > 0$ und ein Maximum $M < \infty$ an. Hieraus folgt

$$\forall x \in \mathbb{K}^n : \ m\|x\|_\infty \leq \|x\| \leq M\|x\|_\infty. \tag{25}$$

Dies sieht man so: Es sei $x \in K^n$. Im Fall $x = 0$ ist die Behauptung klar. Wir dürfen also $x \neq 0$ und damit $\|x\|_\infty > 0$ annehmen. Setzen wir $y := x/\|x\|_\infty$, so folgt $\|y\|_\infty = \|x\|_\infty/\|x\|_\infty = 1$, also $y \in A$ und damit $m \leq \|y\| \leq M$. Wegen $\|y\| = \|x\|/\|x\|_\infty$ folgt die Behauptung 25, also nach Übung 1.64 die Äquivalenz der Normen $\|\cdot\|$ und $\|\cdot\|_\infty$.

□

Übung 1.170 (Kompaktheit der euklidischen Einheitssphäre) *1. Für $n \in \mathbb{N}$ sei $S^n := \{x \in \mathbb{R}^n |\ \|x\|_2 = 1\}$ die euklidische Einheitssphäre. Zeigen Sie: S^n ist kompakt in der Standardtopologie auf $\mathbb{R}^n$.*

2. Folgern Sie, dass für jede stetige Funktion $f : \mathbb{R}^3 \to \mathbb{R}$ die Menge

$$M = \{f(x, y, z)|\ x, y, z \in \mathbb{R},\ x^2 + y^2 + z^2 = 1\}$$

ein Minimum und ein Maximum besitzt.

Übung 1.171 (Stetigkeit linearer Abbildungen auf $\mathbb{R}^m$) *Es seien $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, $m, n \in \mathbb{N}$, $\|\cdot\|$ eine beliebige Norm auf $\mathbb{K}^m$ und $\|\cdot\|'$ eine beliebige Norm auf $\mathbb{K}^n$. Weiter sei $L : \mathbb{K}^m \to \mathbb{K}^n$ eine $\mathbb{K}$-lineare Abbildung. Zeigen Sie, dass $L : (\mathbb{K}^m, \|\cdot\|) \to (\mathbb{K}^n, \|\cdot\|')$ stetig ist.*

Übung 1.172 (Charakterisierung kompakter Einheitskugeln) Es sei $(V, \|\cdot\|)$ ein normierter Raum über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Zeigen Sie die Äquivalenz der folgenden beiden Aussagen:

1. V ist endlichdimensional.
2. Die abgeschlossene Einheitskugel $B_1(0) = \{x \in V|\ \|x\| \leq 1\}$ ist kompakt.

Übung 1.173 Zeigen Sie: Ist (M, d_N) ein totalbeschränkter halbmetrischer Raum und $f : (M, d_M) \to (N, d_N)$ eine gleichmäßig stetige Abbildung mit Werten in einem halbmetrischen Raum (N, d_N), so ist auch das Bild $f[M]$ in (N, d_N) totalbeschränkt.

Übung 1.174 Zeigen Sie:

1. Der Abschluss $\overline{A}$ einer totalbeschränkten Teilmenge $A \subseteq M$ in einem halbmetrischen Raum (M, d) ist totalbeschränkt.
2. Die Vervollständigung eines totalbeschränkten halbmetrischen Raums ist kompakt.

Aus dieser Übung zusammen mit Übung 1.162 folgt: Für beliebige Primzahlen ist die Vervollständigung $\mathbb{Z}_p$ der ganzen Zahlen bezüglich der p-adischen Metrik kompakt.

Übung 1.175 (Stetige Funktionen auf kompakten Mengen sind gleichmäßig stetig.) Es seien (M, d_M) ein kompakter halbmetrischer Raum, (N, d_N) ein halbmetrischer Raum und $f : (M, d_M) \to (N, d_N)$ eine stetige Abbildung. Beweisen Sie, dass f gleichmäßig stetig ist. Lassen Sie sich dazu von dem aus der Analysis 1 bekannten Spezialfall $M, N \subseteq \mathbb{R}$ inspirieren.

Übung 1.176 Es sei

$$M = \{f \in C([0,1], \mathbb{R}) |\ \|f\|_\infty \leq 1, \forall x, y \in [0,1] :\ |f(x) - f(y)| \leq |x - y|\}$$

der Raum aller durch 1 beschränkten, global Lipschitz-stetigen Funktionen mit Lipschitzkonstante 1. Zeigen Sie, dass M in $(C([0,1], \mathbb{R}), \|\cdot\|_\infty)$ abgeschlossen und totalbeschränkt und damit kompakt ist.

Übung 1.177 (Abgeschlossenheit in kompakten Hausdorffräumen.) *Es sei $(X, \mathcal{T})$ ein kompakter Hausdorffraum und $K \subseteq X$ eine Teilmenge. Zeigen Sie, dass die Menge K genau dann kompakt ist, wenn sie abgeschlossen ist. Hinweis zur Richtung "$\Rightarrow$": Gegeben $x \in X \setminus K$, wählen Sie für jedes $y \in K$ offene Umgebungen U_y von y und V_y von x mit $U_y \cap V_y = \emptyset$ aus. Betrachten Sie die offene Überdeckung $(U_y)_{y \in K}$ von K und verwenden Sie die Kompaktheit von K.*

Hier ist ein einfaches hinreichendes Kriterium zum Nachweis, dass eine Abbildung $f : (X, \mathcal{T}) \to (Y, \mathcal{S})$ final ist:

Lemma 1.178 (Finaltopologie und Kompaktheit) *Jede surjektive stetige Abbildung $f : (X, \mathcal{T}) \to (Y, \mathcal{S})$ von einem kompakten topologischen Raum $(X, \mathcal{T})$ auf einen Hausdorffraum $(Y, \mathcal{S})$ ist final.*

Beweis: Wegen der Stetigkeit von f gilt

$$\mathcal{S} \subseteq \{U \subseteq Y |\ f^{-1}[U] \in \mathcal{T}\}.$$

Zum Nachweis der umgekehrten Inklusion

$$\mathcal{S} \supseteq \{U \subseteq Y |\ f^{-1}[U] \in \mathcal{T}\}$$

sei $U \subseteq Y$ mit $f^{-1}[U] \in \mathcal{T}$ gegeben. Dann ist $f^{-1}[Y \setminus U] = X \setminus f^{-1}[U]$ in $(X, \mathcal{T})$ abgeschlossen, also kompakt, da $(X, \mathcal{T})$ kompakt ist. Da f stetig ist, ist auch das Bild $f[f^{-1}[Y \setminus U]]$ kompakt, also abgeschlossen bzgl. $\mathcal{S}$, da $(Y, \mathcal{S})$ ein Hausdorffraum ist. Nun gilt $f[f^{-1}[Y \setminus U]] = Y \setminus U$, da f surjektiv ist. Aus der Abgeschlossenheit von $Y \setminus U$ folgt: U ist offen, d.h. $U \in \mathcal{S}$.

□

Übung 1.179 Es seien (M_i, d_i), $i = 1, \ldots, n$ totalbeschränkte halbmetrische Räume und (M, d) das kartesische Produkt davon, versehen mit der Produkthalbmetrik. Zeigen Sie, dass auch (M, d) totalbeschränkt ist. Folgern Sie, dass (M, d) kompakt ist, wenn alle (M_i, d_i) kompakt sind.

Ausblick: Allgemeiner gilt der **Satz von Tychonoff:** Ein beliebiges kartesisches Produkt kompakter topologischer Räume, versehen mit der Produkttopologie, ist kompakt. Wir beweisen das hier nicht; der Beweis beruht wesentlich auf dem Auswahlaxiom.

1.13 Der Approximationssatz von Stone-Weierstraß

Es sei $(X, \mathcal{T})$ ein nichtleerer kompakter Hausdorffraum und $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.
Wir betrachten den normierten Raum $(C(X, \mathbb{K}), \|\cdot\|_\infty)$ aller bezüglich $\mathcal{T}$ stetigen Funktionen $f : X \to \mathbb{R}$, versehen mit der Supremumsnorm

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

Man beachte, dass $\|f\|_\infty$ wegen der Kompaktheit von X endlich ist. Gegeben $A \subseteq C(X, \mathbb{K})$, besprechen wir nun ein einfaches hinreichendes Kriterium, dass A dicht in $(C(X, \mathbb{K}), \|\cdot\|_\infty)$ ist.

Definition 1.180 (punktetrennend) Eine Menge $A \subseteq C(X, \mathbb{K})$ stetiger Funktionen wird *punktetrennend* genannt, wenn es für alle $x, y \in X$ mit $x \neq y$ ein $f \in A$ mit $f(x) \neq f(y)$ gibt.

Wir betrachten zunächst den Fall $\mathbb{K} = \mathbb{R}$. Einen wichtigen Schritt liefert folgender Satz:

Satz 1.181 (Satz von Kakutani-Krein) *Gegeben seien ein nichtleerer kompakter Hausdorffraum $(X, \mathcal{T})$ und eine Menge $A \subseteq C(X, \mathbb{R})$ mit den folgenden Eigenschaften:*

1. *A ist abgeschlossen in $(C(X, \mathbb{R}), \|\cdot\|_\infty)$.*
2. *Die konstante Funktion $1 : X \to \mathbb{R}$ mit dem Wert 1 ist ein Element von A.*
3. *Für alle $f, g \in A$ ist $f + g \in A$.*
4. *Für alle $f \in A$ und alle $\alpha \in \mathbb{R}$ ist $\alpha f \in A$.*
5. *Für alle $f \in A$ ist $|f| \in A$, wobei $|f|(x) := |f(x)|$ für $x \in X$.*
6. *A ist punktetrennend.*

Dann ist $A = C(X, \mathbb{R})$.

Beweis: Für alle $f, g \in A$ ist $f - g = f + (-1)g \in A$. Wir definieren für $f, g \in C(X, \mathbb{R})$: $f \vee g, f \wedge g : X \to \mathbb{R}$, $(f \vee g)(x) := \max\{f(x), g(x)\}$ und $(f \wedge g)(x) := \min\{f(x), g(x)\}$.
Wir beobachten zunächst: Für alle $f, g \in A$ ist $f \vee g, f \wedge g \in A$, denn

$$f \vee g = \frac{1}{2}(f + g + |f - g|),$$
$$f \wedge g = \frac{1}{2}(f + g - |f - g|).$$

Nun sei $f \in C(X, \mathbb{R})$ gegeben. Es genügt zu zeigen:

$$\forall \epsilon > 0 \; \exists g \in A : \; \|f - g\|_\infty < \epsilon,$$

denn wegen der Abgeschlossenheit von A folgt dann auch $f \in A$. Hierzu sei $\epsilon > 0$ gegeben. Wir können für alle $x, y \in X$ ein $h_{x,y} \in A$ mit $h_{x,y}(x) = f(x)$ und $h_{x,y}(y) = f(y)$ wählen. In der Tat: Im Fall $x = y$ nehmen wir die konstante Funktion $h_{x,x} := f(x)1$. Im Fall $x \neq y$ wählen wir mit der Punktetrennungseigenschaft von A ein $k_{x,y} \in A$ mit $k_{x,y}(x) \neq k_{x,y}(y)$ aus und setzen

$$h_{x,y} := \frac{f(x)}{k_{x,y}(x) - k_{x,y}(y)}(k_{x,y} - k_{x,y}(y)1) + \frac{f(y)}{k_{x,y}(y) - k_{x,y}(x)}(k_{x,y} - k_{x,y}(x)1).$$

Nach den vorausgesetzten Eigenschaften 2.,3.,4. ist $h_{x,y} \in A$. Die Funktion $h_{x,y}$ leistet das Gewünschte.

Wir fixieren für den Moment ein $x \in X$. Für alle $y \in X$ ist die Menge

$$U_{x,y} := \{z \in X | \, h_{x,y}(z) < f(z) + \epsilon\} \subseteq X$$

als Urbild der offenen Menge $]-\infty, \epsilon[$ unter der stetigen Abbildung $h_{x,y} - f$ offen. Weiter gilt $y \in U_{x,y}$. Also ist $(U_{x,y})_{y \in X}$ eine offene Überdeckung von X. Wegen der Kompaktheit von X können wir eine endliche Teilmenge $E_x = \{y_1, \dots, y_n\}$, $n \in \mathbb{N}$, von X mit

$$\bigcup_{m=1}^{n} U_{x,y_m} = X$$

wählen. Natürlich kann n von x abhängen, doch x ist im Moment fixiert. Man beachte, dass E wegen $X \neq \emptyset$ nichtleer ist. Wir setzen

$$g_x := h_{x,y_1} \wedge \dots \wedge h_{x,y_n}.$$

Weil A abgeschlossen unter Minimumbildung ist, folgt auch $g_x \in A$. Weiter gilt $g_x < f + \epsilon$. In der Tat: Gegeben $z \in X$, finden wir ein $m \in \{1, \dots, n\}$ mit $z \in U_{x,y_m}$. Dann folgt

$$g_x(z) \leq h_{x,y_m}(z) < f(z) + \epsilon,$$

wobei wir bei der ersten Ungleichung die Definition von g_x und bei der zweiten Ungleichung die Definition von U_{x,y_m} verwendet haben. Weiter gilt

$$g_x(x) = \min\{h_{x,y_1}(x), \dots, h_{x,y_n}(x)\} = f(x),$$

wegen $h_{x,y_m}(x) = f(x)$ für $m = 1, \dots, n$.
Weil $g_x(x) = f(x)$ gilt, ist

$$V_x := \{z \in X | \, g_x(z) > f(z) - \epsilon\}$$

eine offene Umgebung von x, denn V_x ist das Urbild der offenen Menge $]-\epsilon, \infty[$ unter der stetigen Abbildung $g_x - f$.

Nun lassen wir $x \in X$ variieren. Die Familie $(V_x)_{x\in X}$ ist eine offene Überdeckung von X, besitzt also wegen der Kompaktheit von X eine endliche Teilüberdeckung $(V_x)_{x\in F}$ mit endlichem, nichtleerem $F = \{x_1, \ldots, x_N\} \subseteq X$. Analog zum Obigen setzen wir

$$g := g_{x_1} \vee \ldots \vee g_{x_N}.$$

Dann ist auch $g \in A$, da A abgeschlossen unter Maximumbildung ist. Weiter gilt

$$f - \epsilon < g < f + \epsilon$$

Das sieht man so: Einerseits ist $g < f + \epsilon$, da $g_{x_m} < f + \epsilon$ für alle $m \in \{1, \ldots, N\}$ gilt. Um andererseits $g > f - \epsilon$ zu zeigen, sei $z \in X$ gegeben. Dann finden wir ein $m \in \{1, \ldots, N\}$ mit $z \in V_{x_m}$, da die Familie $(V_{x_m})_{m=1,\ldots,N}$ die Menge X überdeckt. Es folgt

$$g(z) \geq g_{x_m}(z) > f(z) - \epsilon,$$

wobei wir bei der ersten Ungleichung die Definition von g und bei der zweiten Ungleichung die Definition von V_{x_m} verwendet haben.
Damit ist bewiesen: $|f(z) - g(z)| < \epsilon$ für alle $z \in X$, also auch $\|f - g\|_\infty \leq \epsilon$. Das war zu zeigen.

$\square$

Satz 1.182 (Satz von Stone-Weierstraß – reelle Version) *Gegeben seien ein nichtleerer kompakter Hausdorffraum $(X, \mathcal{T})$ und eine Menge $A \subseteq C(X, \mathbb{R})$ mit den folgenden Eigenschaften:*

1. *Die konstante Funktion $1 : X \to \mathbb{R}$ mit dem Wert 1 ist ein Element von A.*

2. *Für alle $f, g \in A$ ist $f + g \in A$.*

3. *Für alle $f \in A$ und alle $\alpha \in \mathbb{R}$ ist $\alpha f \in A$.*

4. *Für alle $f, g \in A$ ist $f \cdot g \in A$.*

5. *A ist punktetrennend.*

Dann ist A dicht in $(C(X, \mathbb{R}), \|\cdot\|_\infty)$.

Beweis: Es genügt zu zeigen, dass der Abschluss $\overline{A}$ von A in $(C(X, \mathbb{R}), \|\cdot\|_\infty)$ die Voraussetzungen des Satzes von Kakutani-Krein erfüllt, denn dann folgt die Behauptung $\overline{A} = C(X, \mathbb{R})$. Nach Voraussetzung ist $1 \in A \subseteq \overline{A}$. Nun sind die Addition $+ : C(X, \mathbb{R}) \times C(X, \mathbb{R}) \to C(X, \mathbb{R})$ und die Multiplikation $\cdot : C(X, \mathbb{R}) \times C(X, \mathbb{R}) \to C(X, \mathbb{R})$ stetig bezüglich der Supremumsnorm und der zugehörigen Produktmetrik. Im Fall der Addition folgt dies aus Lemma 1.62, und im Fall der Multiplikation aus Lemma 1.70 wegen

$$\forall f, g \in C(X, \mathbb{R}) : \ \|f \cdot g\|_\infty \leq \|f\|_\infty \|g\|_\infty.$$

Nun gilt nach der Übung 1.113 $\overline{A} \times \overline{A} = \overline{A \times A}$, also $(f,g) \in \overline{A \times A}$ für alle $f, g \in \overline{A}$. Aus Übung 1.51 folgt dann $f + g \in \overline{A + A} \subseteq \overline{A}$ und ebenso $f \cdot g \in \overline{A \cdot A} \subseteq \overline{A}$ für diese f, g. Also ist $\overline{A}$ abgeschlossen unter $+$ und $\cdot$. Wegen $\alpha 1 \in A \subseteq \overline{A}$ für alle $\alpha \in \mathbb{R}$ impliziert dies auch die Abgeschlossenheit von $\overline{A}$ unter der Multplikation mit Skalaren. Weil A punktetrennend ist, ist auch $\overline{A} \supseteq A$ punktetrennend. Gegeben $f \in A$, bleibt noch zu zeigen: $|f| \in A$. Das sieht man so: Weil $\overline{A}$ abgeschlossen unter "+", "·", und Skalarmultiplikation ist und die konstante Funktion 1 enthält, ist $p \circ f \in \overline{A}$ für jedes $f \in A$ jede Polynomfunktion $p : \mathbb{R} \to \mathbb{R}$, $p \circ f(x) = \sum_{k=0}^{n} \alpha_k f(x)^k$. Aus der Analysis 1 wissen Sie:

$$\sup_{-1 \le y \le 1} \left| |y| - \sum_{n=0}^{k} \binom{\frac{1}{2}}{n} (y^2 - 1)^n \right| \overset{k \to \infty}{\longrightarrow} 0.$$

Etwas gröber gesagt: Für alle $\epsilon > 0$ gibt es eine Polynomfunktion $p_\epsilon : \mathbb{R} \to \mathbb{R}$, so dass gilt:

$$\sup_{-1 \le y \le 1} ||y| - p_\epsilon(y)| < \epsilon,$$

d.h. der Absolutbetrag lässt sich auf $[-1, 1]$ gleichmäßig durch Polynome approximieren. Für das gegebene $f \in \overline{A}$ nehmen wir ein $M > \|f\|_\infty$. Insbesondere ist $M > 0$. Dann gilt $-1 \le f(x)/M \le 1$ für alle $x \in X$, also für alle $\epsilon > 0$

$$\sup_{x \in X} \big| |f(x)| - M p_{\epsilon/M}(f(x)/M) \big| = M \sup_{x \in X} \big| |f(x)/M| - p_{\epsilon/M}(f(x)/M) \big| < M \frac{\epsilon}{M} = \epsilon.$$

Das bedeutet $\|f - p_{\epsilon,M} \circ f\|_\infty < \epsilon$ mit der Polynomfunktion $p_{\epsilon,M} : \mathbb{R} \to \mathbb{R}$, $z \mapsto M p_{\epsilon/M}(z/M)$. Wegen $p_{\epsilon,M} \circ f \in \overline{A}$ ist damit gezeigt: $|f| \in \overline{\overline{A}} = \overline{A}$.
Die Voraussetzungen des Satzes von Kakutani-Krein sind damit für $\overline{A}$ verifiziert.

□

Übung 1.183 (Satz von Stone-Weierstraß – komplexe Version) Gegeben seien ein nichtleerer kompakter Hausdorffraum $(X, \mathcal{T})$ und eine Menge $A \subseteq C(X, \mathbb{C})$ mit den folgenden Eigenschaften:

1. Die konstante Funktion $1 : X \to \mathbb{R}$ mit dem Wert 1 ist ein Element von A.
2. Für alle $f, g \in A$ ist $f + g \in A$.
3. Für alle $f \in A$ und alle $\alpha \in \mathbb{C}$ ist $\alpha f \in A$.
4. Für alle $f, g \in A$ ist $f \cdot g \in A$.
5. Für alle $f \in A$ ist $\overline{f} \in A$, wobei $\overline{f}$ das konjugiert Komplexe von f bezeichnet.
6. A ist punktetrennend.

Zeigen Sie, dass dann A dicht in $(C(X, \mathbb{C}), \| \cdot \|_\infty)$ ist.
Hinweis: Zeigen Sie zuerst, dass $\operatorname{Re} A = \{\operatorname{Re} f |\ f \in A\}$ die Voraussetzungen der reellen Version des Satzes von Stone-Weierstraß erfüllt, und dass $\{\operatorname{Re} f |\ f \in A\} = \{\operatorname{Im} f |\ f \in A\}$ gilt. Erinnern Sie sich dazu an $\operatorname{Re} f = \frac{1}{2}(f + \overline{f})$ und $\operatorname{Im} f = \frac{1}{2i}(f - \overline{f})$.

Korollar 1.184 (Approximationssatz von Weierstraß) *Es sei* $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$*. Für alle kompakten Intervalle* $[a,b] \subseteq \mathbb{R}$*,* $a \leq b$*, ist die Menge* A *aller Polynomfunktionen* $p : [a,b] \to \mathbb{K}$ *dicht in* $(C([a,b],\mathbb{K}), \|\cdot\|_\infty)$*.*

Beweis: Offensichtlich erfüllt der Raum A der Polynomfunktionen alle Voraussetzungen des Satzes von Stone-Weierstraß.

□

Korollar 1.185 (Approximation stetiger Funktionen auf dem Einheitskreis) *Es sei* $S^1 = \{z \in \mathbb{C} |\, |z| = 1\}$ *der Einheitskreis und* A *die Menge aller Funktionen* $f : S^1 \to \mathbb{C}$ *der Gestalt*

$$f(z) = \sum_{k=-n}^{n} \alpha_k z^k$$

mit beliebigem $n \in \mathbb{N}_0$ *und beliebigen Koeffizienten* $\alpha_k \in \mathbb{C}$*.* A *ist also der von den Funktionen* $e_k : S^1 \ni z \mapsto z^k$*,* $k \in \mathbb{Z}$*, aufgespannte* $\mathbb{C}$*-Vektorraum. Dann ist der Raum* A *dicht in* $(C(S^1,\mathbb{C}), \|\cdot\|_\infty)$*.*

Beweis: S^1 ist abgeschlossen und beschränkt, also kompakt. Offensichtlich ist A abgeschlossen unter Addition und Skalarmultiplikation, enthält die Eins: $1 = e_0$, und wegen $e_k \cdot e_l = e_{k+l}$ für $k, l \in \mathbb{Z}$ ist A auch abgeschlossen unter Multiplikation. Nun gilt für alle $z \in S^1$ die Gleichung $\overline{z} = z^{-1}$. Es folgt $\overline{e_k} = e_{-k}$ für alle $k \in \mathbb{Z}$. Das impliziert, dass A auch abgeschlossen unter Komplexkonjugation ist. Weil A die Identität $\mathrm{id}_{S^1} = e_1$ enthält, ist A punktetrennend. Also sind alle Voraussetzungen der komplexen Version des Satzes von Stone-Weierstraß erfüllt, und es folgt die Behauptung.

□

Übung 1.186 (Beziehung zwischen Funktionen auf S^1 und periodischen Funktionen) Eine Funktion $f : \mathbb{R} \to \mathbb{C}$ heißt a-periodisch, $a \in \mathbb{R}$, wenn für alle $x \in \mathbb{R}$ gilt: $f(x+a) = f(x)$. Es bezeichne

$$C_{\mathrm{per}}(\mathbb{R},\mathbb{C}) = \{f : \mathbb{R} \to \mathbb{C} |\, f \text{ ist stetig und } 2\pi\text{-periodisch}\}$$

den Raum aller stetigen, 2π-periodischen Funktionen auf $\mathbb{R}$. Zeigen Sie, dass die Abbildung $\iota : (C(S^1,\mathbb{C}), \|\cdot\|_\infty) \to (C_{\mathrm{per}}(\mathbb{R},\mathbb{C}), \|\cdot\|_\infty)$, $\iota(f)(x) = f(e^{ix})$ für $f \in C(S^1,\mathbb{C})$ und $x \in \mathbb{R}$ eine bijektive lineare Isometrie ist. Folgern Sie, dass der von den Funktionen $\mathbb{R} \ni x \mapsto e^{ikx}$, $k \in \mathbb{Z}$, aufgespannte $\mathbb{C}$-Vektorraum dicht in $(C_{\mathrm{per}}(\mathbb{R},\mathbb{C}), \|\cdot\|_\infty)$ ist.

Übung 1.187 Es sei $B := \{z \in \mathbb{C} |\, |z| \leq 1\}$ die abgeschlossene Einheitskreisscheibe.

1. Es sei $\mathbb{C}[z,\overline{z}]$ die Menge der Polynomfunktionen $p : B \to \mathbb{C}$ in der Variablen z und ihrer komplex Konjugierten $\overline{z}$, also Funktionen der Gestalt

$$p(z) = \sum_{k,l \in \mathbb{N}_0} \alpha_{k,l} z^k \overline{z}^l$$

mit komplexen Zahlen $\alpha_{k,l}$, von denen nur endlich viele von 0 verschieden sind. Zeigen Sie, dass $\mathbb{C}[z,\overline{z}]$ dicht in $(C(B,\mathbb{C}), \|\cdot\|_\infty)$ ist.

2. Nun sei $\mathbb{C}[z]$ die Menge der Polynomfunktionen $p : B \to \mathbb{C}$ in z, also

$$p(z) = \sum_{k=0}^{m} \alpha_k z^k$$

mit komplexen Zahlen α_k und $m \in \mathbb{N}_0$. Zeigen Sie, dass $\mathbb{C}[z]$ *nicht* dicht in $(C(B, \mathbb{C}), \| \cdot \|_\infty)$ ist.
Hinweis: Zeigen Sie, dass

$$L : (C(B, \mathbb{C}), \| \cdot \|_\infty) \to (\mathbb{C}, | \cdot |), \quad L(f) := \frac{1}{2\pi} \int_0^{2\pi} f(e^{ix})\, dx - f(0)$$

eine stetige Linearform ist, die auf $\mathbb{C}[z]$ konstant 0 ist, die aber einen Wert $L(z \mapsto z\bar{z}) = 1$ besitzt.

1.14 ℓ^2-Theorie von Fourierreihen

Die Approximation von stetigen Funktionen auf dem Einheitskreis durch Linearkombinationen der Funktionen $e_k(z) = z^k$, $k \in \mathbb{Z}$, wird viel klarer und einfacher, wenn wir statt der Supremumsnorm $\| \cdot \|_\infty$ eine andere Norm $\| \cdot \|_2$ verwenden, die von einem Skalarprodukt herkommt. In diesem Abschnitt untersuchen wir Approximationen im Sinne dieser 2-Norm.
Weil für Anwendungen auch Funktionen mit Sprungstellen wichtig sind, lassen wir nicht nur stetige Funktionen, sondern Riemann-integrierbare Funktionen zu. Dabei wird eine Funktion $f : S^1 \to \mathbb{C}$ Riemann-integrierbar genannt, wenn $[0, 2\pi] \ni x \mapsto f(e^{ix})$ Riemann-integrierbar ist.

Auf dem Raum $\mathcal{R}(S^1, \mathbb{C})$ aller komplexwertigen Riemann-integrierbaren Funktionen auf dem Einheitskreis wird eine positiv semidefinite hermitesche Sesquilinearform durch

$$\langle f, g \rangle := \frac{1}{2\pi} \int_0^{2\pi} \overline{f(e^{ix})} g(e^{ix})\, dx$$

definiert. Die zugehörige Seminorm wird mit

$$\|f\|_2 = \sqrt{\langle f, f \rangle}$$

bezeichnet. Wir stellen uns Funktionen f, g mit $\|f - g\|_2 = 0$ im Sinne von Lemma 1.9 als miteinander identifiziert vor, so dass $\langle \cdot, \cdot \rangle$ nach diesem "Verkleben" zu einem Skalarprodukt wird. Die Vervollständigung des Raums $(\mathcal{R}(S^1, \mathbb{C}), \|\cdot\|_2)$ nennen wir $(L^2(S^1, \mathbb{C}), \| \cdot \|_2)$. Weil $C(S^1, \mathbb{C})$ dicht in $(\mathcal{R}(S^1, \mathbb{C}), \|\cdot\|_2)$ liegt (siehe Übungen 1.43 und 1.126), ist $(L^2(S^1, \mathbb{C}), \| \cdot \|_2)$ auch die Vervollständigung von $(C(S^1, \mathbb{C}), \| \cdot \|_2)$. Mit dem wie in Beispiel 1.121 hochgehobenen Skalarprodukt wird $L^2(S^1, \mathbb{C})$ zu einem Hilbertraum.

Lemma 1.188 (Fourier-Orthonormalitätsrelation) *Die Familie der Abbildungen* $(e_k : S^1 \to \mathbb{C})_{k\in\mathbb{Z}}$, $e_k(z) = z^k$, *bildet ein orthonormales System bezüglich* $\langle\cdot,\cdot\rangle$. *Das bedeutet:*

$$\forall k, l \in \mathbb{Z} : \langle e_k, e_l\rangle = \delta_{k,l},$$

wobei $\delta_{k,l} = 1_{\{k=l\}}$ *das Kronecker-Delta bezeichnet.*

Beweis: Für gegebene $k, l \in \mathbb{Z}$ berechnen wir:

$$\langle e_k, e_l\rangle = \frac{1}{2\pi}\int_0^{2\pi} \overline{e^{ikx}} e^{ilx}\, dx = \frac{1}{2\pi}\int_0^{2\pi} e^{-ikx} e^{ilx}\, dx = \frac{1}{2\pi}\int_0^{2\pi} e^{i(l-k)x}\, dx$$

Für $k = l$ ist $e^{i(l-k)x} = e^0 = 1$, und es folgt

$$\langle e_k, e_k\rangle = \frac{1}{2\pi}\int_0^{2\pi} 1\, dx = 1.$$

Für $k \neq l$ erhalten wir dagegen

$$\langle e_k, e_l\rangle = \left[\frac{1}{2\pi}\frac{e^{i(l-k)x}}{i(l-k)}\right]_{x=0}^{x=2\pi} = \frac{1}{2\pi i(l-k)}[e^{2\pi i(l-k)} - e^0] = \frac{1}{2\pi i(l-k)}[1-1] = 0.$$

□

Für $m \in \mathbb{N}_0$ sei

$$A_m := \left\{ \sum_{k=-m}^{m} \alpha_k e_k \,\middle|\, \alpha_{-m}, \ldots, \alpha_m \in \mathbb{C} \right\}$$

der von $e_{-m}, \ldots, e_m$ aufgespannte $\mathbb{C}$-Vektorraum. Wir definieren

$$\Pi_m : \mathcal{R}(S^1, \mathbb{C}) \to A_m, \quad \Pi_m(f) = \sum_{k=-m}^{m} \langle e_k, f\rangle\, e_k.$$

Offensichtlich ist Π_m $\mathbb{C}$-linear.
Die Orthonormalitätsrelationen spielen die wesentliche Rolle im Beweis des folgenden Lemmas:

Lemma 1.189 1. *Für alle* $f \in \mathcal{R}(S^1, \mathbb{C})$ *und alle* $g \in A_m$ *gilt* $\langle g, f - \Pi_m(f)\rangle = 0$ *und* $\Pi_m(g) = g$. *Anders gesagt:* Π_m *ist die orthogonale Projektion von* $\mathcal{R}(S^1, \mathbb{C})$ *auf* A_m.

2. *Für alle* $f, g \in \mathcal{R}(S^1, \mathbb{C})$ *gilt*

$$\langle \Pi_m(f), \Pi_m(g)\rangle = \sum_{k=-m}^{m} \overline{\langle e_k, f\rangle}\, \langle e_k, g\rangle\,,$$

$$\|\Pi_m(f)\|^2 = \sum_{k=-m}^{m} |\, \langle e_k, f\rangle\, |^2.$$

Beweis *Zu 1.* Wir beweisen die Aussage zunächst für den Spezialfall $g = e_l$, $l = -m, \ldots, m$. Hier gilt:

$$\langle e_l, \Pi_m(f)\rangle = \left\langle e_l, \sum_{k=-m}^{m} \langle e_k, f\rangle\, e_k \right\rangle = \sum_{k=-m}^{m} \langle e_k, f\rangle \langle e_l, e_k\rangle = \sum_{k=-m}^{m} \langle e_k, f\rangle\, \delta_{l,k} = \langle e_l, f\rangle$$

und daher

$$\langle e_l, f - \Pi_m(f)\rangle = \langle e_l, f\rangle - \langle e_l, \Pi_m(f)\rangle = 0.$$

Weiter gilt

$$\Pi_m(e_l) = \sum_{k=-m}^{m} \langle e_k, e_l\rangle\, e_k = \sum_{k=-m}^{m} \delta_{k,l} e_k = e_l.$$

Der allgemeine Fall $g = \sum_{l=-m}^{m} \alpha_l e_l$ folgt hiermit so:

$$\langle g, f - \Pi_m(f)\rangle = \sum_{l=-m}^{m} \overline{\alpha_l}\, \langle e_l, f - \Pi_m(f)\rangle = 0$$

und

$$\Pi_m(g) = \sum_{l=-m}^{m} \alpha_l \Pi_m(e_l) = \sum_{l=-m}^{m} \alpha_l e_l = g.$$

Zu 2.

$$\langle \Pi_m(f), \Pi_m(g)\rangle = \left\langle \sum_{l=-m}^{m} \langle e_l, f\rangle\, e_l, \sum_{k=-m}^{m} \langle e_k, g\rangle\, e_k \right\rangle = \sum_{l=-m}^{m} \sum_{k=-m}^{m} \overline{\langle e_l, f\rangle}\, \langle e_k, g\rangle \langle e_l, e_k\rangle$$

$$= \sum_{l=-m}^{m} \sum_{k=-m}^{m} \overline{\langle e_l, f\rangle}\, \langle e_k, g\rangle\, \delta_{k,l} = \sum_{k=-m}^{m} \overline{\langle e_k, f\rangle}\, \langle e_k, g\rangle\,.$$

Die zweite Behauptung ist der Spezialfall $f = g$ hiervon.

$\square$

Wir erhalten:

Satz 1.190 (Besselsche Ungleichung) Für jedes $f \in \mathcal{R}(S^1, \mathbb{C})$ und alle $m \in \mathbb{N}_0$ gilt die folgende *"Besselsche Ungleichung"*

$$\sum_{k=-m}^{m} |\,\langle e_k, f\rangle\,|^2 \leq \|f\|_2^2.$$

Es folgt auch

$$\sum_{k\in\mathbb{Z}} |\,\langle e_k, f\rangle\,|^2 \leq \|f\|_2^2.$$

Setzen wir $\mathcal{F}(f) := (\langle e_k, f\rangle)_k \in \mathbb{C}^{\mathbb{Z}}$ für diese f, so können wir dies auch so ausdrücken:

$$\mathcal{F} : (\mathcal{R}(S^1,\mathbb{C}), \|\cdot\|_2) \to (\ell^2(\mathbb{Z},\mathbb{C}), \|\cdot\|_2)$$

ist eine (gleichmäßig) stetige lineare Abbildung mit

$$\forall f \in \mathcal{R}(S^1,\mathbb{C}): \ \|\mathcal{F}(f)\|_2 \le \|f\|_2. \tag{26}$$

Beweis: Aus Lemma 1.189 erhalten wir wegen $\Pi_m(f) \in A_m$ mit der Abkürzung $\Delta_m(f) := f - \Pi_m(f)$:

$$\langle \Pi_m(f), \Delta_m(f)\rangle = 0.$$

Anschaulich gesprochen bedeutet das: $\Delta_m(f) = f - \Pi_m(f)$ und $\Pi_m(f)$ stehen senkrecht aufeinander. Wir erhalten:

$$\begin{aligned}\|f\|_2^2 &= \langle f, f\rangle = \langle \Delta_m(f) + \Pi_m(f), \Delta_m(f) + \Pi_m(f)\rangle \\ &= \|\Delta_m(f)\|_2^2 + \langle \Delta_m(f), \Pi_m(f)\rangle + \langle \Pi_m(f), \Delta_m(f)\rangle + \|\Pi_m(f)\|_2^2 \\ &= \|\Delta_m(f)\|_2^2 + \|\Pi_m(f)\|_2^2 \ge \|\Pi_m(f)\|_2^2 = \sum_{k=-m}^{m} |\,\langle e_k, f\rangle\,|^2 \overset{m\to\infty}{\longrightarrow} \sum_{k\in\mathbb{Z}} |\,\langle e_k, f\rangle\,|^2,\end{aligned}$$

also die Besselsche Ungleichung. Insbesondere ist die Abbildung

$$\mathcal{F} : \mathcal{R}(S^1,\mathbb{C}) \to \ell^2(\mathbb{Z},\mathbb{C})$$

wohldefiniert und erfüllt wegen

$$\sum_{k\in\mathbb{Z}} |\,\langle e_k, f\rangle\,|^2 = \|\mathcal{F}(f)\|_2^2 \text{ für } f \in \mathcal{R}(S^1,\mathbb{C})$$

die Version (26) der Besselschen Ungleichung. Offensichtlich ist die Abbildung $\mathcal{F}$ auch $\mathbb{C}$-linear. Aus der Besselschen Ungleichung und Lemma 1.63 folgt, dass

$$\mathcal{F} : (\mathcal{R}(S^1,\mathbb{C}), \|\cdot\|_2) \to (\ell^2(\mathbb{Z},\mathbb{C}), \|\cdot\|_2)$$

gleichmäßig stetig ist.

$\square$

Weil $(\ell^2(\mathbb{Z},\mathbb{C}), \|\cdot\|_2)$ auf Grund von Satz 1.110 vollständig ist, können wir nach Satz 1.115 die Abbildung $\mathcal{F} : (\mathcal{R}(S^1,\mathbb{C}), \|\cdot\|_2) \to (\ell^2(\mathbb{Z},\mathbb{C}), \|\cdot\|_2)$ zu einer gleichmäßig stetigen Abbildung

$$\mathcal{F} : (L^2(S^1,\mathbb{C}), \|\cdot\|_2) \to (\ell^2(\mathbb{Z},\mathbb{C}), \|\cdot\|_2)$$

fortsetzen. Mit den Dichtheitsargumenten wie in Abschnitt 1.8 erhalten wir, dass auch diese Fortsetzung die Besselsche Ungleichung

$$\forall f \in L^2(S^1,\mathbb{C}): \ \|\mathcal{F}(f)\|_2 \le \|f\|_2. \tag{27}$$

erfüllt.

Definition 1.191 Die Abbildung $\mathcal{F} : (L^2(S^1, \mathbb{C}), \|\cdot\|_2) \to (\ell^2(\mathbb{Z}, \mathbb{C}), \|\cdot\|_2)$ wird *Fourieranalyse* genannt. Die Koeffizienten

$$(\mathcal{F}(f))_k = \langle e_k, f \rangle \,, \quad k \in \mathbb{Z} \tag{28}$$

heißen *Fourierkoeffizienten* von $f \in L^2(S^1, \mathbb{C})$.

Die Gleichung (28) gilt auch für *alle* $f \in L^2(S^1, \mathbb{C})$ mit dem Dichtheitsargument aus Übung 1.75.
In der Version (27) der Besselschen Ungleichung gilt sogar Gleichheit:

Satz 1.192 (Parseval-Gleichung, Fouriersynthese) Die Fourieranalyse

$$\mathcal{F} : (L^2(S^1, \mathbb{C}), \|\cdot\|_2) \to (\ell^2(\mathbb{Z}, \mathbb{C}), \|\cdot\|_2)$$

ist eine *unitäre* Abbildung, d.h. sie ist eine lineare, bijektive Isometrie. Insbesondere gilt die *Parseval-Gleichung*

$$\forall f, g \in L^2(S^1, \mathbb{C}) : \ \langle \mathcal{F}(f), \mathcal{F}(g) \rangle = \langle f, g \rangle \,, \tag{29}$$

speziell

$$\forall f \in L^2(S^1, \mathbb{C}) : \ \|\mathcal{F}(f)\|_2 = \|f\|_2, \tag{30}$$

wobei Skalarprodukt bzw. Norm links in $\ell^2(\mathbb{Z}, \mathbb{C})$ und rechts in $L^2(S^1, \mathbb{C})$ zu verstehen ist.
Die Umkehrung $\mathcal{F}^{-1}$ der Fourieranalyse $\mathcal{F} : L^2(S^1, \mathbb{C}) \to \ell^2(\mathbb{Z}, \mathbb{C})$ wird *Fouriersynthese* genannt. Sie wird wie folgt gegeben:

$$\begin{aligned} &\mathcal{F}^{-1} : \ell^2(\mathbb{Z}, \mathbb{C}) \to L^2(S^1, \mathbb{C}), \\ &\mathcal{F}^{-1}((\alpha_k)_{k\in\mathbb{Z}}) = \sum_{k\in\mathbb{Z}} \alpha_k e_k := \lim_{m\to\infty} \sum_{k=-m}^{m} \alpha_k e_k, \end{aligned} \tag{31}$$

wobei der Limes in $(L^2(S^1, \mathbb{C}), \|\cdot\|_2)$ zu verstehen ist. Die Reihe $\sum_{k\in\mathbb{Z}} \alpha_k e_k$ heißt *Fourierreihe* .

Beweis: Wir beweisen die Parseval-Gleichung (30) für die Norm zunächst für alle $f \in A_m$, $m \in \mathbb{N}_0$, dann für $f \in C(S^1, \mathbb{C})$, und schließlich für $f \in L^2(S^1, \mathbb{C})$.
Für $f \in A_m$ gilt wegen Lemma 1.189 $f = \Pi_m(f)$, also

$$\|f\|_2^2 = \|\Pi_m(f)\|_2^2 = \sum_{k=-m}^{m} |\langle e_k, f \rangle|^2 = \sum_{k\in\mathbb{Z}} |\langle e_k, f \rangle|^2 = \|\mathcal{F}(f)\|_2^2$$

wobei wir beim vorletzten Gleichheitszeichen verwendet haben, dass für $f = \sum_{k=-m}^{m} \alpha_k e_k \in A_m$ und $l \in \mathbb{Z} \setminus \{-m, \ldots, m\}$ aus der Orthonormalitätsrelation folgt:

$$\langle e_l, f \rangle = \sum_{k=-m}^{m} \alpha_k \langle e_l, e_k \rangle = 0.$$

Nun betrachten wir den "Gültigkeitsbereich der Parseval-Gleichung"

$$M := \{f \in L^2(S^1, \mathbb{C}) |\ \|f\|_2 = \|\mathcal{F}(f)\|_2^2\}.$$

Die Menge M ist abgeschlossen in $(L^2(S^1, \mathbb{C}), \|\cdot\|_2)$ als Urbild der 0 unter der stetigen Abbildung $L^2(S^1, \mathbb{C}) \to \mathbb{R}$, $f \mapsto \|f\|_2 - \|\mathcal{F}(f)\|_2^2$, wobei wir die Stetigkeit von $\mathcal{F}$ aus Satz 1.190 verwenden. Setzen wir $A = \bigcup_{m \in \mathbb{N}_0} A_m$, so gilt nach dem eben Gezeigten $A \subseteq M$. Es folgt $\overline{A} \subseteq M$, wobei der Abschluss in $(L^2(S^1, \mathbb{C}), \|\cdot\|_2)$ gemeint ist. Wir zeigen nun $C(S^1, \mathbb{C}) \subseteq \overline{A}$. Die Identität $\mathrm{id} : (C(S^1, \mathbb{C}), \|\cdot\|_\infty) \to (C(S^1, \mathbb{C}), \|\cdot\|_2)$ ist stetig. Dies folgt aus Lemma 1.63, denn id ist linear und es gilt für $f \in C(S^1, \mathbb{C})$:

$$\|f\|_2 = \sqrt{\frac{1}{2\pi} \int_0^{2\pi} |f(e^{ix})|^2 \, dx} \leq \sqrt{\frac{1}{2\pi} \int_0^{2\pi} \|f\|_\infty^2 \, dx} = \|f\|_\infty.$$

Nach Korollar 1.185 ist A dicht in $(C(S^1, \mathbb{C}), \|\cdot\|_\infty)$, also auch dicht in $(C(S^1, \mathbb{C}), \|\cdot\|_2)$, da die stetige Abbildung $\mathrm{id} : (C(S^1, \mathbb{C}), \|\cdot\|_\infty) \to (C(S^1, \mathbb{C}), \|\cdot\|_2)$ nach Übung 1.51 "Berührpunkte nicht abreißt". Damit ist gezeigt: $C(S^1, \mathbb{C}) \subseteq \overline{A} \subseteq M$, d.h. die Parseval-Gleichung für die Norm gilt auch auf $C(S^1, \mathbb{C})$. Nun ist $C(S^1, \mathbb{C})$ dicht in $(L^2(S^1, \mathbb{C}), \|\cdot\|_2)$, denn $(L^2(S^1, \mathbb{C}), \|\cdot\|_2)$ ist die Vervollständigung von $(C(S^1, \mathbb{C}), \|\cdot\|_2)$. Es folgt $L^2(S^1, \mathbb{C}) \subseteq \overline{A} \subseteq M$, d.h. die Parseval-Gleichung für die Norm gilt auf $L^2(S^1, \mathbb{C})$.
Um die Parseval-Gleichung (29) für das Skalarprodukt aus der Parseval-Gleichung für die Norm (30) herzuleiten, verwenden wir die folgende Beziehung zwischen Skalarprodukt und Norm:

$$\langle f, g \rangle = \frac{1}{4} \left(\|f + g\|_2^2 - \|f - g\|_2^2 + i\|f - ig\|_2^2 - i\|f + ig\|_2^2 \right),$$

die in jedem Prähilbertraum gilt. Man sieht diese Beziehung so: Zunächst gilt

$$\|f + g\|_2^2 = \|f\|_2^2 + \langle f, g \rangle + \langle g, f \rangle + \|g\|_2^2 = \|f\|_2^2 + 2 \operatorname{Re} \langle f, g \rangle + \|g\|_2^2.$$

Setzen wir $-g$ statt g ein, erhalten wir hieraus

$$\|f - g\|_2^2 = \|f\|_2^2 - 2 \operatorname{Re} \langle f, g \rangle + \|g\|_2^2,$$

also

$$\operatorname{Re} \langle f, g \rangle = \frac{1}{4} \left(\|f + g\|_2^2 - \|f - g\|_2^2 \right).$$

Setzen wir hier $-ig$ statt g ein, folgt

$$\operatorname{Im} \langle f, g \rangle = \operatorname{Re}(-i \langle f, g \rangle) = \operatorname{Re} \langle f, -ig \rangle = \frac{1}{4} \left(\|f - ig\|_2^2 - \|f + ig\|_2^2 \right)$$

und daher die Behauptung wegen

$$\langle f, g \rangle = \operatorname{Re} \langle f, g \rangle + i \operatorname{Im} \langle f, g \rangle.$$

Wir erhalten hiermit mit Parseval-Gleichung für die Norm und der Linearität von $\mathcal{F}$ für $f, g \in L^2(S^1, \mathbb{C})$:

$$\begin{aligned}\langle \mathcal{F}(f), \mathcal{F}(g) \rangle &= \frac{1}{4}\left(\|\mathcal{F}(f+g)\|_2^2 - \|\mathcal{F}(f-g)\|_2^2 + i\|\mathcal{F}(f-ig)\|_2^2 - i\|\mathcal{F}(f+ig)\|_2^2\right) \\ &= \frac{1}{4}\left(\|f+g\|_2^2 - \|f-g\|_2^2 + i\|f-ig\|_2^2 - i\|f+ig\|_2^2\right) = \langle f, g \rangle\end{aligned}$$

Weil $(L^2(S^1, \mathbb{C}), \|\cdot\|_2)$ ein normierter Raum und

$$\mathcal{F} : (L^2(S^1, \mathbb{C}), \|\cdot\|_2) \to (\ell^2(\mathbb{Z}, \mathbb{C}), \|\cdot\|_2)$$

nach der Parseval-Gleichung eine Isometrie ist folgt einerseits: $\mathcal{F}$ ist injektiv. Andererseits folgt hieraus, weil $(L^2(S^1, \mathbb{C}), \|\cdot\|_2)$ vollständig ist, dass auch das Bild $\mathcal{F}[L^2(S^1, \mathbb{C})] \subseteq \ell^2(\mathbb{Z}, \mathbb{C})$ vollständig ist, also nach Lemma 1.106 abgeschlossen im normierten Raum $(\ell^2(\mathbb{Z}, \mathbb{C}), \|\cdot\|_2)$. Nun ist das Bild $\mathcal{F}[A]$ des Raums $A = \bigcup_{m \in \mathbb{N}_0} A_m = \operatorname{span}\{e_k |\ k \in \mathbb{Z}\}$ genau die Menge $\mathbb{C}^{(\mathbb{Z})}$ aller $(\alpha_k)_{k \in \mathbb{Z}} \in \mathbb{C}^{\mathbb{Z}}$, für die $\alpha_k \neq 0$ höchstens für endlich viele k gilt. Um dies zu sehen, beachte man, dass $\mathcal{F}(e_k)$ für $k \in \mathbb{Z}$ der k-te kanonische Einheitsvektor in $\ell^2(\mathbb{Z}, \mathbb{C})$ ist. Weiter ist $\mathbb{C}^{(\mathbb{Z})}$ nach Übung 1.42 dicht in $\ell^2(\mathbb{Z}, \mathbb{C})$. Also ist auch $\mathcal{F}[L^2(S^1, \mathbb{C})]$ dicht in $(\ell^2(\mathbb{Z}, \mathbb{C}), \|\cdot\|_2)$. Es folgt $\mathcal{F}[L^2(S^1, \mathbb{C})] = \ell^2(\mathbb{Z}, \mathbb{C})$, da $\mathcal{F}[L^2(S^1, \mathbb{C})]$ abgeschlossen und dicht in $(\ell^2(\mathbb{Z}, \mathbb{C}), \|\cdot\|_2)$ ist. Das bedeutet, dass $\mathcal{F}$ auch surjektiv ist. Damit ist gezeigt, dass $\mathcal{F}$ unitär ist.
Wir zeigen nun die Gültigkeit der Umkehrformel (31). Hierzu sei $\alpha = (\alpha_k)_{k \in \mathbb{Z}} \in \ell^2(\mathbb{Z}, \mathbb{C})$ und $\alpha^{(m)} = (1_{\{|k| \le m\}} \alpha_k)_{k \in \mathbb{Z}} \in \mathbb{C}^{(\mathbb{Z})}$ für $m \in \mathbb{N}_0$. Dann gilt wegen $\sum_{k \in \mathbb{Z}} |\alpha_k|^2 < \infty$ nach dem Satz von der dominierten Konvergenz:

$$\|\alpha - \alpha^{(m)}\|_2^2 = \sum_{k \in \mathbb{Z} \setminus \{-m, \ldots, m\}} |\alpha_k|^2 \overset{m \to \infty}{\longrightarrow} 0,$$

also auch

$$\|\mathcal{F}^{-1}(\alpha) - \mathcal{F}^{-1}(\alpha^{(m)})\|_2^2 \overset{m \to \infty}{\longrightarrow} 0,$$

d.h. $\lim_{m \to \infty} \mathcal{F}^{-1}(\alpha^{(m)}) = \mathcal{F}^{-1}(\alpha)$ (Limes bezüglich $\|\cdot\|_2$ gemeint), weil $\mathcal{F}$ unitär ist. Die Umkehrformel (31) folgt nun wegen

$$\mathcal{F}^{-1}(\alpha^{(m)}) = \sum_{k=-m}^{m} \alpha_k e_k, \quad m \in \mathbb{N}_0.$$

□

Ausführlich geschrieben lautet die Parseval-Gleichung für $f, g \in \mathcal{R}([0, 2\pi], \mathbb{C})$ mit $f(0) = f(2\pi)$, $g(0) = g(2\pi)$ und Fourierkoeffizienten

$$\begin{aligned} f_k &:= \frac{1}{2\pi} \int_0^{2\pi} e^{-ikx} f(x)\, dx, \\ g_k &:= \frac{1}{2\pi} \int_0^{2\pi} e^{-ikx} g(x)\, dx \end{aligned}$$

wie folgt:

$$\frac{1}{2\pi}\int_0^{2\pi} \overline{f(x)}g(x)\,dx = \sum_{k\in\mathbb{Z}} \overline{f_k}g_k,$$
$$\frac{1}{2\pi}\int_0^{2\pi} |f(x)|^2\,dx = \sum_{k\in\mathbb{Z}} |f_k|^2.$$

Beispiel 1.193 Wir wenden die Parseval-Gleichung an auf die Funktion $f : [0, 2\pi] \to \mathbb{R}$ mit $f(x) = 1$ für $0 < x < \pi$ und $f(x) = 0$ sonst. Auf der linken Seite erhalten wir

$$\frac{1}{2\pi}\int_0^{2\pi} |f(x)|^2\,dx = \frac{1}{2}.$$

Berechnen wir die Fourierkoeffizienten f_k, $k \in \mathbb{Z}$, von f:

$$f_0 = \frac{1}{2\pi}\int_0^{2\pi} f(x)\,dx = \frac{1}{2}$$

und für $k \in \mathbb{Z}\setminus\{0\}$:

$$f_k = \frac{1}{2\pi}\int_0^{2\pi} e^{-ikx} f(x)\,dx = \frac{1}{2\pi}\int_0^{\pi} e^{-ikx}\,dx = \frac{1}{2\pi}\frac{e^{-ik\pi}-1}{-ik}$$
$$= \begin{cases} (ik\pi)^{-1} & \text{für } k \in 2\mathbb{Z}+1, \\ 0 & \text{für } k \in 2\mathbb{Z}\setminus\{0\}. \end{cases}$$

Die Parseval-Gleichung liefert

$$\frac{1}{2} = \frac{1}{2\pi}\int_0^{2\pi} |f(x)|^2\,dx = \sum_{k\in\mathbb{Z}} |f_k|^2 = \frac{1}{4} + \sum_{k\in 2\mathbb{Z}+1} \frac{1}{(k\pi)^2}$$

also

$$\frac{\pi^2}{4} = \sum_{k\in 2\mathbb{Z}+1} \frac{1}{k^2},$$

anders geschrieben:

$$\frac{\pi^2}{8} = \sum_{k\in 2\mathbb{N}-1} \frac{1}{k^2}.$$

Nun gilt

$$\sum_{k\in\mathbb{N}} \frac{1}{k^2} = \sum_{k\in 2\mathbb{N}-1} \frac{1}{k^2} + \sum_{k\in 2\mathbb{N}} \frac{1}{k^2} = \frac{\pi^2}{8} + \sum_{k\in\mathbb{N}} \frac{1}{(2k)^2} = \frac{\pi^2}{8} + \frac{1}{4}\sum_{k\in\mathbb{N}} \frac{1}{k^2}.$$

Wir erhalten die erstaunliche Formel

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

Übung 1.194 Beweisen Sie analog

$$\sum_{k=1}^{\infty} \frac{1}{k^4} = \frac{\pi^4}{90},$$

indem Sie beide Seiten der Parseval-Gleichung für die "Dreiecksfunktion" $f : [0, 2\pi] \to \mathbb{C}$ mit $f(x) = x$ für $0 \leq x \leq \pi$ und $f(x) = 2\pi - x$ für $\pi \leq x \leq 2\pi$ ausrechnen.

Übung 1.195 Berechnen Sie die Fourierkoeffizienten der Funktionen $f_n, g_n : S^1 \to \mathbb{C}$, $f_n(z) = (\operatorname{Re} z)^n$ und $g_n(z) = (\operatorname{Im} z)^n$, $n \in \mathbb{N}_0$.

Übung 1.196 Es sei $f : \mathbb{R} \to \mathbb{C}$ eine stetig differenzierbare, 2π-periodische Funktion. Zeigen Sie, dass die Fourier-Partialsumme

$$g_n(x) = \sum_{k=-n}^{n} f_k e^{ikx}$$

mit $n \in \mathbb{N}_0$, $x \in \mathbb{R}$ und

$$f_k = \frac{1}{2\pi} \int_0^{2\pi} e^{-ikx} f(x)\, dx$$

für $n \to \infty$ *gleichmäßig* gegen f konvergiert.
Hinweis: Integrieren Sie partiell und verwenden Sie die Besselsche Ungleichung für f' und die Cauchy-Schwarz-Ungleichung, um $\sum_{k\in\mathbb{Z}} |f_k| < \infty$ zu zeigen.

Übung 1.197 (Reelle Version der Fourierreihe) Es sei $f : \mathbb{R} \to \mathbb{R}$ eine stetige 2π-periodische Funktion. Zeigen Sie, dass die Funktionen

$$g_n : [0, 2\pi] \to \mathbb{R}, \quad g_n(x) = c_0 + \sum_{k=1}^{n} [a_k \cos(kx) + b_k \sin(kx)]$$

mit

$$\begin{aligned} c_0 &= \frac{1}{2\pi} \int_0^{2\pi} f(x)\, dx, \\ a_k &= \frac{1}{\pi} \int_0^{2\pi} \cos(kx) f(x)\, dx, \\ b_k &= \frac{1}{\pi} \int_0^{2\pi} \sin(kx) f(x)\, dx \end{aligned}$$

für $n \to \infty$ in $(C([0, 2\pi]), \|\cdot\|_2)$ gegen f konvergieren.
Hinweis: Drücken Sie a_k, b_k und c_0 mit den komplexen Fourierkoeffizienten

$$f_k = \frac{1}{2\pi} \int_0^{2\pi} e^{-ikx} f(x)\, dx, \quad k \in \mathbb{Z}$$

aus und folgern Sie

$$g_n(x) = \sum_{k=-n}^{n} f_k e^{ikx} \text{ für } n \in \mathbb{N}_0,\ x \in \mathbb{R}.$$

Bemerkung: Unter der Zusatzvoraussetzung, dass f *stetig differenzierbar* ist, folgt aus aus Übung 1.196 auch *gleichmäßige* Konvergenz der reellen Version der Fourierreihe, also Konvergenz in $(C([0, 2\pi]), \| \cdot \|_\infty)$.

2 Differentialrechnung mehrerer Variablen

2.1 Visualisierung von Funktionen mehrerer Variablen

In der Analysis mehrerer Variablen ist es essentiell, eine gute geometrische Intuition für Funktionen mehrerer Variablen zu entwickeln. Doch was soll man sich unter einer Funktion $f : \mathbb{R}^m \to \mathbb{R}^n$ anschaulich-geometrisch vorstellen? Für nicht zu große m und n kann man mit verschiedenen graphischen Darstellungstechniken noch gut eine anschauliche Vorstellung von solchen Funktionen f gewinnen. Wir besprechen nun einige dieser Techniken zur Veranschaulichung von Funktionen f "vom Typ $\mathbb{R}^m \to \mathbb{R}^n$". Die "Typangabe" soll hier bedeuten, dass der Definitionsbereich von f eine Teilmenge von $\mathbb{R}^m$ und der Wertebereich eine Teilmenge von $\mathbb{R}^n$ ist.

Funktionen vom Typ $\mathbb{R} \to \mathbb{R}^n$. Funktionen $f = (f_1, \ldots, f_n) : \mathbb{R} \to \mathbb{R}^n$ kann man sich als Kurve im $n+1$-dimensionalen Raum $\mathbb{R}^{n+1}$ vorstellen, und zwar als Graph

$$\{(x, f_1(x), \ldots, f_n(x)) |\, x \in \mathbb{R}\} \subset \mathbb{R}^{n+1}$$

Das ist für $n = 1, 2$ einfach, weil unser Anschauungsraum dreidimensional ist, wird aber für $n \geq 3$ schwieriger.
Indem man die Zeit als vierte Dimension zu unserem Anschauungsraum hinzunimmt, gewinnt man noch eine Dimension mehr für die Anschauung: Man stellt sich f dann mit einem Punkt im Raum $\mathbb{R}^n$ vor, der sich mit der Zeit entlang einer Kurve im Raum $\mathbb{R}^n$ bewegt:

$$\mathbb{R} \ni \text{ Zeit } t \mapsto \text{ Ort } f(t) \in \mathbb{R}^n$$

Das funktioniert für $n = 1, 2, 3$ gut und wird erst für $n \geq 4$ schwieriger. Dann bleibt natürlich immer noch die Möglichkeit, sich die Komponenten $f_1, \ldots, f_n$ einzeln oder zu kleineren Blöcken zusammengefasst vorzustellen, z.B. (f_1, f_2) und (f_3, f_4) für $n = 4$.

Funktionen vom Typ $\mathbb{R}^m \to \mathbb{R}$.

Beispiel 2.1 Man kann sich die "Nordhalbkugel" und "Südhalbkugel" der Einheitssphäre $S^2 = \{x \in \mathbb{R}^3 |\, \|x\|_2\} = 1$ als Graphen

$$\{(x, y, f_\pm(x, y)) \in \mathbb{R}^3 | x^2 + y^2 \leq 1\}$$

der Funktionen

$$f_\pm : \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 \leq 1\} \to \mathbb{R}, \quad f_\pm(x, y) = \pm\sqrt{1 - x^2 - y^2}$$

vorstellen. Diese Beschreibung endet am Äquator; sie wird dort "*singulär*" durch das Auftreten vertikaler Tangentialebenen und weil Punkte (x, y) mit $x^2 + y^2 = 1$ am Rand des Definitionsgebiets von $f_\pm$ liegen. Oft entfernt man diese "Singularitäten" aus dem Definitionsbereich und arbeitet auf offenen Mengen. Natürlich ist das nur eine *Beschreibungssingularität* oder *Kartensingularität* der "Erdkugel", denn geometrisch gesehen ist die Erde am Äquator ebenso "*regulär*" wie an allen anderen Stellen auch.

Allgemeiner: Für nicht zu große m kann man sich auch $f : \mathbb{R}^m \to \mathbb{R}$ als Graph

$$\{(x, f(x))|\, x \in \mathbb{R}^m\} = \{(x_1, \dots, x_m, f(x_1, \dots, x_m))|\, x \in \mathbb{R}^m\} \subset \mathbb{R}^{m+1}$$

vorstellen. Insbesondere stellt man sich damit Funktionen $f : \mathbb{R}^2 \to \mathbb{R}$ als Höhenprofil eines Gebirges im Anschauungsraum $\mathbb{R}^3$ vor. Für zweidimensionale Graphiken im Druck oder auf dem Bildschirm funktioniert das auch noch mehr oder weniger gut, indem man die Fläche perspektivisch darstellt, evtl. mit Schattierungen, z.B. in Panoramakarten.
Aus Wanderkarten kennen Sie auch eine präzisere Darstellung von Höhenprofilen durch Höhenlinien oder Niveaulinien: Orte gleicher Höhe werden durch eine Höhenlinie markiert, zum Beispiel eine Höhenlinie pro 50 Meter Höhenunterschied. Die Wanderer unter Ihnen wissen, dass ein Gelände umso steiler ist, je dichter die Höhenlinien auf der Wanderkarte zusammen liegen, dass man eine gemütliche Wanderung auf gleichem Höhenniveau erwarten kann, wenn der Weg auf einer Höhenlinie läuft, aber dass man einen steilen Aufstieg/Abstieg erwarten muss, wenn man senkrecht zu den Höhenlinien aufwärts/abwärts läuft.
Auch in der Wetterkunde verwendet man diese Darstellungstechnik, zum Beispiel zur Veranschaulichung des Luftdrucks in Wetterkarten: Isobarenlinien sind die Linien gleichen Luftdrucks. Liegen die Isobaren dicht zusammen, hat man hohe Luftdruckgegensätze und muss mit einem Sturm rechnen, während man fast Windstille erwarten kann, wenn die Isobaren weit auseinander liegen oder gar ein ganzes Flächenstück zu einer Isobaren gehört.
Formal gesehen ist die Niveaulinie $N_h(f)$ einer Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ zum Niveau $h \in \mathbb{R}$ das Urbild

$$N_h(f) = f^{-1}[\{h\}],$$

also der Lösungsraum der Gleichung

$$f(x, y) = h.$$

Zum Beispiel sind die Niveaulinien

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x, y) = x^2 + y^2$$

zu positiven Niveaus $h > 0$ Kreise um $(0, 0)$ mit Radius $\sqrt{h}$. Das Niveau $h = 0$ ist eine Ausnahme, eine *Singularität*: Es ist keine Linie mehr, sondern nur ein Punkt. Für

$$g : \mathbb{R}^2 \to \mathbb{R}, \quad g(x, y) = xy$$

sind die Niveaulinien $g^{-1}[\{h\}]$ für $h \neq 0$ Hyperbeln, für $h = 0$ jedoch die beiden Koordinatenachsen, die sich im Nullpunkt schneiden. Auch hier ist der Nullpunkt eine *Singularität*.
Für Abbildungen $f : \mathbb{R}^3 \to \mathbb{R}$ ist das Niveaugebilde $f^{-1}[\{y\}]$ "in regulären Fällen" eine (meist gekrümmte) Fläche. (Untypische, "singuläre" Ausnahmen, wie zum Beispiel bei einer konstanten Funktion f, gibt es natürlich auch.) Zum Beispiel ist für die Abbildung

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x, y, z) = x^2 + y^2 + z^2$$

das Niveaugebilde $f^{-1}[\{h\}]$ zum Niveau $h > 0$ eine Sphäre (Kugeloberfläche) mit Radius $\sqrt{h}$, für $h = 0$ jedoch keine Fläche, sondern nur der Nullpunkt. Der Nullpunkt ist also eine "Singularität" von f. Die Beschreibung der "Erdoberfläche" als Niveaufläche $f^{-1}[\{1\}]$ ist nun *singularitätenfrei* – anders als oben die Beschreibung mit Nord- und Südhalbkugel als Graph. Für die Abbildung

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

ist die Niveaufläche $g^{-1}[h]$ für Niveaus $h > 0$ ein zweischaliges Hyperboloid, für $h < 0$ ein einschaliges Hyperboloid und für $h = 0$ ein Doppelkegel mit Spitze in $(0, 0, 0)$. Am Nullpunkt ist das Niveaugebilde also keine glatte Fläche; der Nullpunkt ist hier singulär. Man kann sich die Niveauflächen von g vorstellen, indem man die Niveaugebilde von $(x, y) \mapsto x^2 - y^2$ im dreidimensionalen Raum um die x-Achse rotiert.
Auch für $m > 3$ ist das Niveaugebilde $f^{-1}[\{h\}]$ einer Abbildung $f : \mathbb{R}^m \to \mathbb{R}$, also der Lösungsraum der Gleichung

$$f(x_1, \ldots, x_m) = h$$

"in regulären Fällen" ein $(m - 1)$-dimensionales Gebilde, "Hyperfläche" genannt. Das Präfix *"Hyper-"* bezieht sich darauf, dass das Gebilde nicht zweidimensional sein muss, sondern *eine* Dimension weniger als der Grundraum $\mathbb{R}^m$ hat, also *Kodimension* 1 besitzt. Solch eine (reguläre) Hyperfläche nennt man auch eine $(m-1)$-*dimensionale Untermannigfaltigkeit* von $\mathbb{R}^m$.
Besonders wichtig für die Differentialrechnung mehrerer Variablen ist es, eine gute geometrisch-anschauliche Vorstellung von *Linearformen*, also von linearen Abbildungen $f : \mathbb{R}^m \to \mathbb{R}$ zu entwickeln. Hier sind (im "regulären" Fall $f \neq 0$) die Niveauflächen parallele Hyperebenen (also parallele Geraden in der Ebene für $m = 2$ und parallele Ebenen im Raum für $m = 3$). Diese Hyperebenen liegen äquidistant zueinander, wenn man die Niveaus äquidistant wählt. Man kann sich diese Hyperebenen physikalisch als Wellenfronten einer ebenen Welle im Raum oder in der Ebene vorstellen. Eine weitere Veranschaulichung von Linearformen erhält man, indem man sich f in der Form $f : \mathbb{R}^m \ni x \mapsto \langle k, x\rangle \in \mathbb{R}$ gegeben vorstellt, wobei $k \in \mathbb{R}^m$ und $\langle \cdot, \cdot \rangle$ das euklidische Skalarprodukt bezeichnet. In der Linearen Algebra lernen Sie, dass jede Linearform auf $\mathbb{R}^m$ genau eine solche Darstellung besitzt, denn die Abbildung $\mathbb{R}^m \to (\mathbb{R}^m)' = \mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^m, \mathbb{R})$, $k \mapsto \langle k, \cdot\rangle$ vom Raum $\mathbb{R}^m$ in seinen Dualraum $(\mathbb{R}^m)'$ ist ein Isomorphismus. Der "Normalenvektor" oder "Wellenvektor" k steht senkrecht auf den Niveau(hyper)ebenen $\{x \in \mathbb{R}^m | \langle k, x\rangle = h\}$ und zeigt in Richtung ansteigender Werte. Je größer sein Betrag, desto dichter liegen die Niveauebenen (bei gleichem Niveauabstand) zusammen, desto "steiler" das Höhenprofil, desto kleiner die Wellenlänge im Wellenbild.
Weitere Visualisierungsmöglichkeiten von Abbildungen f vom Typ $\mathbb{R}^2 \to \mathbb{R}$ erhält man, indem man Funktionswerte $f(x)$ durch Farben kodiert, z.B. durch Graustufen oder mit einer Farbskala. Auch diese Technik kennen Sie von der Temperaturdarstellung in Wetterkarten: Rot für heiße Regionen, Gelb-Grün für moderat warme Gegenden und Blau für kalte Gebiete. Auch in der Medizin wird diese Darstellungstechnik bei bildgebenden Verfahren (z.B. Computertomographie, Kernspintomographie, Ultraschall) häufig eingesetzt.

Eine weitere Dimension in der Darstellungsmöglichkeit gewinnt man wieder durch Hinzunehmen der Zeit: Ein "Wetterfilm" zur Veranschaulichung der Temperaturänderungen, bei dem sich die Farbe im Bild örtlich und zeitlich verändert, veranschaulicht eine Abbildung vom Typ $\mathbb{R}^3 \to \mathbb{R}$,

$$(\text{Längengrad}, \text{Breitengrad}, \text{Zeit}) \mapsto \text{Temperatur}.$$

Eine Abbildung vom Typ $\mathbb{R}^4 \to \mathbb{R}$ kann man sich mit einer Schar von sich zeitlich verändernden Niveaugebilden

$$N(h,t) = \{(x,y,z) \in \mathbb{R}^3 | \; f(t,x,y,z) = h\}$$

im Raum vorstellen, wobei h das Niveau und t die Zeit bedeutet, oder, wenn man will, auch als farbige Gebilde im Raum $\mathbb{R}^3$, die sich zeitlich verändern.[10]

Funktionen vom Typ $\mathbb{R}^2 \to \mathbb{R}^2$ oder $\mathbb{R}^2 \to \mathbb{R}^3$. Die einfachste Möglichkeit, sich Abbildungen $f : \mathbb{R}^2 \to \mathbb{R}^2$ vorzustellen, ist es, sich jede der beiden Komponenten (f_1, f_2) von f einzeln vorzustellen, z.B. indem man die Höhenlinien von f_1 und von f_2 in das gleiche Bild zusammen einzeichnet. Diese Technik ist insbesondere in der (komplexen) Funktionentheorie wertvoll, in der es um "analytische" oder "holomorphe" Funktionen, also um im komplexen Sinn differenzierbare Funktionen f vom Typ $\mathbb{C} \to \mathbb{C}$, also ebenfalls vom Typ $\mathbb{R}^2 \to \mathbb{R}^2$, geht. (In der Vorlesung "Funktionentheorie" lernen Sie, dass sich die Höhenlinien von Real- und Imaginärteil von f dann senkrecht schneiden.) Die Technik, die Höhenlinien der drei Komponenten (f_1, f_2, f_3) einer Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^3$ alle in ein gemeinsames Bild einzuzeichnen, ist dagegen wenig nützlich, weil das schnell zu unübersichtlich wird.
Man kann sich eine Abbildung f vom Typ $\mathbb{R}^2 \to \mathbb{R}^2$ oder $\mathbb{R}^2 \to \mathbb{R}^3$ dagegen vorstellen, indem man sich $\mathbb{R}^2$ wie ein Millimeterpapier als eine mit Koordinatenlinien bedruckte Gummimembran vorstellt, die dann mit f deformiert in der Ebene $\mathbb{R}^2$ oder im Raum $\mathbb{R}^3$ liegt.
Man stelle sich zum Beispiel die Erdkugel mit Längen- und Breitengradlinien markiert vor: Dies liefert eine Veranschaulichung der Abbildung

$$f : \mathbb{R}^2 \to \mathbb{R}^3, \quad f(\theta, \phi) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta).$$

(ϕ = Längengrad $\in [0, 2\pi]$ und θ = Breitengrad $\in [0, \pi]$, etwas abweichend von der Konvention in der Geographie.) Man beachte, dass der Nordpol $(0, 0, 1)$ und der Südpol $(0, 0, -1)$ bei dieser Beschreibung singuläre Stellen werden, weil die Meridiane (Längengradlinien) sich dort schneiden. (Natürlich ist das nur eine *Beschreibungssingularität* der

[10] Bedenkt man, dass der Raum aller Farben dreidimensional ist, mit den Koordinaten "Rotanteil, Grünanteil, Blauanteil", etwa beim Farbfernsehen, oder auch - nach einem Koordinatenwechsel - mit den Koordinaten "Helligkeit, Sättigungsgrad, Farbton", so kann man sich mit dieser Technik zum Spaß sogar noch Abbildungen vom Typ $\mathbb{R}^3 \to \mathbb{R}^3$ oder - unter Hinzunahme der Zeit - vom Typ $\mathbb{R}^4 \to \mathbb{R}^3$ als farbige, verschieden helle und verschiedene Pastellfarben verwendende, sich zeitlich verändernde Gebilde im Raum vorstellen, auch wenn das zugegebenermaßen schwierig ist und nicht viel nützt.

Erdkugel, denn geometrisch gesehen ist am Nordpol und am Südpol die Erdoberfläche ebenso regulär wie an allen anderen Stellen auch.) Außerdem verliert man einen Meridian in der Beschreibung, wenn man zum offenen Kern $]0,\pi[\times]0,2\pi[\ni(\theta,\phi)$ übergeht.
(Wer will, kann sich auch Abbildungen vom Typ $\mathbb{R}^3\to\mathbb{R}^3$ mit dieser Technik als zeitlich veränderliche Deformationen einer Gummimembran vorstellen, oder eine Abbildung vom Typ $\mathbb{R}^2\to\mathbb{R}^3$ als eine zeitlich veränderliche Kurve im Raum.)

Funktionen vom Typ $\mathbb{R}^n\to\mathbb{R}^n$. Solche Abbildungen kann man sich (unter gewissen Regularitätsvoraussetzungen) als *Koordinatenwechsel* vorstellen, also als Umrechnung zwischen kartesischen und krummlinigen Koordinaten. Beispiele liefern die Polarkoordinatendarstellung vom Typ $\mathbb{R}^2\to\mathbb{R}^2$, $(r,\phi)\mapsto(x,y)=(r\cos\phi,r\sin\phi)$ und die Kugelkoordinatendarstellung vom Typ $\mathbb{R}^3\to\mathbb{R}^3$, $(r,\theta,\phi)\mapsto(x,y,z)=(r\sin\theta\cos\phi,r\sin\theta\sin\phi,r\cos\theta)$.
Eine ganz andere, sehr nützliche Veranschaulichung von Abbildungen f vom Typ $\mathbb{R}^2\to\mathbb{R}^2$ oder $\mathbb{R}^3\to\mathbb{R}^3$ liefern *Vektorfelder*: Man erhält sie, indem man sich für jede Stelle x den Vektor $f(x)$ als einen Pfeil vorstellt, der mit Fußpunkt bei x in eine Graphik eingezeichnet wird. Diese Technik kennen Sie ebenfalls von Wetterkarten bei der Darstellung von Wind: An jeder Stelle wird ein Pfeil in Windrichtung gezeichnet, dessen Länge die Windstärke an dieser Stelle darstellt. Die gleiche graphische Darstellungstechnik wird oft in der Physik und Elektrotechnik zur Veranschaulichung des elektrischen Feldes oder auch des magnetischen Feldes benutzt. (Wer möchte, kann sich so auch eine Abbildung $\mathbb{R}^3\to\mathbb{R}^2$ oder $\mathbb{R}^4\to\mathbb{R}^3$ als zeitlich veränderliche Vektorfelder vorstellen.) Mit etwas abstrahierter geometrischer Intuition kann man sich so auch für $n>3$ Abbildungen $f:\mathbb{R}^n\to\mathbb{R}^n$ als Vektorfelder (und $f:\mathbb{R}\times\mathbb{R}^n\to\mathbb{R}^n$ als zeitlich veränderliche Vektorfelder) im n-dimensionalen Raum $\mathbb{R}^n$ vorstellen. Jedenfalls liefert sie den Anlass für folgende Sprechweise:

Definition 2.2 *Es sei $U\subseteq\mathbb{R}^n$ mit $n\in\mathbb{N}$. Eine Abbildung $f:U\to\mathbb{R}^n$ wird Vektorfeld auf U genannt. Oft beschränkt man sich auf den Fall, dass U offen und f stetig ist.*

Die Veranschaulichung durch Vektorfelder liefert uns auch ein geometrisches Bild von Differentialgleichungssystemen $\dot{y}(t)=f(y(t))$ mit einem Vektorfeld $f:\mathbb{R}^n\to\mathbb{R}^n$ und einer Lösungskurve $y:\mathbb{R}\to\mathbb{R}^n$: Stellen wir uns f als Vektorfeld der Windgeschwindigkeit im Raum ($n=3$) oder in der Ebene ($n=2$) vor. Nehmen wir ein "Teilchen" (etwa eine Vogelfeder oder ein Luftmolekül), das mit dem Wind mitgetragen wird, also "mitfließt". Seine mit der Zeit t parametrisierte Bahnkurve $t\mapsto y(t)$, also *Flusslinie*, wird dann durch eine Lösung y des Differentialgleichungssystems beschrieben. Die Gesamtheit der Lösungen kann man sich dann auch als Flussabbildung Φ vom Typ $\mathbb{R}\times\mathbb{R}^n\to\mathbb{R}^n$, $\Phi:(\text{Zeitdauer } t,\text{Startpunkt } a)\mapsto\text{Endpunkt } y(t)$ vorstellen. Umgekehrt kann man dieses "Flusslinienbild" auch verwenden, um sich Vektorfelder vorzustellen. In der Physik, insbesondere in der Schule, verwendet man dieses Flusslinienbild oft zur Veranschaulichung elektrischer oder magnetischer Felder. Natürlich enthält es nur dann die gesamte Information über das Vektorfeld, wenn man auch die Flussgeschwindigkeit graphisch darstellt, z.B. in einem Film. Die Lösungen y eines allgemeineren Differentialgleichungssystems $\dot{y}(t)=f(t,y(t))$ mit einem *zeitabhängigen* Vektorfeld $f:\mathbb{R}\times\mathbb{R}^n\to\mathbb{R}^n$ kann man sich ebenso als Bahnkurven von Teilchen vorstellen, die in einer Flüssigkeit oder in einem Gas suspendiert sind, deren Geschwindigkeitsfeld sich räumlich und zeitlich ändert; man denke an Filme zur Entwicklung des Winds im Wetterbericht.

Abbildungen vom Typ $\mathbb{R}^n \to (\mathbb{R}^n)'$. Eine anschauliche Intuition zu *Linearformenfeldern* zu gewinnen, also von Abbildungen $\omega : \mathbb{R}^n \to (\mathbb{R}^n)'$ von $\mathbb{R}^n$ in seinen *Dualraum*, ist schon eine echte Herausforderung an die geometrische Vorstellungskraft. Solche Abbildungen werden auch *1-Formen*, *Differentialformen 1. Grades* oder *Pfaffsche Formen* genannt. Sie spielen in der Differentialrechnung mehrerer Variablen eine zentrale Rolle. Wer will, kann sich eine 1-Form als ein Feld von Niveauflächen oder "Wellenfronten" im Raum vorstellen, das aber von Punkt zu Punkt variiert. Die Vorstellung ist auch schon im Fall $n = 2$ oder $n = 3$ schwierig, weil Linearformenfelder ja Abbildungen vom Typ $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $(x, y) \mapsto \omega_x(y)$ kodieren, die im y-Argument linear sind: Das sind also Abbildungen von $2n$ Variablen! Hilfreich ist hier die Übersetzung von 1-Formen in Vektorfelder, die sich aus dem Isomorphismus $\mathbb{R}^n \to (\mathbb{R}^n)'$, $k \mapsto \langle k, \cdot\rangle$ ergibt: Dann kann man sich 1-Formen einfach als Vektorfelder vorstellen, muss aber beachten, dass diese Vorstellung an der Wahl eines Skalarprodukts hängt. Eine andere Veranschaulichung bekommt man, wenn man sich an jeder Stelle x im Raum den *Kern* $\operatorname{Ker} \omega_x = \{y \in \mathbb{R}^n |\ \omega_x(y) = 0\}$ als (Hyper-)Ebene, verschoben um x und angeheftet bei x, vorstellt. Dieses "Hyperebenenfeld" steht also senkrecht zum obigen Vektorfeld. Bei dieser Veranschaulichung fehlt natürlich die Information, wie dicht die Niveauflächen von ω_x zusammenliegen.

Zusammenfassung: Beschreibung von Flächen im Raum. Wir haben nun drei verschiedene Darstellungen von Flächen im Raum besprochen:

- als Graph einer Abbildung vom Typ $\mathbb{R}^2 \to \mathbb{R}$ (Beispiel Nord- und Südhalbkugel),
- als Bild einer "Parameterdarstellung" vom Typ $\mathbb{R}^2 \to \mathbb{R}^3$ (Beispiel Längen- und Breitenkreisbeschreibung),
- als Niveaugebilde einer Abbildung f vom Typ $\mathbb{R}^3 \to \mathbb{R}$, also als Lösungsraum einer Gleichung $f(x, y, z) = 0$ (Beispiel $x^2 + y^2 + z^2 = 1$).

Bei allen drei Beschreibungen sind "Singularitäten" möglich.
Allgemeiner kann man m-dimensionale "*Untermannigfaltigkeiten*" im n-dimensionalen Raum, also $(n - m)-$*kodimensionale* Gebilde (unter Ausschluss von "Singularitäten") ebenfalls auf drei verschiedene Weisen beschreiben:

- mit "*Graphen*" $\{(x, f(x))|x \in U \subseteq \mathbb{R}^m\}$ von Abbildungen f vom Typ $\mathbb{R}^m \to \mathbb{R}^{n-m}$,
- mit "*Parameterdarstellungen*" als Wertebereiche von Abbildungen vom Typ $\mathbb{R}^m \to \mathbb{R}^n$,
- mit "*Niveaugebilden*" $f^{-1}[\{y\}]$ mit festem $y \in \mathbb{R}^{n-m}$ von Abbildungen f vom Typ $\mathbb{R}^n \to \mathbb{R}^{n-m}$, anders gesagt als Lösungsraum von $n - m$ Gleichungen mit n Unbe-

kannten $x_1, \dots, x_n$, z.B. für $y = 0$ durch

$$\begin{aligned} f_1(x_1, \dots, x_n) =& 0, \\ f_2(x_1, \dots, x_n) =& 0, \\ \vdots \quad & \vdots \\ f_{n-m}(x_1, \dots, x_n) =& 0. \end{aligned}$$

Dies wird uns später noch intensiv beschäftigen.

2.2 Partielle Ableitungen

Als ersten und einfachsten mehrdimensionalen Ableitungsbegriff besprechen wir die "partielle Ableitung". Hierbei handelt es sich um die Ableitung nach einer Variablen bei festgehaltenen anderen Variablen:

Definition 2.3 (Partielle Ableitung) Es seien $n \in \mathbb{N}$, $U \subseteq \mathbb{R}^n$, $j \in \{1, \dots, n\}$, $f : U \to \mathbb{C}$ eine Abbildung und $x = (x_1, \dots, x_n) \in U$ ein Punkt. Weiter bezeichne $e_j = (\delta_{j,k})_{k=1,\dots,n}$ den j-ten kanonischen Einheitsvektor in $\mathbb{R}^n$. Die Abbildung f heißt *partiell differenzierbar* in x nach der j-ten Koordinate, wenn die Abbildung $\{t \in \mathbb{R} | \, x + te_j \in U\} \ni t \mapsto f(x + e_j t)$ in einer Umgebung von 0 definiert und an dieser Stelle differenzierbar ist. Die Ableitung

$$D_j f(x) := \frac{d}{dt} f(x + e_j t)\Big|_{t=0}$$

an der Stelle 0 wird dann *partielle Ableitung* von f an der Stelle x nach der j-ten Koordinate (oder auch nach x_j) genannt. Anders gesagt:

$$D_j f(x) = \lim_{h \to 0} \frac{f(x_1, \dots, x_{j-1}, x_j + h, x_{j+1}, \dots, x_n) - f(x_1, \dots, x_n)}{h}.$$

Ist eine Funktion nach allen Koordinaten partiell differenzierbar, so heißt sie partiell differenzierbar (ohne Zusatz). Der (Spalten-)vektor aus den partiellen Ableitungen von f wird *Gradient* von f genannt:[11]

$$\operatorname{grad} f(x) := \nabla f(x) := \begin{pmatrix} D_1 f(x) \\ D_2 f(x) \\ \vdots \\ D_n f(x) \end{pmatrix}.$$

Für die partielle Ableitung sind viele andere Notationen im Gebrauch, z.B.:

$$D_j f(x) = \frac{\partial}{\partial x_j} f(x) = \frac{\partial f}{\partial x_j}(x) = \partial_j f(x) = \nabla_j f(x) = f_{x_j}$$

[11] Das Symbol "∇" wird als "Nabla" gelesen.

Beispiel 2.4 Die partielle Ableitung $\frac{\partial}{\partial x} f(x,y)$ einer Funktion f zweier Variablen beschreibt also, wie stark f in x- Richtung ansteigt, während $\frac{\partial}{\partial y} f(x,y)$ beschreibt, wie stark sie in y-Richtung ansteigt.

1. Für $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = xy$ ist

$$D_1 f(x,y) = \frac{\partial}{\partial x} f(x,y) = y, \quad D_2 f(x,y) = \frac{\partial}{\partial y} f(x,y) = x,$$

   also

$$\operatorname{grad} f(x,y) = \begin{pmatrix} y \\ x \end{pmatrix}.$$

2. Für $x, y \in \mathbb{R}$ mit $y \neq 0$ ist

$$\frac{\partial}{\partial x} \frac{x}{y} = \frac{1}{y}, \quad \frac{\partial}{\partial y} \frac{x}{y} = -\frac{x}{y^2}.$$

3. Für $x > 0, y \in \mathbb{R}$ erhalten wir

$$\frac{\partial}{\partial x} x^y = y x^{y-1}, \quad \frac{\partial}{\partial y} x^y = \frac{\partial}{\partial y} e^{y \log x} = x^y \log x.$$

4. Für $x = (x_1, \ldots, x_n) \in \mathbb{R}^n \setminus \{0\}$ und $j = 1, \ldots, n$ erhalten wir

$$\frac{\partial}{\partial x_j} \|x\|_2 = \frac{\partial}{\partial x_j} \sqrt{\sum_{k=1}^n x_k^2} = \frac{2x_j}{2\sqrt{\sum_{k=1}^n x_k^2}} = \frac{x_j}{\|x\|_2}.$$

   Das Gleiche in Vektornotation:

$$\nabla \|x\|_2 = \|x\|_2^{-1} x.$$

Übung 2.5 Berechnen Sie den Gradienten der folgenden Funktionen:

1. $f : \mathbb{R}^2 \times \mathbb{R}^+ \to \mathbb{R}$, $f(x,y,z) = \frac{1}{\sqrt{2\pi z}} e^{-(x-y)^2/(2z)}$,
2. $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = \langle a, x \rangle$ mit gegebenem $a \in \mathbb{R}^n$, wobei $\langle ., . \rangle$ das euklidische Skalarprodukt bezeichnet,
3. $f : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$, $f(x) = \|x\|_2^\alpha$ mit gegebenem $\alpha \in \mathbb{R}$,
4. $f : \mathbb{R}^3 \setminus \{0\} \to \mathbb{R}$, $f(x) = \frac{\langle a, x \rangle}{\|x\|_2^3}$ mit gegebenem $a \in \mathbb{R}^3$.

Die partielle Ableitung ist der Spezialfall des folgenden Begriffs der *Richtungsableitung* für Koordinatenrichtungen:

Definition 2.6 Es seien $n \in \mathbb{N}, U \subseteq \mathbb{R}^n$, $f : U \to \mathbb{C}$ eine Abbildung, $x = (x_1, \ldots, x_n) \in U$ ein Punkt und $a \in \mathbb{R}^n$ so, dass $I := \{t \in \mathbb{R} |\ x+ta \in U\}$ eine Umgebung von 0 ist. Existiert die Ableitung

$$D_a f(x) := \frac{d}{dt} f(x+ta)\Big|_{t=0},$$

so wird sie *Richtungsableitung* von f in x in Richtung des Vektors a genannt. Ist allgemeiner $k : I \to U$ eine "parametrisierte Kurve durch x" (also I ein Intervall mit $0 \in I^\circ$, alle Komponenten von c differenzierbar und $k(0) = x$), so wird die Ableitung $(f \circ k)'(0)$ Richtungsableitung von f in x entlang der Kurve k genannt.

Man kann sich die Richtungsableitung veranschaulichen, indem man sich $f : U \to \mathbb{R}$ als Höhenprofil und die Kurve k als einen Wanderweg vorstellt: $(f \circ k)' > 0$ bedeutet, dass die Wanderung bergauf führt, $(f \circ k)' < 0$ bedeutet, dass es bergab geht.
Partielle Ableitungen sind also Richtungsableitungen in Richtung der kanonischen Einheitsvektoren e_j, und eine Richtungsableitung in x in Richtung eines Vektors a ist eine Richtungableitung entlang der Geraden k durch x, die mit dem Geschwindigkeitsvektor a durchlaufen wird.
Anders als bei Ableitungen in einer Dimension folgt aus partieller Differenzierbarkeit *nicht* die Stetigkeit. Hierzu ein Gegenbeispiel: Die Abbildung

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

ist partiell differenzierbar. In der Tat: Für $(x,y) \neq (0,0)$ gilt

$$D_1 f(x,y) = \frac{y}{x^2+y^2} - \frac{2x^2y}{(x^2+y^2)^2}, \quad D_2 f(x,y) = \frac{x}{x^2+y^2} - \frac{2xy^2}{(x^2+y^2)^2}.$$

Doch auch in $(0,0)$ ist f partiell differenzierbar, denn es gilt $f(x,0) = 0 = f(0,y)$ für $x, y \in \mathbb{R}$ und daher $D_1 f(0,0) = 0 = D_2 f(0,0)$. Die Abbildung f ist jedoch unstetig in 0, denn es gilt

$$\lim_{n\to\infty} f(\frac{1}{n}, \frac{1}{n}) = \lim_{n\to\infty} \frac{1}{2} = \frac{1}{2} \neq 0 = f(0,0).$$

Man kann sich dieses Beispiel einfacher anschaulich vorstellen, wenn man Polarkoordinaten $x = r\cos\phi$, $y = r\sin\phi$ verwendet. Für $r > 0$, $\phi \in \mathbb{R}$ gilt

$$f(r\cos\phi, r\sin\phi) = \frac{r^2\cos\phi\sin\phi}{r^2(\cos^2\phi + \sin^2\phi)} = \cos\phi\sin\phi$$

Die Werte von f sind also entlang von Strahlen durch den Nullpunkt konstant, hängen aber von der Richtung des Strahls ab.
Partielle Ableitungen vertauschen unter recht allgemeinen Voraussetzungen mit dem Integral. Wir besprechen im Moment nur eine einfache Variante davon, später in der Analysis 3 jedoch eine weitgehende Verallgemeinerung davon.

Lemma 2.7 (Vertauschung von Integral und partieller Ableitung) *Ist $f : [a,b]\times]c,d[\to \mathbb{C}$ (mit $a < b$ und $c < d$) eine stetige, nach dem 2. Argument partiell differenzierbare Abbildung mit stetiger Ableitung $D_2 f$, so gilt für alle $y \in]c,d[$:*

$$\frac{d}{dy}\int_a^b f(x,y)\,dx = \int_a^b \frac{\partial}{\partial y} f(x,y)\,dx$$

Beweis: Da $D_2 f$ stetig ist, ist es auch gleichmäßig stetig auf jedem *kompakten* Rechteck $[a,b]\times[c',d'] \subset [a,b]\times]c,d[$. Gegeben $y \in]c,d[$, folgt nach dem Mittelwertsatz der Differentialrechnung die folgende *gleichmäßige* Konvergenz:

$$\sup_{x\in[a,b]} \left| \frac{f(x,t)-f(x,y)}{t-y} - \frac{\partial}{\partial y} f(x,y) \right| \xrightarrow{t\to y} 0$$

und daher

$$\begin{aligned}\frac{d}{dy}\int_a^b f(x,y)\,dx &= \lim_{t\to y}\frac{1}{t-y}\left(\int_a^b f(x,t)\,dx - \int_a^b f(x,y)\,dx\right)\\ &= \lim_{t\to y}\int_a^b \frac{f(x,t)-f(x,y)}{t-y}\,dx = \int_a^b \lim_{t\to y}\frac{f(x,t)-f(x,y)}{t-y}\,dx\\ &= \int_a^b \frac{\partial}{\partial y} f(x,y)\,dx.\end{aligned}$$

□

Übung 2.8 Berechnen Sie den Gradienten der folgenden Funktion $f : \mathbb{R}^2\times\mathbb{R}^+ \to \mathbb{R}$, $f(a,b,c) = \int_a^b e^{-cx^2}\,dx$. Im Ergebnis darf zwar ein Integral, aber keine Ableitung stehen.

Für höhere partielle Ableitungen, also für partielle Ableitungen von partiellen Ableitungen gibt es wieder mehrere Notationen:

$$D_i D_j f(x) = \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} f(x) = \frac{\partial^2}{\partial x_i\,\partial x_j} f(x) = D^2_{i,j} f(x)$$

(analog für Ableitungen noch höherer Stufe). Für die m-te partielle Ableitung nach der j-ten Koordinate schreibt man auch

$$D_j^m f(x) = \frac{\partial^m}{\partial {x_j}^m} f(x) = \underbrace{\frac{\partial}{\partial x_j}\cdots\frac{\partial}{\partial x_j}}_{m \text{ Ableitungen}} f(x)$$

Eine Abbildung f wird m-fach partiell differenzierbar genannt, wenn alle partiellen Ableitung bis zur m-ten Stufe existieren. Sind diese partiellen Ableitungen sogar stetig, so nennt man sie m-fach stetig partiell differenzierbar. Eine beliebig oft stetig partiell differenzierbare Abbildung nennen wir auch *glatt*. Der Raum der m-fach stetig partiell differenzierbaren Funktionen auf U mit Werten in $\mathbb{K}\in\{\mathbb{R},\mathbb{C}\}$ wird mit $C^m(U,\mathbb{K})$ bezeichnet; der analoge Raum der glatten Funktionen mit $C^\infty(U,\mathbb{K})$.

Definition 2.9 Ist $f : U \to \mathbb{C}$ eine in $x \in U$ zweimal partiell differenzierbare Funktion, $U \subseteq \mathbb{R}^n$, so wird die Matrix aus den partiellen Ableitungen

$$D^2 f(x) := (D_i D_j f(x))_{i,j=1,\ldots,n} \in \mathbb{C}^{n\times n}$$

die *Hessematrix* von f in x genannt. Ihre Spur, also die Summe der Diagonaleinträge, heißt *Laplaceoperator* und wird mit

$$\Delta f(x) := \sum_{j=1}^{n} \frac{\partial^2}{\partial {x_j}^2} f(x)$$

bezeichnet.

Veranschaulichung des Laplaceoperators in $\mathbb{R}^2$*:* Stellen wir uns den Graphen von f als Fläche vor, so gibt der Laplaceoperator $\Delta f(x)$ an, ob die Fläche in x "im Mittel" eher nach oben ($\Delta f(x) > 0$) oder eher nach unten ($\Delta f(x) < 0$) gekrümmt ist. Stellen wir uns den Graphen als eine in einen horizontalen Rahmen eingespannte, ein wenig deformierte Gummimembran vor, so bedeutet $\Delta f(x) < 0$, dass eine Kraft die Membran an der Stelle x nach oben drückt.

Beispiel 2.10 Über $\mathbb{R}^n$ gilt:

$$\Delta(\|x\|^2) = \sum_{j=1}^{n} \frac{\partial^2}{\partial {x_j}^2} \sum_{k=1}^{n} x_k^2 = \sum_{j=1}^{n} 2 = 2n.$$

Insbesondere ist $\Delta(x^2 + y^2) = 4$ (nach oben geöffnetes Rotationsparaboloid, nach oben gekrümmt), $\Delta(-x^2 - y^2) = -4$ (nach unten geöffnetes Rotationsparaboloid, nach unten gekrümmt), $\Delta(x^2 - y^2) = 0$ (Hyperboloid, sowohl nach oben als auch nach unten gekrümmt) .

Übung 2.11 Es sei $f : \mathbb{R}^d \setminus \{0\} \to \mathbb{R}$, $f(x) = \|x\|_2^{2-d}$. Beweisen Sie $\Delta f = 0$. Insbesondere gilt in drei Dimensionen

$$\Delta \frac{1}{\|x\|_2} = 0.$$

Übung 2.12 Gegeben sei die Funktion

$$f : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}, \quad f(x,t) = \frac{1}{\sqrt{t^n}} e^{-\frac{\|x\|_2^2}{2t}},$$

Zeigen Sie, dass f die "Wärmeleitungsgleichung"

$$\frac{\partial}{\partial t} f(x,t) = \frac{1}{2} \Delta f(x,t)$$

löst, wobei sich hier der Laplaceoperator nur auf $x = (x_1, \ldots, x_n)$, nicht jedoch auf t bezieht. Versuchen Sie sich im Fall $n = 1$ anschaulich vorzustellen, dass f "zerfließende Gaußsche Glocken" beschreibt, indem Sie $x \mapsto f(x,t)$ für verschiedene t grob skizzieren.

Wie schon der Name besagt, beschreibt die Wärmeleitungsgleichung die zeitliche Temperaturentwicklung in einem wärmeleitenden Material: Mit der Interpretation, dass $f(x,t)$ die Temperatur am Ort x zur Zeit t beschreibt, bedeutet die Wärmeleitungsgleichung grob gesagt: "lokaler Temperaturanstieg, wenn die Umgebung im Mittel wärmer als der betrachtete Punkt ist". Diese Gleichung und Varianten davon modellieren auch Diffusionsvorgänge, wie sie z.B. bei der Entwicklung von Marktpreisen auftreten. Varianten der Wärmeleitungsgleichung spielen daher in der Finanzmathematik bei der Bewertung von Finanzderivaten eine wichtige Rolle.
Partielle Ableitungen können in Ausnahmefällen *nicht* vertauscht werden. Hierzu ein
Gegenbeispiel:
Ist

$$f:\mathbb{R}^2\to\mathbb{R},\quad f(x,y)=\begin{cases} xye^{-x^2/y^2} & \text{für } y\neq 0,\\ 0 & \text{für } y=0,\end{cases}$$

so gilt für $x\in\mathbb{R}$ und $y\neq 0$:

$$D_1f(x,y)=ye^{-x^2/y^2}-2\frac{x^2}{y}e^{-x^2/y^2},$$

aber auch $D_1f(x,0)=0$, also $D_1f(0,y)=y$. Es folgt $D_2D_1f(0,0)=1$. Andererseits gilt für alle festen $x\in\mathbb{R}$: $|f(x,y)|=|xye^{-x^2/y^2}|\ll|y|$ für $y\to 0$, also $D_2f(x,0)=0$. Es folgt

$$D_1D_2f(0,0)=0\neq 1=D_2D_1f(0,0).$$

Die partiellen Ableitungen vertauschen in diesem Gegenbeispiel also *nicht*.
Allerdings ist dieses Gegenbeispiel eine *untypische Ausnahme*. Es gilt nämlich:

Lemma 2.13 (Lemma von Schwarz zur Vertauschbarkeit partieller Ableitungen)
Es seien $n\geq 2$, $U\subseteq\mathbb{R}^n$ offen, $x\in U$, $f:U\to\mathbb{R}$ und $i,j\in\{1,\ldots,n\}$. Wenn die partiellen Ableitungen D_iD_jf und D_jD_if über U existieren und in x stetig sind, so gilt $D_iD_jf(x)=D_jD_if(x)$.
Etwas vergröbert formuliert: Ist f zweimal stetig partiell differenzierbar, so ist die Hessematrix D^2f symmetrisch, d.h. für ihre Transponierte gilt $(D^2f)^t=D^2f$.

Beweis: Für $i=j$ ist nichts zu zeigen. Weil alle Koordinaten außer x_i und x_j festgehalten werden, dürfen wir o.B.d.A. $n=2$, $i=1$ und $j=2$ annehmen. Wir schreiben nun (x,y) statt x.[12] Wir betrachten eine beliebige Umgebung $V\subseteq U$ von (x,y) und nehmen ein Rechteck mit den Ecken (x,y), (x',y), (x',y'), (x,y'), das ganz zu V gehört, wobei $x'>x$ und $y'>y$. Nach dem Mittelwertsatz der Differentialrechnung, angewandt in 1-Richtung, gibt es ein ξ zwischen x und x', für das gilt:

$$\begin{aligned}[f(x',y')-f(x',y)]-[f(x,y')-f(x,y)] &= \frac{\partial}{\partial\xi}[f(\xi,y')-f(\xi,y)](x'-x)\\ &= D_1f(\xi,y')-D_1f(\xi,y)](x'-x)\end{aligned}$$

[12]Man lasse sich nicht durch den Typ- und Bedeutungswechsel von x verwirren.

Nochmal mit dem Mittelwertsatz der Differentialrechnung, diesmal angewandt in 2-Richtung, erhalten wir für geeignetes η zwischen y und y':

$$D_1 f(\xi, y') - D_1 f(\xi, y) = D_2 D_1 f(\xi, \eta)(y' - y)$$

Zusammen erhalten wir:

$$\exists \xi \in [x, x'] \; \exists \eta \in [y, y'] : \; \frac{f(x', y') - f(x', y) - f(x, y') + f(x, y)}{(x' - x)(y' - y)} = D_2 D_1 f(\xi, \eta).$$

Mit vertauschten Rollen von “x” und “y” erhalten wir ebenso für geeignetes η' zwischen y und y' und geeignetes ξ' zwischen x und x':

$$\begin{aligned}
&[f(x', y') - f(x, y')] - [f(x', y) - f(x, y)] = \frac{\partial}{\partial \eta'}[f(x', \eta') - f(x, \eta')](y' - y) \\
&= D_2 f(x', \eta') - D_1 f(x, \eta')](y' - y) = D_1 D_2 f(\xi', \eta')(x' - x)(y' - y)
\end{aligned}$$

und daher zusammen

$$D_2 D_1 f(\xi, \eta) = \frac{f(x', y') - f(x', y) - f(x, y') + f(x, y)}{(x' - x)(y' - y)} = D_1 D_2 f(\xi', \eta').$$

Damit ist gezeigt, dass es für jede Umgebung $V \subseteq U$ von (x, y) zwei Punkte $(\xi, \eta), (\xi', \eta') \in V$ mit $D_2 D_1 f(\xi, \eta) = D_1 D_2 f(\xi', \eta')$ gibt. Aus der vorausgesetzten Stetigkeit von $D_2 D_1 f$ und von $D_1 D_2 f$ in (x, y) folgt damit $D_2 D_1 f(x, y) = D_1 D_2 f(x, y)$.

$\square$

Multiindexnotation für höhere partielle Ableitungen. Da es bei m-fach stetig partiell differenzierbaren Funktionen f von n Variablen $(x_1, \ldots, x_n)$ nicht auf die Reihenfolge der partiellen Ableitungen ankommt, liegt es nahe, nur die Anzahl der Ableitungen nach jeder Variablen und nicht deren Reihenfolge zu notieren: Gegeben eine Liste $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$ von Ableitungsstufen mit $|\alpha| := \sum_{j=1}^n \alpha_j \leq m$, setzen wir

$$D^\alpha f := \frac{\partial^{\alpha_1}}{\partial {x_1}^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial {x_2}^{\alpha_2}} \cdots \frac{\partial^{\alpha_n}}{\partial {x_n}^{\alpha_n}} f.$$

Die Zahl $|\alpha|$ nennen wir auch den *Grad* des Differentialoperators D^α. Hier ein Beispiel für diese Notation für eine 3-fach stetig partiell differenzierbare Abbildung $f : \mathbb{R}^3 \to \mathbb{R}$:

$$D^{(1,0,2)} f(x, y, z) = \frac{\partial}{\partial x} \frac{\partial^2}{\partial z^2} f(x, y, z).$$

2.3 Differenzierbarkeit und Ableitung

Die Kernidee der Differentialrechnung ist es, eine Funktion nahe an einer Stelle durch eine affin-lineare Funktion zu approximieren, also zu "*linearisieren*". Diese Idee kennen Sie schon aus der Analysis 1, wo wir differenzierbare Funktionen $f : \mathbb{R} \to \mathbb{R}$ nahe bei $x \in \mathbb{R}$ durch ihre Tangente $\tilde{x} \mapsto f(x) + f'(x)(\tilde{x} - x)$ approximiert haben:

$$f(\tilde{x}) = f(x) + f'(x)(\tilde{x} - x) + o(\tilde{x} - x) \text{ für } \tilde{x} \to x.$$

Die Idee der Linarisierung lässt sich auch auf höherdimensionale Situationen übertragen und ist dort ebenfalls von fundamentaler Bedeutung:

Definition 2.14 (Ableitung, Differential) Es seien $(V, \|\cdot\|)$ und $(W, \|\cdot\|)$ zwei normierte Räume, $V \neq \{0\}$, $U \subseteq V$ offen und $x \in U$. Eine Abbildung $f : U \to W$ wird *differenzierbar* an der Stelle x genannt, wenn es eine stetige lineare Abbildung $L : V \to W$ gibt, für die gilt:

$$\lim_{\tilde{x} \to x} \frac{\|f(\tilde{x}) - f(x) - L(\tilde{x} - x)\|}{\|\tilde{x} - x\|} = 0. \qquad (32)$$

Die stetige lineare Abbildung L ist dann eindeutig bestimmt und wird *Ableitung*, *Linearisierung* oder auch *Differential* von f an der Stelle x genannt und mit

$$df_x := df(x) := L$$

bezeichnet. Die Abbildung f heißt *differenzierbar*, wenn sie an jeder Stelle differenzierbar ist.

Mit anderen Worten gesagt: f ist differenzierbar in x mit Ableitung $L \in \mathcal{B}(V, W)$, wenn es eine in 0 stetige Abbildung $R : U - x \to W$ mit $R(0) = 0$ gibt, so dass für alle $h \in U - x$ gilt:

$$f(x + h) = f(x) + L(h) + \|h\| R(h).$$

Die Abbildung f heißt *stetig differenzierbar*, wenn sie differenzierbar ist und die Ableitung $df : U \to \mathcal{B}(V, W)$, $x \mapsto df_x$ stetig ist, wobei der Raum $\mathcal{B}(V, W)$ aller stetigen linearen Abbildungen $L : V \to W$ mit der zugehörigen Operatornorm versehen wird.

Der **wichtigste Fall** ist $V = \mathbb{R}^m$ und $W = \mathbb{R}^n$. In diesem Fall kommt es auf die Wahl der Normen nicht an, da alle Normen auf $\mathbb{R}^m$ bzw. $\mathbb{R}^n$ äquivalent sind. Auch ist in diesem Fall *jede* lineare Abbildung $L : \mathbb{R}^m \to \mathbb{R}^n$ stetig, so dass man hier auf die Forderung der Stetigkeit von L verzichten kann. In diesem Fall wird L auch durch Multiplikation mit einer Matrix $A \in \mathbb{R}^{n \times m}$ (m Spalten, n Zeilen) beschrieben:

$$L(y) = Ay \text{ für alle } y \in \mathbb{R}^m.$$

Ist $L = df_x$ das Differential von f bei x, so wird die zugehörige Matrix A die *Jacobimatrix* von f an der Stelle x genannt und mit $Df(x)$, Df_x oder (für $m > 1$ selten) mit $f'(x)$ bezeichnet.

Gelegentlich wird es zweckmäßig sein, auf die Bedingung, dass U offen ist, zu verzichten, z.B. bei der Ableitung an Punkten am Rand eines Halbraums, in Analogie zur links- und rechtsseitigen Ableitung in der Analysis 1 am Rand von Intervallen. Im Fall, dass es eine eindeutige Ableitung gibt, verwenden wir auch in diesem Fall die gleichen Notationen.
Für $m = n = 1$ stimmt die Jacobimatrix mit der Ableitung aus der Analysis 1 überein, nun aufgefasst als 1×1-Matrix.

Zur Eindeutigkeit der Ableitung df_x. Ist $L : V \to W$ eine stetige lineare Abbildung, für die (32) gilt, so folgt für alle $h \in V \setminus \{0\}$:

$$0 = \lim_{t\to 0} \frac{1}{t\|h\|}[f(x+th) - f(x) - L(th)] = \frac{1}{\|h\|}\left(\lim_{t\to 0}\frac{1}{t}[f(x+th) - f(x)] - L(h)\right),$$

also

$$L(h) = \lim_{t\to 0}\frac{1}{t}[f(x+th) - f(x)].$$

Die Abbildung $L = df_x$ wird also durch f eindeutig bestimmt.

Lokale affin-lineare Approximation von f. Die affin-lineare Abbildung

$$\mathbb{R}^m \ni \tilde{x} \mapsto f(x) + df_x(\tilde{x} - x) = f(x) + Df_x \cdot (\tilde{x} - x)$$

approximiert also die Funktion f wie folgt:

$$f(\tilde{x}) = f(x) + Df_x \cdot (\tilde{x} - x) + o(\|\tilde{x} - x\|) \text{ für } \tilde{x} \to x \tag{33}$$

wobei das Landau-Symbol $o(\|\tilde{x} - x\|)$ analog zur Verwendung in der Analysis 1 für einen Term $R(\tilde{x})$ steht, der $\|R(\tilde{x})\|/\|\tilde{x} - x\| \to 0$ für $\tilde{x} \to x$ erfüllt. Beim Umgang mit Matrixoperationen, insbesondere in der Darstellung (33), fassen wir die Elemente von $\mathbb{R}^m$ und von $\mathbb{R}^n$ als *Spaltenvektoren* auf; insbesondere wird $\tilde{x} - x$ in (33) als *Spaltenvektor* aufgefasst. Wir identifizieren also $\mathbb{R}^n \equiv \mathbb{R}^{n\times 1}$. Die obige affin-lineare Abbildung verallgemeinert die Beschreibung von Tangenten aus der Analysis 1 auf multidimensionale Situationen.

Beispiel 2.15 Ist $f : V \to W$ linear und stetig, so gilt $df_x = f$ für alle $x \in V$. In der Tat ist $\|f(\tilde{x}) - f(x) - f(\tilde{x} - x)\| = 0$ für alle $\tilde{x}, x = 0$. Anders gesagt: Die Linearisierung einer stetigen linearen Funktion f ist f selbst. Im Spezialfall $V = \mathbb{R}^m$, $W = \mathbb{R}^n$ bedeutet das: Für jede Matrix $A \in \mathbb{R}^{n\times m}$ besitzt die lineare Abbildung $L_A : \mathbb{R}^m \to \mathbb{R}^n$, $L_A(x) = Ax$ die Jacobimatrix $DL_A(x) = A$ an jeder Stelle $x \in \mathbb{R}^m$. Das verallgemeinert die Ableitungsregel $\frac{d}{dx}ax = a$ aus der Analysis 1.
Ein besonders einfaches, aber trotzdem wichtiges Beispiel erhält man im Fall $V = \mathbb{R}^m$ für die kanonischen Projektionen $x = (x_1, \ldots, x_m) \mapsto x_j$, $j = 1, \ldots, m$. In der Differentialrechnung bezeichnet man diese Projektionen oft mit dem gleichen Namen wie die Koordinaten selbst, also mit $x_j(x_1, \ldots, x_m) = x_j$. Insbesondere ist dx_j an jeder Stelle x die j-te kanonische Projektion:

$$dx_j(x) : (y_1, \ldots, y_m) \mapsto y_j.$$

Die zugehörige Jacobimatrix ist konstant gleich dem j-ten kanonischen Einheits*zeilenvektor*

$$Dx_j(x) = e_j^t = (0, \ldots 0, \underbrace{1}_{\text{Stelle } j}, 0, \ldots, 0) \in \mathbb{R}^{1\times m}$$

Beispiel 2.16 (Verallgemeinerung der Ableitungsregel $\frac{d(x^2)}{dx} = 2x$.) Es sei $(V, \langle \cdot, \cdot \rangle)$ ein Prähilbertraum über $\mathbb{R}$ und $f : V \to \mathbb{R}$, $f(x) = \|x\|^2 = \langle x, x \rangle$. Dann ist f differenzierbar in jeder Stelle $x \in V$ mit der Ableitung

$$df_x = 2\langle x, \cdot \rangle \tag{34}$$

ausführlich geschrieben $df_x(y) = 2\langle x, y \rangle$ für $y \in V$. Man beachte, dass diese Ableitung eine *Linearform* auf V ist, also eine lineare Abbildung von V in den Grundkörper $\mathbb{R}$. In der Tat: Aus der Cauchy-Schwarz-Ungleichung folgt

$$\forall x, y \in V: \ |2\langle x, y \rangle| \leq 2\|x\|\|y\|,$$

also ist die lineare Abbildung $2\langle x, \cdot \rangle : V \to \mathbb{R}$ stetig. Weiter gilt für $\tilde{x} \in V$:

$$\begin{aligned} &f(\tilde{x}) - f(x) - 2\langle x, \tilde{x} - x \rangle = \langle \tilde{x}, \tilde{x} \rangle - \langle x, x \rangle - 2\langle x, \tilde{x} - x \rangle \\ &= \langle \tilde{x}, \tilde{x} \rangle - 2\langle x, \tilde{x} \rangle + \langle x, x \rangle = \|\tilde{x} - x\|^2 = o(\|\tilde{x} - x\|) \end{aligned} \tag{35}$$

für $\tilde{x} \to x$, also

$$\frac{|f(\tilde{x}) - f(x) - 2\langle x, \tilde{x} - x \rangle|}{\|\tilde{x} - x\|} \xrightarrow{\tilde{x} \to x} 0.$$

Im Spezialfall $V = \mathbb{R}^m$, versehen mit dem euklidischen Skalarprodukt, lautet die Jacobimatrix von f also

$$Df(x) = 2(x_1, \ldots, x_n) = 2x^t,$$

denn $2\langle x, y \rangle = 2x^t y$ für $y \in \mathbb{R}^m$. Die Transposition ist nötig, da wir den *Vektor* x durch einen *Spaltenvektor* darstellen, die *Linearform* $df(x)$ jedoch durch den *Zeilenvektor* $Df(x)$.

Die Jacobimatrix wird aus den partiellen Ableitungen der Komponenten wie folgt gebildet:

Lemma 2.17 *Gegeben sei eine Funktion $f : \mathbb{R}^m \supseteq U \to \mathbb{R}^n$, $f(\tilde{x}) = (f_1(\tilde{x}), \ldots, f_n(\tilde{x}))$, die differenzierbar in $x = (x_1, \ldots, x_m) \in U$ ist. Dann sind alle Komponentenfunktionen $f_i : U \to \mathbb{R}$, $i = 1, \ldots, n$ in x partiell differenzierbar, und es gilt:*

$$Df(x) = \left(\frac{\partial f_i}{\partial x_j}(x)\right)_{i=1,\ldots,n;\, j=1,\ldots,m} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \ldots & \frac{\partial f_1}{\partial x_m}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & \ldots & \frac{\partial f_2}{\partial x_m}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(x) & \frac{\partial f_n}{\partial x_2}(x) & \ldots & \frac{\partial f_n}{\partial x_m}(x) \end{pmatrix}$$

Beweis: O.B.d.A. arbeiten wir mit der Norm $\|\cdot\| = \|\cdot\|_\infty$ auf $\mathbb{R}^m$ bzw. $\mathbb{R}^n$. Es gilt für $i = 1, \ldots, n$ und $j = 1, \ldots, m$ für $t \in \mathbb{R} \setminus \{0\}$ nahe bei 0:

$$\begin{aligned}
&|t^{-1}[f_i(x+te_j) - f_i(x)] - (Df_x)_{ij}| = |(t^{-1}[f(x+te_j) - f(x) - Df_x \cdot te_j]_i| \\
&\le |t^{-1}|\, \|f(x+te_j) - f(x) - Df_x \cdot te_j]\| = \frac{\|f(x+te_j) - f(x)] - df_x(te_j)\|}{\|te_j\|} \overset{t\to 0}{\longrightarrow} 0.
\end{aligned}$$

Das bedeutet

$$\frac{\partial f_i}{\partial x_j}(x) = (Df_x)_{ij}.$$

□

Kürzen wir ab: $y_i = f_i(x)$, $\tilde{y}_i = f_i(\tilde{x})$, so können wir Gleichung (33) in Komponentenschreibweise so schreiben:

$$\boxed{\tilde{y}_i - y_i = \sum_{j=1}^{m} \frac{\partial y_i}{\partial x_j} \cdot (\tilde{x}_j - x_j) + o(\tilde{x}_j - x_j)}$$

für $i = 1, \ldots, n$ und $\tilde{x} \to x$.

Warnung: Ein häufiger Anfängerfehler ist es, Zeilen- und Spaltenindizes in der Jacobimatrix zu verwechseln, insbesondere im Fall $m = n$, also die Jacobimatrix mit ihrer Transponierten zu verwechseln. Merken Sie sich zur Vermeidung dieses Fehlers für die Einträge $a_{ij} = \partial y_i / \partial x_j$ der Jacobimatrix:

Linker Index i = Zeilenindex = Index der Ausgabevariablen
Rechter Index j = Spaltenindex = Index der Eingabevariablen

Beispiel 2.18 *Für die Transformation* $f : \mathbb{R}^2 \to \mathbb{R}^2$, $(x, y) = f(r, \phi) := (r\cos\phi, r\sin\phi)$ *gilt*

$$Df(r,\phi) = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \phi} \end{pmatrix} = \begin{pmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{pmatrix}.$$

Übung 2.19 Berechnen Sie die Jacobimatrix $Df(r, \theta, \phi)$ der Transformation von Kugelkoordinaten nach kartesischen Koordinaten

$$f : \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^3, \quad f(r, \theta, \phi) = (r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta).$$

Vektoren und Linearformen als Ableitungen

1. **Abbildungen vom Typ $\mathbb{R} \to \mathbb{R}^n$.** Eine differenzierbare Abbildung

$$f : \mathbb{R} \supseteq U \to \mathbb{R}^n, \quad f(t) = (f_i(t))_{i=1,\ldots,n}$$

können wir uns als eine durch die Zeit parametrisierte Kurve im Raum vorstellen. Die Jacobimatrix ist der aus der Ableitung der Komponenten gebildete *Spaltenvektor*

$$f'(t) = \begin{pmatrix} f_1'(t) \\ \vdots \\ f_n'(t) \end{pmatrix},$$

den wir uns als *Geschwindigkeitsvektor* zur Zeit t vorstellen. Die Linearisierung von f bei t ist die *vektorwertige* Abbildung

$$df_t : \mathbb{R} \to \mathbb{R}^n, \quad s \mapsto sf'(t).$$

2. **Abbildungen vom Typ $\mathbb{R}^m \to \mathbb{R}$.** Im Spezialfall von differenzierbaren Abbildungen $f : \mathbb{R}^m \supseteq U \to \mathbb{R}$ ist die Ableitung $df_x : \mathbb{R}^m \to \mathbb{R}$ eine *Linearform*; sie wird durch den *Zeilenvektor*

$$Df(x) = \left(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_m}(x)\right)$$

dargestellt, in Kurznotation:

$$\boxed{Df = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_m}\right)}$$

In Differentialnotation können wir das Gleiche in der Form

$$\boxed{df = \sum_{j=1}^{m} \frac{\partial f}{\partial x_j}\, dx_j}$$

schreiben; erinnern Sie sich, dass dx_j die Projektion auf die j-te Koordinate bezeichnet. Die in der Analysis 1 manchmal zweckmäßige symbolische Notation $dy = f'(x)\, dx$ für $f : \mathbb{R} \to \mathbb{R}$, $y = f(x)$ bekommt damit einen rigorosen Sinn. Ein Vergleich des Gradienten

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_m} \end{pmatrix}$$

mit der Jacobimatrix Df zeigt: *Der Gradient ist die Transponierte der Jacobimatrix:* $\nabla f = (Df)^t$. In Differentialnotation kann man das mit dem euklidischen Skalarprodukt $\langle \cdot, \cdot \rangle$ auch so schreiben: $df_x(y) = \langle \nabla f(x), y \rangle$ für $x \in U$, $y \in \mathbb{R}^m$, oder kurz

$$\boxed{df = \langle \nabla f, \cdot \rangle}$$

Der Übersetzung vom Gradienten zum Differential liegt also der vom euklidischen Skalarprodukt induzierte Isomorphismus $\mathbb{R}^n \to (\mathbb{R}^n)'$, $a \mapsto \langle a, \cdot \rangle$ vom Raum $\mathbb{R}^n$ in seinen Dualraum $(\mathbb{R}^n)'$ zugrunde.

Das Wechselspiel von Vektoren und Linearformen in der Differentialrechnung mehrerer Variablen mag vielleicht beim ersten Kennenlernen verwirrend wirken; darum folgende Tabelle als Merkregel:

Typ von f	**Ableitung** $df(x)$	**dargestellt durch** $Df(x)$
$\mathbb{R} \to \mathbb{R}^n$	Vektor (genauer: vektorwertige Abb.)	Spaltenvektor
$\mathbb{R}^m \to \mathbb{R}$	Linearform (Differentialform)	Zeilenvektor

Das Lemma 2.17 hat den Nachteil, dass es kein Kriterium für Differenzierbarkeit liefert, denn es hat die Differenzierbarkeit von f als Voraussetzung.
In der Tat folgt aus partieller Differenzierbarkeit *nicht* die Differenzierbarkeit, wie folgendes Gegenbeispiel zeigt:

Beispiel 2.20 *Die Abbildung*

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

ist partiell differenzierbar, auch in $(0,0)$*. Für die Richtungsableitung in* $(0,0)$ *in Richtung* $a \in \mathbb{R}^2$ *gilt*

$$D_a f(0,0) = \frac{d}{dt} f(ta)\Big|_{t=0} = \frac{d}{dt} t f(a)\Big|_{t=0} = f(a),$$

was jedoch keine lineare Funktion von a *ist. Also ist* f *nicht in* 0 *differenzierbar.*

Hier ist ein einfaches hinreichendes Kriterium für Differenzierbarkeit:

Lemma 2.21 (Stetige partielle Differenzierbarkeit $\Rightarrow$ Differenzierbarkeit) *Ist* $U \subseteq \mathbb{R}^m$ *offen und* $f : U \to \mathbb{R}^n$ *partiell differenzierbar mit in* $x \in U$ *stetigen partiellen Ableitungen aller Komponenten, so ist* f *in* x *differenzierbar. Insbesondere ist jede stetig partiell differenzierbare Abbildung* $f : U \to \mathbb{R}^n$ *differenzierbar.*

Beweis: Es genügt, jede Komponente f_i von f einzeln zu betrachten; daher dürfen wir o.B.d.A. $n = 1$ annehmen. Gegeben $\epsilon > 0$, wählen wir mit der Stetigkeit von ∇f in x ein $\delta > 0$ so klein, dass für alle $\tilde{x} \in \mathbb{R}^m$ mit $\|\tilde{x} - x\|_\infty < \delta$ gilt: $\tilde{x} \in U$ und $\|\nabla f(\tilde{x}) - \nabla f(x)\|_\infty < \epsilon/m$. Gegeben solch ein $\tilde{x}$, definieren wir Vektoren $z^j = (z_i^j)_{1 \le i \le m} \in \mathbb{R}^m$ für $j = 0, \ldots, m$ durch

$$z_i^j = \begin{cases} x_i & \text{für } j < i \le m, \\ \tilde{x}_j & \text{für } 1 \le i \le j. \end{cases}$$

Insbesondere ist $z^0 = x$, $z^m = \tilde{x}$ und $z^j - z^{j-1} = (\tilde{x}_j - x_j) e_j$ sowie $\|z^j - x\|_\infty < \delta$ für $j = 1, \ldots, m$. Nach dem Mittelwertsatz der Differentialrechnung folgt für ein $t_j \in]0,1[$:

$$f(z^j) - f(z^{j-1}) = D_j f((1-t_j) z^{j-1} + t_j z^j)\,(\tilde{x}_j - x_j)$$

also mit der Abkürzung $y^j := (1-t_j)z^{j-1} + t_j z^j \in U_\delta^{\|\cdot\|_\infty}(x)$:

$$f(z^j) - f(z^{j-1}) - D_j f(x)\,(\tilde{x}_j - x_j) = [D_j f(y^j) - D_j f(x)]\,(\tilde{x}_j - x_j).$$

Summieren wir das über j, folgt

$$f(\tilde{x}) - f(x) - \langle \nabla f(x), \tilde{x} - x\rangle = f(z^m) - f(z^0) - \sum_{j=1}^m D_j f(x)\,(\tilde{x}_j - x_j)$$

$$= \sum_{j=1}^m [f(z^j) - f(z^{j-1}) - D_j f(x)\,(\tilde{x}_j - x_j)] = \sum_{j=1}^m [D_j f(y^j) - D_j f(x)]\,(\tilde{x}_j - x_j).$$

und daher

$$|f(\tilde{x}) - f(x) - \langle \nabla f(x), \tilde{x} - x\rangle\,| \le \sum_{j=1}^m \|\nabla f(y^j) - \nabla f(x)\|_\infty |\tilde{x}_j - x_j| \le m\frac{\epsilon}{m}\|\tilde{x} - x\|_\infty = \epsilon\|\tilde{x} - x\|_\infty.$$

Weil $\epsilon > 0$ beliebig war, zeigt das, dass f in x differenzierbar mit der Ableitung $df(x) = \langle \nabla f(x), \cdot\rangle$ ist.

□

Beispiel 2.22 (Abstraktion der Produktregel) Als ein etwas abstraktes Beispiel betrachten wir drei normierte Räume U, V, W (Normen mit $\|\cdot\|$ bezeichnet) und die Kompositionsabbildung

$$\mathrm{comp} : \mathcal{B}(V,W) \times \mathcal{B}(U,V) \to \mathcal{B}(U,W), \quad \mathrm{comp}(F,G) = F \circ G$$

wobei die Räume jeweils mit der Operatornorm und das kartesische Produkt mit dem Maximum der Operatornormen der beiden Einträge versehen wird. Wer will, kann sich Matrizen statt der Operatoren und das Matrixprodukt statt der Komposition vorstellen. Die Abbildung comp ist differenzierbar mit der Ableitung

$$d\,\mathrm{comp}_{(F,G)}(F', G') = F' \circ G + F \circ G'$$

für[13] $F, F' \in \mathcal{B}(V,W)$ und $G, G' \in \mathcal{B}(V,W)$. In der Tat:

$$\begin{aligned}&\|(F+F') \circ (G+G') - F \circ G - [F' \circ G + G' \circ F]\|_{U\to W} = \|F' \circ G'\|_{U\to W}\\ &\le \|F'\|_{V\to W}\|G'\|_{U\to V} = o(\max\{\|F'\|_{V\to W}, \|G'\|_{U\to V}\})\\ &\text{für } \max\{\|F'\|_{V\to W}, \|G'\|_{U\to V}\} \to 0.\end{aligned}$$

Zudem ist die Abbildung

$$\mathcal{B}(V,W) \times \mathcal{B}(U,V) \to \mathcal{B}(U,W), \quad (F', G') \mapsto F' \circ G + F \circ G'$$

stetig nach Lemma 1.63, denn sie ist linear und es gilt

$$\begin{aligned}&\|F' \circ G + F \circ G'\|_{U\to W} \le \|F' \circ G\|_{U\to W} + \|F \circ G'\|_{U\to W}\\ &\le \|F'\|_{V\to W}\|G\|_{U\to V} + \|F\|_{V\to W}\|G'\|_{U\to W}\\ &\le (\|G\|_{U\to V} + \|F\|_{V\to W})\max\{\|F'\|_{V\to W}, \|G'\|_{U\to W}\}.\end{aligned}$$

□

[13] F' und G' sind hier nur anschauliche Namen für Variablen, der $'$ steht hier nicht für “Ableitung”.

2.4 Veranschaulichung der Ableitung mit Tangentialräumen

Übung 2.23 *Betrachten Sie die Funktion* $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = x^2 + y^2$, *den Punkt* $(x_0, y_0) = (3,4)$ *mit dem Wert* $z_0 = f(x_0, y_0) = 25$ *und die Linearisierung* $g : \mathbb{R}^2 \to \mathbb{R}$, $g(x,y) = z_0 + df_{(x_0,y_0)}(x - x_0, y - y_0)$ *von* f *bei* (x_0, y_0). *Zeichnen Sie die Niveaulinien von* f *nahe bei* (x_0, y_0) *mit einem Zirkel und die Niveaugeraden von* g *nahe bei* (x_0, y_0) *mit einem Lineal, z.B. für die Niveaus* $z = 21, 22, 23, 24, 25, 26, 27, 28, 29$. *Beobachten Sie, wie die Niveaugeraden von* g *nahe bei* (x_0, y_0) *die Niveaukreise von* f *nahe bei* (x_0, y_0) *approximieren.*

Erinnern Sie sich, dass man m-dimensionale Gebilde im n-dimensionalen Raum, zum Beispiel Flächen im Anschauungsraum, auf drei verschiedene Weisen durch Abbildungen beschreiben kann. Durch Linearisierung erhält man daraus drei verschiedene Beschreibungen von *Tangentialräumen*, also zum Beispiel der Tangentialebene an eine Fläche in einem Punkt:

1. Ersetzt man im Graphen $G = \{(x, f(x)) | x \in U \subseteq \mathbb{R}^m\}$ einer differenzierbaren Abbildung f vom Typ $\mathbb{R}^m \to \mathbb{R}^{n-m}$ die Abbildung f durch ihre affin-lineare Näherung $g : \mathbb{R}^m \to \mathbb{R}$, $g(x) = f(x_0) + df_{x_0}(x - x_0)$ in einem Punkt $x_0 \in U$, so erhält man den *Tangentialraum*
$$T_{(x_0, f(x_0))} G := \{(x, g(x)) | x \in \mathbb{R}^m\}$$
von G im Punkt $(x_0, f(x_0)) \in G$, dargestellt als Graph. Z.B. für eine Fläche $G = \{(x, y, f(x,y)) | (x,y) \in U \subseteq \mathbb{R}^2\}$ im Raum $\mathbb{R}^3$ erhält man somit die Tangentialebene $T_{(x_0,y_0,z_0)} G$ an G im Punkt $(x_0, y_0, z_0) \in G$ (wobei $(x_0, y_0) \in U$ und $z_0 = f(x_0, y_0)$), dargestellt als Graph:
$$T_{(x_0,y_0,z_0)} G := \{(x, y, g(x,y)) | x, y \in \mathbb{R}\}$$
mit der affin-linearen Abbildung
$$g(x,y) = f(x_0, y_0) + Df(x_0, y_0) \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} = z_0 + D_1 f(x_0, y_0)(x - x_0) + D_2 f(x_0, y_0)(y - y_0).$$
Beispiel Nordhalbkugel, $f(x,y) = \sqrt{1 - x^2 + y^2}$,
$$df(x,y) = -\frac{x}{\sqrt{1 - x^2 + y^2}}\, dx - \frac{y}{\sqrt{1 - x^2 + y^2}}\, dy.$$
Die Tangentialebene in $(x_0, y_0, \sqrt{1 - x_0^2 - y_0^2})$ wird als Graph von
$$g(x,y) = \sqrt{1 - x_0^2 - y_0^2} - \frac{x_0}{\sqrt{1 - x_0^2 + y_0^2}}(x - x_0) - \frac{y_0}{\sqrt{1 - x_0^2 + y_0^2}}(y - y_0)$$
dargestellt.

2. Gegeben sei eine Parameterdarstellung eines m-dimensionalen Gebildes G in $\mathbb{R}^n$ als Wertebereich $G = f[U]$, $U \subseteq \mathbb{R}^m$ offen, mit einer injektiven differenzierbaren Abbildung $f : U \to \mathbb{R}^n$. Ersetzt man die Abbildung f durch ihre affin-lineare Näherung
$$g : \mathbb{R}^m \to \mathbb{R}^n, \quad g(x) = f(x_0) + df_{x_0}(x - x_0)$$
bei $x_0 \in U$, so erhält man den Tangentialraum $T_{f(x_0)}G = g[\mathbb{R}^m]$ an G in Parameterdarstellung, *falls* die *Regularitätsbedingung* $\mathrm{Rang}(df_{x_0}) = m$ erfüllt ist. Ist diese Regularitätsbedingung verletzt, ist $g[\mathbb{R}^m]$ nicht m-dimensional.
Z.B. für eine Fläche $G = \{f(s,t) | (s,t) \in U \subseteq \mathbb{R}^2\}$, $f = (f_1, f_2, f_3) : U \to \mathbb{R}^3$, im Raum: Hier erhält man die Tangentialebene $T_{f(s_0,t_0)}G$ in Parameterdarstellung:
$$T_{f(s_0,t_0)}G = \left\{ \left. \begin{pmatrix} f_1(s_0,t_0) \\ f_2(s_0,t_0) \\ f_3(s_0,t_0) \end{pmatrix} + Df(s_0,t_0) \cdot \begin{pmatrix} s - s_0 \\ t - t_0 \end{pmatrix} \right| s, t \in \mathbb{R} \right\}$$
Die Regularitätsbedingung $\mathrm{Rang}(df_{(s_0,t_0)}) = 2$ garantiert, dass tatsächlich eine Ebene parametrisiert wird, nicht etwa nur eine Gerade oder ein Punkt.

   Beispiel Erdkugeldarstellung mit Längen- und Breitenkreisen: Hier ist $f(\theta, \phi) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ mit
$$Df(\theta,\phi) = \begin{pmatrix} \cos\theta\cos\phi & -\sin\theta\sin\phi \\ \cos\theta\sin\phi & \sin\theta\cos\phi \\ -\sin\theta & 0 \end{pmatrix},$$
und man erhält
$$T_{f(\theta_0,\phi_0)}S^2 = \left\{ \left. \begin{pmatrix} \sin\theta_0\cos\phi_0 \\ \sin\theta_0\sin\phi_0 \\ \cos\theta_0 \end{pmatrix} + \begin{pmatrix} \cos\theta_0\cos\phi_0 & -\sin\theta_0\sin\phi_0 \\ \cos\theta_0\sin\phi_0 & \sin\theta_0\cos\phi_0 \\ -\sin\theta_0 & 0 \end{pmatrix} \begin{pmatrix} \theta - \theta_0 \\ \phi - \phi_0 \end{pmatrix} \right| \theta, \phi \in \mathbb{R} \right\}$$
als Beschreibung der Tangentialebene im Punkt $f(\theta_0, \phi_0)$ mit Breiten- und Längengrad (θ_0, ϕ_0). Man beachte, dass die Regularitätsbedingung $\mathrm{Rang}(df_{(\theta_0,\phi_0)}) = 2$ im Nord- und Südpol $(0, 0, \pm 1)$, also für $\cos\theta_0 = \pm 1$ verletzt ist; an diesen "Kartensingularitäten" erhält man in der Tat nicht die volle Tangentialebene.

3. Beschreibt man ein m-dimensionales Gebilde G in $\mathbb{R}^n$ als Niveaugebilde $G = f^{-1}[\{h\}]$ mit festem $h \in \mathbb{R}^{n-m}$ mit einer Abbildung f vom Typ $\mathbb{R}^n \to \mathbb{R}^{n-m}$, so erhält man den Tangentialraum $T_{x_0}G$ in einem Punkt $x_0 \in G$ als Niveau-h-Gebilde der affinen Linearisierung
$$g : \mathbb{R}^n \to \mathbb{R}^{n-m}, \quad g(x) = h + df_{x_0}(x - x_0),$$
also
$$T_{x_0}G = \{x \in \mathbb{R}^n | \, df_{x_0}(x - x_0) = 0\},$$
falls die *Regularitätsbedingung* $\mathrm{Rang}(df_{x_0}) = m - n$ erfüllt ist. Ist zum Beispiel G eine Fläche im Raum, die durch eine reelle Gleichung $f(x, y, z) = h$ mit festem

$h \in \mathbb{R}$ beschrieben wird, so wird die Tangentialebene im Punkt $(x_0, y_0, z_0) \in G$ als Lösungsraum der Gleichung

$$Df(x_0, y_0, z_0) \cdot \begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} = 0$$

beschrieben.
Beispiel Einheitssphäre S^2 als Lösungsraum der Gleichung

$$f(x, y, z) = x^2 + y^2 + z^2 = 1$$

Hier ist $df(x, y, z) = 2x\,dx + 2y\,dy + 2z\,dz$, also wird die Tangentialebene im Punkt $(x_0, y_0, z_0) \in S^2$ durch die Gleichung

$$2x_0(x - x_0) + 2y_0(y - y_0) + 2z_0(z - z_0) = 0$$

beschrieben. Geometrisch besagt das, dass der Tangentialvektor $(x-x_0, y-y_0, z-z_0)$ senkrecht auf dem Radialvektor (x_0, y_0, z_0) stehen muss, so wie man es anschaulich erwartet. Man beachte, dass die Regularitätsbedingung $\mathrm{Rang}(df(x_0, y_0, z_0)) = 1$, also $(x_0, y_0, z_0) \neq 0$, auf S^2 erfüllt ist, so dass diese Beschreibung der Sphäre singularitätenfrei ist. Allerdings wird die Regularitätsbedingung im Nullpunkt $(0, 0, 0)$ verletzt. Geometrisch ist das klar, denn die Gleichung $x^2 + y^2 + z^2 = 0$ beschreibt keine Fläche im Raum mehr, sondern nur mehr einen Punkt.

Man beachte, dass in allen drei Beispielen die Tangentialebenen einfach durch formales Einsetzen von Differenzen in die Differentiale entstehen. Die Äquivalenz dieser drei Beschreibungen des Tangentialraums und auch die Rolle der Regularitätsbedingungen können wir erst später formal beweisen, wenn uns mehr Hilfsmittel zur Verfügung stehen.

2.5 Die Kettenregel

Die folgende Kettenregel ist das *wichtigste Werkzeug der Differentialrechnung*:

Satz 2.24 (Kettenregel) *Es seien $(U, \|\cdot\|)$, $(V, \|\cdot\|)$ und $(W, \|\cdot\|)$ normierte Räume, $U' \subseteq U$ und $V' \subseteq V$ offen, und $f : U' \to V'$ und $g : V' \to W$ Abbildungen. Weiter sei f in einem gegebenen $x \in U'$ differenzierbar, und g in $f(x)$ differenzierbar. Dann ist auch $g \circ f : U' \to W$ differenzierbar in x, und es gilt:*

$$\boxed{d(g \circ f)_x = dg_{f(x)} \circ df_x}$$

In Worten:

Linearisierung der Komposition = Komposition der Linearisierungen

Der *wichtigste Fall* ist $U = \mathbb{R}^l$, $V = \mathbb{R}^m$ und $W = \mathbb{R}^n$ mit $l, m, n \in \mathbb{N}$. In diesem Fall können wir die Kettenregel mit Jacobimatrizen auch so schreiben:

$$\boxed{D(g \circ f)(x) = Dg(f(x)) \cdot Df(x)}$$

wobei "·" die Multiplikation von Matrizen bedeutet. Schreiben wir $x = (x_1, \ldots, x_l)$, $y = (y_1, \ldots, y_m) = f(x)$ und $z = (z_1, \ldots, z_n) = g(y)$, so können wir die Kettenregel in Komponentenschreibweise auch so formulieren: Für $i = 1, \ldots, n$ und $k = 1, \ldots, l$ gilt:

$$\boxed{\frac{\partial z_i}{\partial x_k} = \sum_{j=1}^{m} \frac{\partial z_i}{\partial y_j} \frac{\partial y_j}{\partial x_k}}$$

In dieser Notation bleibt es implizit, an welcher Stelle die partiellen Ableitungen ausgewertet werden sollen.

Beweis der Kettenregel. Aufgrund der Differenzierbarkeit von f in x können wir für $x + h \in U'$ schreiben:

$$f(x+h) = f(x) + df_x(h) + \|h\| R(h)$$

mit einer in 0 stetigen Funktion $R : U' - x \to V$ mit $R(0) = 0$. Ebenso können wir aufgrund der Differenzierbarkeit von g in $y := f(x)$ für $y + k \in V'$ schreiben:

$$g(y+k) = g(y) + dg_y(k) + \|k\| S(k)$$

mit einer in 0 stetigen Funktion $S : V' - y \to W$ mit $S(0) = 0$. Durch Einsetzen von $y + k = f(x + h)$, also von

$$k = k(h) := f(x+h) - f(x) = df_x(h) + \|h\| R(h),$$

folgt:

$$\begin{aligned} g(f(x+h)) &= g(f(x)) + dg_y(df_x(h) + \|h\| R(h)) + \|k(h)\| S(k(h)) \\ &= g(f(x)) + dg_y \circ df_x(h) + T(h) \end{aligned}$$

mit dem Restterm

$$T(h) = \|h\| dg_y(R(h)) + \|k(h)\| S(k(h)).$$

Es bleibt noch zu zeigen:

$$\lim_{h \to 0} \frac{\|T(h)\|}{\|h\|} = 0.$$

Dies folgt so: Weil dg_y und R in 0 stetig mit den Werten $R(0) = 0$ und $dg_y(0) = 0$ sind, folgt $\lim_{h\to 0} dg_y(R(h)) = 0$. Weiter können wir ein $C > 0$ wählen, so dass für alle h in einer geeigneten Umgebung von $0 \in V$ gilt: $\|k(h)\| \le C\|h\|$, denn mit der Operatornorm $\|df_x\|_{U \to V} < \infty$ von df_x folgt:

$$\|k(h)\| \le \|df_x(h)\| + \|h\| \|R(h)\| \le (\|df_x\|_{U \to V} + \|R(h)\|) \|h\|,$$

und $\|R(h)\|$ ist für in einer Umgebung von 0 beschränkt. Insbesondere erhalten wir $\lim_{h\to 0} k(h) = 0$ und daher auch $\lim_{h \to 0} S(k(h)) = 0$ wegen und der Stetigkeit von S in 0 und $S(0) = 0$. Es folgt die Behauptung: Für $h \neq 0$ in einer Umgebung von 0 gilt:

$$0 \le \frac{\|T(h)\|}{\|h\|} \le \|dg_y(R(h))\| + \frac{\|k(h)\|}{\|h\|} \|S(k(h))\| \le \|dg_y(R(h))\| + C \|S(k(h))\| \xrightarrow{h \to 0} 0.$$

$\square$

Beispiel 2.25 (Richtungsableitung) Es sei $U \subseteq \mathbb{R}^n$ offen, $I \subseteq \mathbb{R}$ ein offenes Intervall mit $0 \in I$ und $x \in U$. Ist $f : U \to \mathbb{R}$ differenzierbar in x und $k : I \to U$ eine parametrisierte differenzierbare Kurve mit $k(0) = x$, folgt für die Richtungsableitung aus der Kettenregel:

$$D(f \circ k)(0) = Df(x) \cdot Dk(0) = \langle \nabla f(x), k'(0) \rangle$$

Im Spezialfall einer Geraden $k(t) = x + at$, $a \in \mathbb{R}^n$, erhalten wir

$$D_a f(x) = Df_x \cdot a = \langle \nabla f(x), a \rangle = df_x(a).$$

Aus der Cauchy-Schwarz-Ungleichung folgt $D_a f(x) \leq \|\nabla f(x)\|_2 \|a\|_2$, mit Gleichheit, wenn $\nabla f(x)$ und a in die gleiche Richtung zeigen. Das liefert uns eine Veranschaulichung des Gradienten: Der Gradient steht senkrecht auf dem Niveaugebilde und zeigt in Richtung des stärksten Anstiegs. Sein Betrag ist umso größer, je steiler das Gelände ist. Läuft man also bei einer Bergwanderung immer mit konstanter Geschwindigkeit in Richtung des Gradienten, so kommt man am schnellsten nach oben (auch wenn das nicht immer der gemütlichste Wanderweg ist).

Wichtiger Rechentrick zur Anwendung der Kettenregel: Hat man einen differenzierbaren Term der Gestalt $f(t, t, \ldots, t)$, in dem eine reellwertige Variable t an n verschiedenen Stellen auftritt, so erhält man die Ableitung $\frac{d}{dt} f(t, t, \ldots, t)$ indem man nach jedem Auftreten von t bei festgehaltenen anderen Auftreten ableitet und die Ableitungen addiert. In Formeln:

$$\boxed{\frac{d}{dt} f(t, t, \ldots, t) = \sum_{j=1}^{n} \frac{\partial}{\partial t_j} f(t_1, t_2, \ldots, t_n) \Big|_{t_1 = t_2 = \ldots = t_n = t}}$$

Zur Begründung dieses Rechentricks betrachten wir die Abbildung $g : \mathbb{R} \to \mathbb{R}^n$, $g(t) = (t, t, \ldots, t)$. Sie besitzt offensichtlich die Jacobimatrix

$$Dg(t) = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^{n \times 1}.$$

Durch Multiplikation mit der Jacobimatrix

$$Df(g(t)) = (D_1 f(g(t)), \ldots, D_n f(g(t)))$$

erhalten wir

$$\frac{d}{dt} f(t, t, \ldots, t) = D(f \circ g)(t) = Df(g(t)) \cdot Dg(t) = \sum_{j=1}^{n} \frac{\partial}{\partial t_j} f(t_1, t_2, \ldots, t_n) \Big|_{t_1 = t_2 = \ldots = t_n = t}.$$

Beispiel 2.26 (Illustrationen zur Kettenregel und zu den verschiedenen Notationen)

1. **Die Produktregel als Spezialfall der Kettenregel.** Für differenzierbare $f, g : \mathbb{R} \to \mathbb{R}$ tritt die Variable t im Produkt $f(t)g(t)$ an zwei Stellen auf:

$$f(\underbrace{t}_{\text{1. Auftr.}})g(\underbrace{t}_{\text{2. Auftr.}})$$

Ableitung nach dem 1. Auftreten bei festgehaltenem 2. Auftreten liefert $f'(t)g(t)$, Ableitung nach dem 2. Auftreten bei festgehaltenem 1. Auftreten dagegen $f(t)g'(t)$. Aufsummieren liefert die Produktregel:

$$\frac{d}{dt}[f(t)g(t)] = f'(t)g(t) + f(t)g'(t).$$

Das Gleiche ausführlich mit den Jacobimatrizen der Abbildungen $F : \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = xy$ und $k : \mathbb{R} \to \mathbb{R}^2$, $k(t) = (f(t), g(t))$ geschrieben:

$$\begin{aligned}
&DF(x, y) = (D_1F(x, y), D_2F(x, y)) = (y, x),\\
&DF(k(t)) = DF(f(t), g(t)) = (g(t), f(t))\\
&Dk(t) = \begin{pmatrix} f'(t) \\ g'(t) \end{pmatrix}\\
&\frac{d}{dt}[f(t)g(t)] = D(F \circ k)(t)\\
&= DF(k(t)) \cdot Dk(t) = (g(t), f(t)) \begin{pmatrix} f'(t) \\ g'(t) \end{pmatrix} = g(t)f'(t) + f(t)g'(t).
\end{aligned}$$

Nochmal das Gleiche in Differentialnotation geschrieben: Setzen wir in

$$dF(x, y) = d(xy) = y\, dx + x\, dy$$

die Substitution $x = f(t)$, $y = g(t)$, also $dx = f'(t)\, dt$, $dy = g'(t)\, dt$ ein, so erhalten wir:

$$d(f(t)g(t)) = g(t)f'(t)\, dt + f(t)g'(t)\, dt = [g(t)f'(t) + f(t)g'(t)]dt,$$

also

$$\frac{d(f(t)g(t))}{dt} = g(t)f'(t) + f(t)g'(t).$$

2. **Die Quotientenregel als Spezialfall der Kettenregel.** Für differenzierbare $f : \mathbb{R} \to \mathbb{R}$, $g : \mathbb{R} \to \mathbb{R} \setminus \{0\}$ tritt die Variable t im Quotienten $f(t)/g(t)$ an folgenden zwei Stellen auf:

$$\frac{f(\overbrace{t}^{\text{1. Auftr.}})}{g(\underbrace{t}_{\text{2. Auftr.}})}$$

Ableitung nach dem 1. Auftreten bei festgehaltenem 2. Auftreten liefert $f'(t)/g(t)$, Ableitung nach dem 2. Auftreten bei festgehaltenem 1. Auftreten dagegen $-f(t)g'(t)/g(t)^2$. Aufsummieren liefert die Quotientenregel:

$$\frac{d}{dt}\frac{f(t)}{g(t)} = \frac{f'(t)}{g(t)} - \frac{f(t)g'(t)}{g(t)^2} = \frac{f'(t)g(t) - f(t)g'(t)}{g(t)^2}.$$

Das Gleiche mit den Jacobimatrizen von $F : \mathbb{R} \times (\mathbb{R} \setminus \{0\}) \to \mathbb{R}$, $F(x,y) = x/y$ und $k : \mathbb{R} \to \mathbb{R}^2$, $k(t) = (f(t), g(t))$ geschrieben:

$$\begin{aligned}
DF(x,y) &= \left(\frac{1}{y}, -\frac{x}{y^2}\right),\\
DF(k(t)) &= \left(\frac{1}{g(t)}, -\frac{f(t)}{g(t)^2}\right)\\
Dk(t) &= \begin{pmatrix} f'(t) \\ g'(t) \end{pmatrix}\\
\frac{d}{dt}\frac{f(t)}{g(t)} &= D(F \circ k)(t) = DF(k(t)) \cdot Dk(t)\\
&= \left(\frac{1}{g(t)}, -\frac{f(t)}{g(t)^2}\right) \begin{pmatrix} f'(t) \\ g'(t) \end{pmatrix} = \frac{f'(t)}{g(t)} - \frac{f(t)g'(t)}{g(t)^2}
\end{aligned}$$

Das Gleiche in Differentialnotation: Wir setzen in

$$d\frac{x}{y} = \frac{1}{y}\,dx + x\,d\frac{1}{y} = \frac{1}{y}\,dx - \frac{x}{y^2}\,dy$$

die Substitution $x = f(t)$, $y = g(t)$, $dx = f'(t)\,dt$, $dy = g'(t)\,dt$ ein:

$$d\frac{f(t)}{g(t)} = \frac{1}{g(t)}\,df(t) + f(t)\,d\frac{1}{g(t)} = \frac{1}{g(t)}f'(t)\,dt - f(t)\frac{g'(t)}{g(t)^2}\,dt,$$

also

$$\frac{d(f(t)/g(t))}{dt} = \frac{f'(t)}{g(t)} - f(t)\frac{g'(t)}{g(t)^2}.$$

3. Wir berechnen $\frac{d}{dx}x^x$ für $x > 0$ mit der Kettenregel: Aus

$$\frac{\partial}{\partial x}x^y = yx^{y-1}, \quad \frac{\partial}{\partial y}x^y = x^y \log x$$

erhalten wir

$$\frac{d}{dx}x^x = yx^{y-1} + x^y \log x\Big|_{y=x} = xx^{x-1} + x^x \log x = (1 + \log x)x^x$$

Das stimmt natürlich mit dem aus der Analysis 1 bekannten Ergebnis

$$\frac{d}{dx}x^x = \frac{d}{dx}e^{x \log x} = (1 + \log x)x^x$$

überein.

4. Es seien $f : \mathbb{R}^2 \to \mathbb{R}$, $g : \mathbb{R} \to \mathbb{R}$ und $h : \mathbb{R} \to \mathbb{R}$ stetig differenzierbare Funktionen. Wir berechnen

$$\frac{d}{dx} \int_{g(x)}^{h(x)} f(x,y)\,dy$$

mit der Kettenregel: Hierzu betrachten wir die Abbildung

$$F : \mathbb{R}^3 \to \mathbb{R}, \quad F(a,b,x) = \int_a^b f(x,y)\,dy.$$

Sie besitzt nach dem Hauptsatz der Differential- und Integralrechnung und nach Lemma 2.7 die Jacobimatrix

$$DF(a,b,x) = \left(-f(x,a), f(x,b), \int_a^b D_1 f(x,y)\,dy \right).$$

Wir komponieren F mit der Abbildung

$$G : \mathbb{R} \to \mathbb{R}^3, \quad G(x) = (g(x), h(x), x),$$

mit der Jacobimatrix

$$DG(x) = \begin{pmatrix} g'(x) \\ h'(x) \\ 1 \end{pmatrix}.$$

Es folgt

$$\begin{aligned}
&\frac{d}{dx} \int_{g(x)}^{h(x)} f(x,y)\,dy = D(F \circ G)(x) = DF(G(x)) \cdot DG(x) \\
&= \left(-f(x,g(x)), f(x,h(x)), \int_{g(x)}^{h(x)} D_1 f(x,y)\,dy \right) \begin{pmatrix} g'(x) \\ h'(x) \\ 1 \end{pmatrix} \\
&= -f(x,g(x))g'(x) + f(x,h(x))h'(x) + \int_{g(x)}^{h(x)} D_1 f(x,y)\,dy.
\end{aligned}$$

Die drei Summanden im Ergebnis entstehen natürlich durch Ableiten nach den drei Auftreten von x in

$$\int_{g(\underbrace{x}_{\text{1. Auftr.}})}^{h(\overbrace{x}^{\text{2. Auftr.}})} f(\underbrace{x}_{\text{3. Auftr.}}, y)\,dy$$

Hier noch einmal die gleiche Rechnung in Differentialnotation: Aus

$$dF(a,b,x) = -f(x,a)\,da + f(x,b)\,db + \left(\int_a^b D_1 f(x,y)\,dy \right) dx \tag{36}$$

erhalten wir mit der Substitution

$$a = g(x), \quad b = h(x) \tag{37}$$

also

$$da = g'(x)\,dx, \quad db = h'(x)\,dx \tag{38}$$

das Ergebnis:

$$\begin{aligned} d\int_{g(x)}^{h(x)} f(x,y)\,dy &= d(F\circ G)(x) = dF(G(x))\circ dG(x) \\ &= -f(x,g(x))g'(x)\,dx + f(x,h(x))h'(x)\,dx + \left(\int_{g(x)}^{h(x)} D_1 f(x,y)\,dy\right)dx \end{aligned}$$

Man beachte, dass dieses Ergebnis formal einfach durch symbolisches Einsetzen von (37) und (38) in (36) entsteht.

5. Betrachten wir eine differenzierbare Abbildung $f : \mathbb{R}^n \to \mathbb{R}$ und eine weitere Abbildung $g = (g_1, \ldots, g_n) : \mathbb{R} \to \mathbb{R}^n$, so folgt

$$\frac{d}{dt} f(g_1(t), \ldots, g_n(t)) = Df(g(t)) \cdot Dg(t) = \sum_{j=1}^{n} D_j f(g(t))\dot{g}_j(t),$$

wobei wir die Ableitung nach der "Zeitvariablen" t in Newtonscher Tradition mit "$\dot{}$" statt mit "$'$" bezeichnen. Mit den Notationen $x_j = g_j(t)$ für $j = 1, \ldots, n$, $y = f(x_1, \ldots, x_n)$ können wir das Gleiche auch symbolisch in der Form

$$\frac{dy}{dt} = \sum_{j=1}^{n} \frac{\partial y}{\partial x_j}\frac{dx_j}{dt}$$

oder mit $dx_j = \dot{g}_j(t)\,dt$ auch in Differentialnotation als

$$dy = \sum_{j=1}^{n} \frac{\partial y}{\partial x_j}\,dx_j = \sum_{j=1}^{n} \frac{\partial y}{\partial x_j}\frac{dx_j}{dt}\,dt$$

schreiben. Betrachten wir als Beispiel die Abbildung $r : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$, $r(x) = \|x\|_2$ und eine parametrisierte differenzierbare Kurve[14] $x : \mathbb{R} \to \mathbb{R}^n \setminus \{0\}$, $t \mapsto x = (x_j)_{j=1,\ldots,n} = x(t)$. Wir stellen uns dies als Bahnkurve eines Teilchens im Raum vor. Aus $\nabla r(x) = x/r(x)$ erhalten wir

$$\frac{d}{dt} r(x(t)) = \sum_{j=1}^{n} \frac{\partial r}{\partial x_j}\frac{dx_j}{dt}(t) = \langle \nabla r(x(t)), \dot{x}(t)\rangle = \frac{\langle x(t), \dot{x}(t)\rangle}{r(x(t))}$$

Das Ergebnis bedeutet anschaulich: Bei der Änderung des Abstands von 0 kommt es nur auf die radiale Komponente des Geschwindigkeitsvektors $\dot{x}$ an.

[14]In diesem Beispiel verwenden wir die traditionelle Notation, den Funktionswert $x = x(t)$ und die Funktion x mit dem gleichen Buchstaben zu bezeichnen. Man lasse sich dadurch nicht verwirren.

Rückzug von Differentialen. Sehen wir uns die "Substitution" oder das "Einsetzen" von Transformationen in Differentiale, das in den Beispielen an mehreren Stellen auftrat, noch einmal vom konzeptuellen Standpunkt an:
Es sei $g : V_0 \to \mathbb{R}$ eine differenzierbare Abbildung auf einer offenen Menge $V_0 \subseteq V = \mathbb{R}^n$, $f : U_0 \to V_0$ eine weitere differenzierbare Abbildung auf einer offenen Menge $U_0 \subseteq U = \mathbb{R}^m$, $x \in U_0$ und $y = f(x) \in V_0$. Wir setzen $\omega := dg$, also $\omega_y = dg_y \in V' = \mathrm{Hom}_{\mathbb{R}}(V, \mathbb{R}) =$ Dualraum von V. Die Ableitung ω_y ist also eine Linearform auf V. Nach der Kettenregel ist

$$d(g \circ f)_x = dg_y \circ df_x = \omega_y \circ df_x.$$

Diese Operation ist in der linearen Algebra als "Adjungierte" bekannt:

Definition 2.27 Sind U und V zwei $\mathbb{R}$-Vektorräume und $L : U \to V$ eine lineare Abbildung, so wird durch

$$L^* : V' \to U', \quad L^*(\phi) = \phi \circ L$$

eine lineare Abbildung vom Dualraum V' in den Dualraum U' definiert. Sie heißt *Adjungierte* oder *duale Abbildung*[15] zu L. Insbesondere gilt für alle $\phi \in V'$ und $u \in U$:

$$\phi(L(u)) = (L^*(\phi))(u).$$

Im Spezialfall $U = \mathbb{R}^m$, $V = \mathbb{R}^n$ wird die lineare Abbildung L durch Multiplikation von *Spaltenvektoren* mit einer Matrix $A \in \mathbb{R}^{n\times m}$ *von links* dargestellt:

$$L(x) = A \cdot x$$

Stellen wir Linearformen in U', V' durch *Zeilenvektoren* dar, so wird dir Adjungierte L^* durch Multiplikation von *Spaltenvektoren* mit A *von rechts* dargestellt:

$$\mathbb{R}^{1\times n} \ni \Omega \mapsto \Omega \cdot A \in \mathbb{R}^{1\times m}.$$

Mit dieser Definition schreiben wir

$$d(g \circ f)_x = (df_x)^*(dg_y) = (df_x)^*(\omega_y).$$

Man nennt $(df_x)^*\omega_y$ auch den *Rückzug* der Linearform ω_y bei x und schreibt dafür kurz

$$(f^*\omega)_x := (df_x)^*(\omega_y) = \omega_y \circ df_x.$$

Der Rückzug $f^*\omega$ ist also eine Abbildung $f^*\omega : U_0 \to U' = \mathrm{Hom}_{\mathbb{R}}(U, \mathbb{R})$. In Komponentenschreibweise ist das einfach die "Substitutions-" oder "Einsetzoperation": In

$$\omega_y = dg_y = \sum_{j=1}^{n} \frac{\partial g}{\partial y_j}(y)\, dy_j$$

[15] Man beachte, dass dieses Bilden der dualen Abbildung die Pfeilrichtung umdreht!

setzen wir ein: $f(x) = (f_1(x), \ldots, f_n(x))$ für y und die Ableitung

$$(df_j)_x = \sum_{k=1}^{m} \frac{\partial f_j}{\partial x_k}(x)\, dx_k$$

für dy_j, so erhalten wir

$$(f^*\omega)_x = \sum_{j=1}^{n} \sum_{k=1}^{m} \frac{\partial g}{\partial y_j}(f(x)) \frac{\partial f_j}{\partial x_k}(x)\, dx_k.$$

Mit der traditionellen Schreibweise, den Wert $y = f(x)$ und die Funktion f mit dem gleichen Symbol y zu bezeichnen, kann man das Gleiche auch intuitiv-stenographisch so schreiben:

$$dg = \sum_{j=1}^{n} \frac{\partial g}{\partial y_j}\, dy_j$$

liefert durch Einsetzen von

$$dy_j = \sum_{k=1}^{m} \frac{\partial y_j}{\partial x_k}\, dx_k$$

den Rückzug

$$f^* dg = d(g \circ f) = \sum_{j=1}^{n} \sum_{k=1}^{m} \frac{\partial g}{\partial y_j} \frac{\partial y_j}{\partial x_k}\, dx_k.$$

Das Gleiche in Matrixnotation: Wird die Linearform ω_y durch den *Zeilenvektor*

$$\Omega_y := Dg(y) = (D_1 g(y), \ldots, D_n g(y))$$

dargestellt, so wird der Rückzug $(f^*\omega)_x$ durch die Multiplikation

$$\Omega_y \cdot Df_x$$

mit der Jacobimatrix der Transformation *von rechts* dargestellt.

Merkregel:

Linearisierung von f $\xrightarrow{\text{Darstellung}}$ Multiplikation mit Df *von links*

Rückzug mit f $\xrightarrow{\text{Darstellung}}$ Multiplikation mit Df *von rechts*

Beispiel 2.28 (Transformation in Polarkoordinaten) Die Abbildung $g : \mathbb{R}^2 \to \mathbb{R}$, $g(x, y) = x^2 y$ besitzt die Ableitung

$$dg(x, y) = 2xy\, dx + x^2\, dy,$$

in Matrixnotation

$$Dg(x, y) = (2xy, x^2).$$

Setzen wir hier die Transformation in Polarkoordinaten

$$f(r,\phi) = (r\cos\phi, r\sin\phi) =: (x,y)$$

ein. Wir rechnen:

$$\begin{aligned} dx &= d(r\cos\phi) = \cos\phi\, dr - r\sin\phi\, d\phi, \\ dy &= d(r\sin\phi) = \sin\phi\, dr + r\cos\phi\, d\phi. \end{aligned}$$

Das Gleiche in Matrixnotation:

$$Df_{(r,\phi)} = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \phi} \end{pmatrix} = \begin{pmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{pmatrix}$$

Wir erhalten:

$$\begin{aligned} f^*dg_{(r,\phi)} = d(g\circ f)_{(r,\phi)} &= 2r\cos\phi \cdot r\sin\phi \cdot (\cos\phi\, dr - r\sin\phi\, d\phi) \\ &+ (r\cos\phi)^2 \cdot (\sin\phi\, dr + r\cos\phi\, d\phi) \\ &= 3r^2\cos^2\phi\sin\phi\, dr + r^3(\cos^3\phi - 2\cos\phi\sin^2\phi)\, d\phi. \end{aligned}$$

Das Gleiche in Matrixnotation:

$$\begin{aligned} D(g\circ f)_{(r,\phi)} = Dg_{f(r,\phi)} \cdot Df_{(r,\phi)} &= (2r\cos\phi \cdot r\sin\phi, (r\cos\phi)^2) \begin{pmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{pmatrix} \\ &= (3r^2\cos^2\phi\sin\phi, r^3(\cos^3\phi - 2\cos\phi\sin^2\phi)). \end{aligned}$$

Übung 2.29 Betrachten Sie die Funktion $g : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$, $g(x,y) = \arctan(y/x)$. Berechnen Sie den Rückzug f^*dg der Ableitung von g unter der Polarkoordinatenabbildung

$$f : \mathbb{R}^+ \times]-\pi/2, \pi/2[\to \mathbb{R}^+ \times \mathbb{R}, \quad (x,y) = f(r,\phi) = (r\cos\phi, r\sin\phi)$$

auf zwei verschiedene Weisen:

a) indem Sie $g\circ f$ ausrechnen und dann ableiten;

b) indem Sie die Ableitung von g bilden und diese dann mit df^* zurückziehen.

Schreiben Sie die Rechnung jeweils sowohl in Differentialnotation als auch in Matrixnotation. Überzeugen Sie sich davon, dass die Ergebnisse übereinstimmen.

Übung 2.30 Berechnen Sie den Rückzug f^*dg für folgende Daten (definiert jeweils auf geeigneten offenen Teilmengen von $\mathbb{R}^3$):

a) $f(x_1,x_2,x_3) = (y_1,y_2,y_3)$ mit $y_1 = e^{x_1x_2-x_3}$, $y_2 = e^{x_1x_2+x_3}$, $y_3 = e^{x_2x_3}$ und $dg(y_1,y_2,y_3) = y_2y_3\, dy_1 + y_1y_3\, dy_2 + y_1y_2\, dy_3$.

b) $f(x_1, x_2, x_3) = (x_1/(x_1 + x_2 + x_3), x_2/(x_1 + x_2 + x_3), x_1 + x_2 + x_3) = (y_1, y_2, y_3)$ und $dg(y_1, y_2, y_3) = e^{y_1 y_3 + 2y_2 y_3}[y_3\, dy_1 + 2y_3\, dy_2 + (y_1 + 2y_2)dy_3]$.

Gehen Sie dazu jeweils auf zwei verschiedene Weisen vor:

1. mit direkter Rechnung durch Anwendung der Adjungierten df^* auf dg,

2. durch Bestimmen einer Funktion g mit der gegebenen Ableitung dg und Berechnen von $d(g \circ f)$.

Überzeugen Sie sich davon, dass die Ergebnisse von 1. und 2. übereinstimmen.

Übung 2.31 (Eine Formel von Heun) Gegeben sei eine Differentialgleichung $y'(x) = f(x, y(x))$ mit einer beliebig oft (partiell) differenzierbaren Abbildung $f : \mathbb{R}^2 \to \mathbb{R}$ und einer beliebig oft differenzierbaren Lösung $y : \mathbb{R} \to \mathbb{R}$, die die Anfangsbedingung $y(0) = b \in \mathbb{R}$ erfüllt. Wir definieren die Näherung $z : \mathbb{R} \to \mathbb{R}$ an y durch

$$z(x) = b + xf(x/2, b + xf(0, b)/2).$$

Beweisen Sie, dass die Taylorpolynome 2. Grades von y und von z um $x_0 = 0$ übereinstimmen.
Bemerkung: Aus dieser Formel und Varianten davon gewinnt man durch Iteration numerische Verfahren zur Berechnung von Näherungslösungen von Differentialgleichungen. Mehr dazu in der Numerischen Mathematik.

Übung 2.32 (Länge von Kurven in krummlinigen Koordinaten) Ein m-dimensionales Gebilde $G \subseteq \mathbb{R}^n$ werde durch eine stetig differenzierbare Abbildung $f : \mathbb{R}^m \supseteq U \to \mathbb{R}^n$ in Parameterdarstellung $G = f[U]$ gegeben. Weiter sei eine stetig differenzierbare Kurve $k : [a, b] \to U$ gegeben. Wir stellen uns k als eine Beschreibung der Kurve $f \circ k$ "in krummlinigen Koordinaten f" vor. Zeigen Sie, dass die Länge der Kurve $f \circ k$ durch

$$\int_a^b \sqrt{k'(s)^t g(k(s)) k'(s)}\, ds$$

gegeben wird, wobei

$$g : U \to \mathbb{R}^{m \times m}, g(x) = Df(x)^t Df(x).$$

Die matrixwertige Abbildung g wird die *Riemannsche Metrik* zur Parametrisierung f genannt. Berechnen Sie die Riemannsche Metrik für die Polarkoordinatenabbildung $f(r, \phi) = (r\cos\phi, r\sin\phi)$ und für die Kugelkoordinatenabbildung

$$f(r, \theta, \phi) = (r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta).$$

Berechnen Sie damit die Länge der Kurve, die in Polarkoordinaten durch $r(s) = e^s$, $\phi(s) = s$, $s \in [0, 1]$ gegeben wird.

Übung 2.33 Die Sauerstoffkonzentration $c(x, t)$ in einem zylindrischen Wasserbecken (Grundfläche A, zeitabhängige Wasserstandshöhe $h(t)$) hänge nur von der Höhe $x \le h(t)$ über dem Boden und von der Zeit t ab. Es bezeichne $m(t)$ die Sauerstoffmenge im Becken zur Zeit t. Drücken Sie (unter geeigneten Differenzierbarkeitsvoraussetzungen) die Änderungsgeschwindigkeit $\dot{m}(t)$ der Sauerstoffmenge mit Hilfe der Kettenregel aus.

Übung 2.34 Berechnen Sie unter geeigneten Differenzierbarkeitsvoraussetzungen an $f, g : \mathbb{R}^2 \to \mathbb{R}^+$ den Gradienten der Funktion $h : \mathbb{R}^2 \ni (s,t) \mapsto f(s,t)^{g(s,t)} \in \mathbb{R}$. Das Ergebnis darf in Matrixnotation geschrieben werden.

Übung 2.35 (Rotationsinvarianz des Laplaceoperators) Es seien $n \in \mathbb{N}$, $V \subseteq \mathbb{R}^n$ offen und $U : \mathbb{R}^{n\times n}$ eine orthogonale Matrix, d.h. $U^t U = \mathrm{Id}$ oder äquivalent $\forall x \in \mathbb{R}^n$: $\|Ux\|_2 = \|x\|_2$. Weiter sei $UV := \{Ux \mid x \in V\}$ und $g \in C^2(UV, \mathbb{R})$ und $f : V \to \mathbb{R}$, $f(x) = g(Ux)$. Zeigen Sie für alle $x \in V$: $\Delta f(x) = \Delta g(Ux)$.

Übung 2.36 (Der Laplaceoperator in Polarkoordinaten) Es sei $f \in C^2(\mathbb{R}^2 \backslash \{0\}, \mathbb{C})$ und $g : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{C}$, $g(r,\phi) = f(r\cos\phi, r\sin\phi)$. Zeigen Sie:

$$\Delta f(r\cos\phi, r\sin\phi) = \frac{\partial^2 g}{\partial r^2}(r,\phi) + \frac{1}{r}\frac{\partial g}{\partial r}(r,\phi) + \frac{1}{r^2}\frac{\partial^2 g}{\partial \phi^2}(r,\phi)$$

Bemerkung: Mit mehr Hilfsmitteln können wir später in der Analysis 3 die Transformation des Laplaceoperators in beliebige krummlinige Koordinaten in n Dimensionen recht einfach beschreiben.

Übung 2.37 (Von der Wärmeleitungsgleichung via Fourierreihe zur Besselschen Differentialgleichung) Es sei $f \in C^\infty(\mathbb{R}^2 \backslash \{0\}, \mathbb{C})$, so dass $g : (\mathbb{R}^2 \backslash \{0\}) \times \mathbb{R} \to \mathbb{C}$, $g(t,x,y) = e^{-t} f(x,y)$ eine exponentiell abfallende Lösung der Wärmeleitungsgleichung[16]

$$\frac{\partial}{\partial t} g(t,x,y) = \Delta g(t,x,y)$$

bildet, wobei sich der Laplaceoperator $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ nur auf x und y, nicht jedoch auf t beziehen soll.

(a) Zeigen Sie: f löst die "*Helmholtzgleichung*" $\Delta f + f = 0$.

(b) Wir setzen für $k \in \mathbb{Z}$ und $r > 0$:

$$f_k(r) = \frac{1}{2\pi} \int_0^{2\pi} e^{-ik\phi} f(r\cos\phi, r\sin\phi)\, d\phi$$

und für $(x,y) = (r\cos\phi, r\sin\phi) \in \mathbb{R}^2 \setminus \{0\}$:

$$g_k(x,y) := f_k(r) e^{ik\phi} = \frac{1}{2\pi} \int_0^{2\pi} e^{-ik\theta} f(x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta)\, d\theta.$$

Überzeugen Sie sich, dass die letzte Gleichung gilt. Zeigen Sie für alle (x,y) von eben:

$$\sum_{k\in\mathbb{Z}} g_k(x,y) = f(x,y).$$

[16] Hier ohne Faktor 1/2

(c) Zeigen Sie für $k \in \mathbb{Z}$, dass auch g_k die Helmholtzgleichung $\Delta g_k + g_k = 0$ erfüllt und folgern Sie, dass f_k die "*Besselsche Differentialgleichung*"

$$r^2 f_k''(r) + r f_k'(r) + (r^2 - k^2) f_k(r) = 0$$

für $r > 0$ erfüllt. Welche Funktionen f_k erhalten Sie im Spezialfall $f(x,y) = i^{-n} e^{ix}$?

Übung 2.38 Anfangswertproblem für die Wärmeleitungsgleichung. *Lösen Sie entweder die Variante (a) oder die Variante (b) der folgenden Aufgabe.*
Gegeben sei eine stetige Funktion $g : \mathbb{R} \to \mathbb{R}$, wobei

(a) $g(x) = \max\{e^x - K, 0\}$ für gegebenes $K > 0$.

(b) $\int_{-\infty}^{\infty} e^{-ax^2} |g(x)|\, dx < \infty$ für alle $a > 0$.

Weiter sei

$$f : \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}, \quad f(x,t) = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{2t}} g(y)\, dy. \tag{39}$$

Zeigen Sie, dass f die Wärmeleitungsgleichung

$$\frac{\partial f}{\partial t} = \frac{1}{2} \frac{\partial^2 f}{\partial x^2}$$

löst.
Hinweis:

(a) Zeigen und verwenden Sie hierzu die Formel

$$f(x,t) = e^{x+\frac{t}{2}} \Phi(t^{-1/2}(x + t - \log K)) - K\Phi(t^{-1/2}(x - \log K))$$

mit der Funktion

$$\Phi : \mathbb{R} \to \mathbb{R}, \quad \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{u^2}{2}}\, du = \frac{1}{\sqrt{2\pi}} \int_{-z}^{\infty} e^{-\frac{u^2}{2}}\, du.$$

Die Funktion Φ wird auch "*Verteilungsfunktion der Standardnormalverteilung*" genannt.

(b) Um zu zeigen, dass das Integral (39) mit den partiellen Ableitungen nach x und t vertauscht werden kann, bilden Sie Differenzenquotienten von f und von $\partial f/\partial x$ und vergleichen Sie diese mit dem Integral der entsprechenden partiellen Ableitung des Integranden. Schätzen Sie die Restterme im Integranden mit der Taylorformel ab.

Zeigen Sie, dass die Anfangsbedingung

$$f(x,t) \to g(x_0) \quad \text{für} \quad (x,t) \to (x_0, 0), t > 0$$

für alle $x_0 \in \mathbb{R}$ erfüllt ist.
Hinweise:

(a) Zeigen und verwenden Sie dabei:

$$\forall x, y \in \mathbb{R}: \ |\Phi(x+y) - \Phi(x)| \le \frac{|y|}{\sqrt{2\pi}}$$

und

$$\lim_{s\to\infty} \Phi(sa) = \begin{cases} 1 & \text{für } a > 0, \\ 1/2 & \text{für } a = 0, \\ 0 & \text{für } a < 0. \end{cases}$$

(b) Spalten Sie den Integrationsbereich des Integrals

$$f(x,t) - g(x) = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{2t}} (g(y) - g(x))\, dy$$

für kleine $\delta > 0$ in die Bereiche $|y - x| > \delta$ und $|y - x| \le \delta$ auf. Beachten Sie, dass der Teil zu $|y - x| > \delta$ im Limes $(x, t) \to (x_0, 0)$, $t > 0$ vernachlässigbar ist.

Sie dürfen die Formel

$$\int_{-\infty}^{\infty} e^{-\frac{z^2}{2}}\, dz = \sqrt{2\pi},$$

also $\lim_{z\to\infty} \Phi(z) = 1$, ohne Beweis verwenden; wir beweisen sie nämlich in der Analysis 3.

Übung 2.39 Optionspreise im Black-Scholes-Modell. Eine "*europäische Call-Option*" ist ein Finanzderivat, das seinem Inhaber das Recht, aber nicht die Pflicht gibt, zu einem vereinbarten zukünftigen Zeitpunkt T eine Einheit eines vereinbarten "Basiswerts" (z.B. einer Aktie, einer Währung, eines Rohstoffs) zu einem vereinbarten Preis K zu kaufen. Die Option verfällt damit wertlos zur Zeit T, wenn der Marktpreis S_T des Basiswerts zur Zeit T kleiner oder gleich K ist; andernfalls hat sie dann den Wert $S_T - K$. Zur Zeit T besitzt die Call-Option also den Wert $\max\{S_T - K, 0\}$. Der Preis C einer solchen Call-Option zu einer früheren Zeit $t < T$ wird vom Wert S des Basiswerts zu dieser Zeit abhängen: $C = C(t, S)$. Betrachten wir den mit dem Marktzins r diskontierten Optionspreis

$$L(t, x) := e^{-rt} C(t, S)$$

in Abhängigkeit vom logarithmierten diskontierten Basiswertpreis

$$x := \log(e^{-rt} S).$$

In einem berühmten finanzmathematischen Modell, dem *Black-Scholes Modell*, wird der diskontierte Preis L durch die *Rückwärts-Wärmeleitungsgleichung mit Drift*

$$\frac{\partial L}{\partial t} + \frac{\sigma^2}{2} \frac{\partial^2 L}{\partial x^2} - \frac{\sigma^2}{2} \frac{\partial L}{\partial x} = 0 \tag{40}$$

für $t < T$ und $x \in \mathbb{R}$ beschrieben. Hierbei ist $\sigma > 0$ ein Modellparameter, die "*Volatilität*" des Basiswerts, der das Ausmaß der stochastischen Fluktuationen des Basiswerts quantifiziert. (Die zugrundeliegenden Modellannahmen und eine Herleitung der Gleichung (40) aus diesen Annahmen mit Methoden der Stochastik können Sie in Vorlesungen zur stochastischen Analysis und Finanzmathematik lernen.)

1. **Driftentfernung, Zeitskalierung und Zeitumkehr.** Zeigen Sie, dass mit der affin-linearen Transformation

$$f(\tilde{t}, \tilde{x}) := L(T - \sigma^{-2}\tilde{t}, \tilde{x} + \tilde{t}/2) \quad \text{mit } \tilde{t} > 0, \tilde{x} \in \mathbb{R}$$

die obige Rückwärts-Wärmeleitungsgleichung (40) äquivalent zur einfachen Wärmeleitungsgleichung

$$\frac{\partial f}{\partial \tilde{t}} = \frac{1}{2}\frac{\partial^2 f}{\partial \tilde{x}^2}$$

wird.

2. **Black-Scholes Differentialgleichung.** Folgern Sie aus der Rückwärts-Wärmeleitungsgleichung (40) die partielle Differentialgleichung

$$\frac{\partial C}{\partial t} + rS\frac{\partial C}{\partial S} + \frac{\sigma^2}{2}S^2\frac{\partial^2 C}{\partial S^2} - rC = 0.$$

Sie wird "Black-Scholes Differentialgleichung" genannt.

3. **Black-Scholes-Preise europäischer Call-Optionen.** Finden Sie eine Lösung $C(t, S)$, $t < T$, $S > 0$ der Black-Scholes Differentialgleichung mit der Randbedingung

$$C(t, S) \to \max\{S_T - K, 0\} \text{ für } (t, S) \to (T, S_T),\ t < T,$$

indem Sie die Lösung

$$f(\tilde{t}, \tilde{x}) = e^{\tilde{x}+\frac{\tilde{t}}{2}}\Phi(\tilde{t}^{-1/2}(\tilde{x} + \tilde{t} - \log K)) - K\Phi(\tilde{t}^{-1/2}(\tilde{x} - \log K))$$

der Wärmeleitungsgleichung aus Aufgabe 2.38 (a) transformieren. Überprüfen zur Probe mit einer direkten Rechnung, dass Ihr Ergebnis wirklich die Black-Scholes Differentialgleichung löst.

4. **Veranschaulichung des Ergebnisses.** Veranschaulichen Sie sich diese Optionspreise $C(t, S)$ im Black-Scholes-Modell für den Fall $r = 0$, $T = 0$, $\sigma = 1$ und $K = 1$, indem Sie den Graphen von $S \mapsto C(t, S)$ für verschiedene $t < 0$ in ein einziges S-C-Diagramm skizzieren. (Eine qualitative Skizze genügt.)

2.6 Die multidimensionale Taylorformel

Wir besprechen nun eine multidimensionale Variante der Taylorformel. Hierzu sind einige Vorbereitungen nützlich. Einige Abkürzungen: Für $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$ und einen Multiindex $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$ setzen wir

$$h^\alpha := \prod_{j=1}^{n} h_j^{\alpha_j}$$

und

$$\alpha! := \prod_{j=1}^{n} \alpha_j!$$

Erinnerung: $|\alpha| = \alpha_1 + \ldots + \alpha_n$.
Zunächst eine kombinatorische Vorbereitung: Wir zählen, auf wie viele verschiedene Weisen man D^α durch Hintereinanderstellen von α_1-mal D_1, α_2-mal D_2, …, α_n-mal D_n schreiben kann.

Beispiel 2.40 *$D^{(1,1)}$ kann man so auf zwei Weisen schreiben, nämlich D_1D_2 und D_2D_1, $D^{(2,0)}$ jedoch nur als D_1D_1. Den Differentialoperator $D^{(2,1,1)}$ kann man so auf* $12 = 4!/(2! \cdot 1! \cdot 1!)$ *Weisen schreiben, nämlich*

$$\begin{gathered} D_1D_1D_2D_3, D_1D_1D_3D_2, D_1D_2D_1D_3, D_1D_3D_1D_2, \\ D_1D_2D_3D_1, D_1D_3D_2D_1, D_2D_1D_1D_3, D_3D_1D_1D_2, \\ D_2D_1D_3D_1, D_3D_1D_2D_1, D_2D_3D_1D_1, D_3D_2D_1D_1. \end{gathered}$$

Wir nennen diese Anzahl von Schreibweisen von D^α den *Multinomialkoeffizienten* von α. Anders gesagt ist er die Anzahl der Abbildungen $\iota : \{1, \ldots, |\alpha|\} \to \{1, \ldots, n\}$, die für alle $k = 1, \ldots, n$ genau α_k-mal den Wert k annehmen. Wir bezeichnen mit $M(\alpha)$ die Menge dieser Abbildungen.

Lemma 2.41 (Multinomialkoeffizient) *Der Multinomialkoeffizient von α beträgt*

$$\frac{|\alpha|!}{\alpha!} = \frac{(\alpha_1 + \ldots + \alpha_n)!}{\alpha_1! \cdot \ldots \cdot \alpha_n!}.$$

Veranschaulichung der Beweisidee: Auf wie viele verschiedene Weisen kann man 3 rote, 4 grüne und 5 blaue Kugeln in eine Reihe legen? Sind die Kugeln unterscheidbar, so gibt es $(3 + 4 + 5)! = 12!$ Möglichkeiten. Sind Kugeln gleicher Farbe jedoch ununterscheidbar, so kann man die roten Kugeln auf 3! Weisen vertauschen, die grünen auf 4! Weisen, und die blauen auf 5! Weisen, ohne dass man das unterscheiden kann. Zusammen sind das $3! \cdot 4! \cdot 5!$ Vertauschungen. Von den ursprünglich 12! Anordnungen bleiben also nur $12!/(3! \cdot 4! \cdot 5!)$ *unterscheidbare* übrig. Formalisieren wir diese Idee:
Beweis von Lemma 2.41: Wir betrachten die Abbildung $\ell : \{1, \ldots, |\alpha|\} \to \{1, \ldots, n\}$, die die ersten α_1 Elemente von $\{1, \ldots, |\alpha|\}$ auf 1 abbildet, die nächsten α_2 Elemente auf 2, usw., und die letzten α_n Elemente auf n. Dann können wir jede Abbildung $\iota \in M(\alpha)$ auf genau $\alpha_1! \cdot \ldots \cdot \alpha_n!$ verschiedene Weisen als $\iota = \ell \circ \sigma$ mit einer Bijektion $\sigma : \{1, \ldots, |\alpha|\} \to \{1, \ldots, |\alpha|\}$ schreiben. Es gibt nämlich für $k = 1, \ldots, n$ genau $\alpha_k!$ verschiedene Weisen, die α_k-elementige Menge $\iota^{-1}[\{k\}]$ auf die gleichmächtige Menge $\ell^{-1}[\{k\}]$ bijektiv abzubilden, also zusammen $\alpha_1! \cdot \ldots \cdot \alpha_n!$ verschiedene Weisen, eine Bijektion $\sigma : \{1, \ldots, |\alpha|\} \to \{1, \ldots, |\alpha|\}$ (synonym: eine Permutation von $\{1, \ldots, |\alpha|\}$) so zu wählen, dass sie für alle $k = 1, \ldots, n$ die Menge $\iota^{-1}[\{k\}]$ auf die gleichmächtige

Menge $\ell^{-1}[\{k\}]$ bijektiv abbildet. So gezählt gibt es also genau $|M(\alpha)|\alpha_1! \cdot \ldots \cdot \alpha_n!$ Permutationen von $\{1, \ldots, |\alpha|\}$, wobei $|M(\alpha)|$ die Mächtigkeit der Menge $|M(\alpha)|$ bezeichnet. Andererseits gibt es genau $|\alpha|!$ Permutationen von $\{1, \ldots, |\alpha|\}$. Es folgt

$$|M(\alpha)|\alpha_1! \cdot \ldots \cdot \alpha_n! = |\alpha|!,$$

also die Behauptung.

□

Übung 2.42 (Multinomialformel) *Zeigen Sie für einen Multiindex $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$, $m \in \mathbb{N}_0$ und für $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$ die folgende Verallgemeinerung der binomischen Formel:*

$$(h_1 + \ldots + h_n)^m = \sum_{\substack{\alpha \in \mathbb{N}_0^n \\ |\alpha| = m}} \frac{|\alpha|!}{\alpha!} h^\alpha.$$

Lemma 2.43 (Iterierte Richtungsableitung) *Gegeben seien $U \subseteq \mathbb{R}^n$ offen, $x \in U$, eine m-fach stetig (partiell) differenzierbare Abbildung $f : U \to \mathbb{R}$ und ein Vektor $h \in \mathbb{R}^n$, wobei $m \in \mathbb{N}_0$ und $n \in \mathbb{N}$. Dann gilt für die m-fach iterierte Richtungsableitung*

$$D_h^m f(x) = \sum_{\iota_1=1}^{n} \ldots \sum_{\iota_m=1}^{n} h_{\iota_1} \ldots h_{\iota_n} D_{\iota_1} \ldots D_{\iota_n} f(x) = \sum_{\substack{\alpha \in \mathbb{N}_0^n \\ |\alpha| = m}} \frac{|\alpha|!}{\alpha!} h^\alpha D^\alpha f(x)$$

Beweis. Nach der Kettenregel gilt für in x stetig differenzierbare Abbildungen $g : U \to \mathbb{R}^n$ die Formel

$$D_h g(x) = \sum_{j=1}^{n} h_j D_j g(x).$$

Die erste Gleichung in der Behauptung folgt durch m-faches Anwenden dieser Formel auf f. Die zweite Gleichung folgt mit der Formel für den Multinomialkoeffizienten, indem wir alle m-fachen partiellen Ableitungen $D_{\iota_1} \ldots D_{\iota_n}$, die zu D^α mit dem gleichen Multiindex α gehören, zusammenfassen.

□

Wir nennen eine Menge $U \subseteq \mathbb{R}^n$ *sternförmig* mit dem Zentrum $x \in U$, wenn für alle $y \in U$ die Verbindungsstrecke

$$[x, y] = \{sy + (1 - s)x |\ 0 \leq s \leq 1\}$$

ganz zu U gehört.
Im Folgenden bezeichnet $\|\cdot\|$ eine beliebige gegebene Norm auf $\mathbb{R}^n$.

Satz 2.44 (Multidimensionale Taylorformel) *Es sei $f : U \to \mathbb{R}$ eine $(m+1)$-fach stetig (partiell) differenzierbare Funktion auf einer offenen, sternförmigen Menge $U \subseteq \mathbb{R}^n$ mit Zentrum x, wobei $m \in \mathbb{N}_0$ und $n \in \mathbb{N}$. Dann gilt für alle $h \in \mathbb{R}^n$ mit $x + h \in U$:*

$$f(x+h) = \sum_{\substack{\alpha\in\mathbb{N}_0^n \\ |\alpha|\le m}} \frac{h^\alpha}{\alpha!} D^\alpha f(x) + (m+1) \sum_{\substack{\alpha\in\mathbb{N}_0^n \\ |\alpha|=m+1}} \frac{h^\alpha}{\alpha!} \int_0^1 (1-s)^m D^\alpha f(x+sh)\, ds.$$

Vergröbert können wir das auch in der Form

$$\boxed{f(x+h) = \sum_{\substack{\alpha\in\mathbb{N}_0^n \\ |\alpha|\le m}} \frac{h^\alpha}{\alpha!} D^\alpha f(x) + O(\|h\|^{m+1}) \qquad \text{für } h \to 0}$$

schreiben. Ist dagegen f wenigstens noch m-mal (partiell) differenzierbar, so gilt folgende etwas schwächere Form des Restglieds:

$$f(x+h) = \sum_{\substack{\alpha\in\mathbb{N}_0^n \\ |\alpha|\le m}} \frac{h^\alpha}{\alpha!} D^\alpha f(x) + o(\|h\|^m) \qquad \text{für } h \to 0$$

Besonders wichtig ist der Fall $m = 2$: Hier gilt für 3-fach stetig (partiell) differenzierbare f:

$$\sum_{\substack{\alpha\in\mathbb{N}_0^m \\ |\alpha|\le m}} \frac{h^\alpha}{\alpha!} D^\alpha f(x) = f(x) + Df(x)\cdot h + \frac{1}{2} h^t \cdot D^2 f(x) \cdot h,$$

also

$$\boxed{f(x+h) = f(x) + Df(x)\cdot h + \frac{1}{2} h^t \cdot D^2 f(x) \cdot h + O(\|h\|^3) \qquad \text{für } h \to 0.}$$

Hat man nur zweifach stetig (partielle) Differenzierbarkeit von f, so gilt für das Restglied wenigstens noch die schwächere Form $o(\|h\|^2)$ für $h \to 0$.

Das multivariate Polynom $\sum_{\alpha\in\mathbb{N}_0^m, |\alpha|\le m} \frac{h^\alpha}{\alpha!} D^\alpha f(x)$ wird *Taylorpolynom* vom Grad m von f an der Stelle x zum Abstandsvektor h genannt.
Beweis der multidimensionalen Taylorformel: Der Beweis beruht auf der eindimensionalen Taylorformel, angewandt auf die Funktion

$$g(s) = f(x+sh), \quad 0 \le s \le 1.$$

Man beachte, dass g nach der Kettenregel $(m+1)$-fach differenzierbar ist. Die eindimensionale Taylorformel liefert:

$$f(x+h) = g(1) = \sum_{j=0}^m \frac{D^j g(0)}{j!} + (m+1) \int_0^1 (1-s)^m \frac{D^{m+1} g(s)}{(m+1)!}\, ds,$$

wobei wir das Lagrange-Restglied mit $m+1$ erweitert haben. Die multidimensionale Taylorformel folgt nun aus der Formel für die iterierte Richtungsableitung:

$$D^j g(s) = D_h^j f(x+sh) = \sum_{\substack{\alpha\in\mathbb{N}_0^n \\ |\alpha|=j}} \frac{|\alpha|!}{\alpha!} h^\alpha D^\alpha f(x+sh). \tag{41}$$

Die vergröberte Form der Taylorformel mit Restglieddarstellung $O(\|h\|^{m+1})$ folgt aus der Beschränktheit von $D^\alpha f$ nahe bei x für $|\alpha| = m+1$ und aus $h^\alpha = O(\|h\|^{|\alpha|})$ für $h \to 0$. Um auch die Restglieddarstellung $o(\|h\|^m)$ der Voraussetzung von nur m-facher stetig (partieller) Differenzierbarkeit herzuleiten, schreiben wir

$$f(x+h) = \sum_{\substack{\alpha\in\mathbb{N}_0^n \\ |\alpha|\le m}} \frac{h^\alpha}{\alpha!} D^\alpha f(x) + (m+1) \sum_{\substack{\alpha\in\mathbb{N}_0^n \\ |\alpha|=m}} \frac{h^\alpha}{\alpha!} \int_0^1 (1-s)^m [D^\alpha f(x+sh) - D^\alpha f(x)]\, ds,$$

man beachte hierbei $(m+1)\int_0^1 (1-s)^m\, ds = 1$. Für $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| = m$ erhalten wir wegen der Stetigkeit von $D^\alpha f$ in x:

$$\lim_{h\to 0} \sup_{s\in[0,1]} |D^\alpha f(x+sh) - D^\alpha f(x)| = 0$$

und $h^\alpha = O(\|h\|^{|\alpha|})$ für $h \to 0$. Es folgt

$$\frac{h^\alpha}{\alpha!} \int_0^1 (1-s)^m [D^\alpha f(x+sh) - D^\alpha f(x)]\, ds = o(\|h\|^{|\alpha|}) \text{ für } h \to 0,$$

also die Behauptung.
Den Spezialfall $m = 2$ der multidimensionalen Taylorformel erhält man aus

$$\sum_{\substack{\alpha\in\mathbb{N}_0^n \\ |\alpha|=0}} \frac{h^\alpha}{\alpha!} D^\alpha f(x) = f(x),$$

$$\sum_{\substack{\alpha\in\mathbb{N}_0^n \\ |\alpha|=1}} \frac{h^\alpha}{\alpha!} D^\alpha f(x) = \sum_{j=1}^n h_j D_j f(x) = Df(x)\cdot h,$$

$$\sum_{\substack{\alpha\in\mathbb{N}_0^n \\ |\alpha|=2}} \frac{h^\alpha}{\alpha!} D^\alpha f(x) = \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n h_i h_j D_i D_j f(x) = \frac{1}{2} h^t \cdot D^2 f(x) \cdot h.$$

Übung 2.45 Betrachten Sie die Determinantenabbildung $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$, $n \in \mathbb{N}$.

1. Zeigen Sie, dass die Ableitung der Determinantenabbildung bei der Einheitsmatrix Id durch die Spur

   $$d\det_{\mathrm{Id}}(A) = \operatorname{Spur} A$$

   für $A \in \mathbb{R}^{n\times n}$ gegeben wird.
   Erinnerung: Die Spur einer quadratischen Matrix ist die Summe der Diagonaleinträge.

2. Folgern Sie für $A, B \in \mathbb{R}^{n\times n}$ mit $\det B \neq 0$:

$$d\det_B(A) = \operatorname{Spur}(B^{-1}A)\det B$$

3. Entwickeln Sie $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$ bei der Einheitsmatrix Id in ein Taylorpolynom vom Grad n. Zeigen Sie damit die Formel

$$\det(\mathrm{Id} + tA) = 1 + \sum_{k=1}^{n} t^k \sum_{\substack{I\subseteq\{1,\dots,n\}\\ |I|=k}} \det A_{I,I}.$$

   für $A \in \mathbb{R}^{n\times n}$, $t \in \mathbb{R}$. Hierbei bezeichnet $|I|$ die Anzahl der Elemente von I und $A_{I,I}$ für eine Matrix $A = (a_{i,j})_{i,j=1,\dots,n} \in \mathbb{R}^{n\times n}$ die Untermatrix $A_{I,I} = (a_{i,j})_{i,j\in I} \in \mathbb{R}^{I\times I}$.

Variante der Übung: Falls Sie eine einfachere Version der Aufgabe wünschen, bearbeiten Sie nur den Spezialfall $n = 2$ oder $n = 3$.

2.7 Stationäre Punkte und lokale Extrema

Erinnern Sie sich an die Analysis 1: Ist $f : \mathbb{R} \to \mathbb{R}$ differenzierbar, so ist an jedem lokalen Minimum oder Maximum die Ableitung gleich 0. Analoges gilt auch multidimensional. Wir definieren dazu

Definition 2.46 Es sei U eine Menge, versehen mit einer Topologie, und $f : U \to \mathbb{R}$ eine Abbildung. Ein Punkt $x \in U$ heißt *lokales Minimum* (bzw. *lokales Maximum*) von f, wenn es eine Umgebung $V \subseteq U$ von x gibt, so dass für alle $y \in V$ gilt: $f(y) \geq f(x)$ (bzw. $f(y) \leq f(x)$). "*Lokales Extremum*" ist ein Synonym für "lokales Minimum oder Maximum".
Nun sei speziell $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ differenzierbar. Nullstellen der Ableitung von f, also Punkte $x \in U$ mit $df(x) = 0$, werden *stationäre Punkte* von f genannt.

Lemma 2.47 (Lokale Extrema im Innern sind stationär) *Ist $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ differenzierbar, so ist jedes lokale Extremum von f ein stationärer Punkt.*

Die Umkehrung davon ist falsch, wie Sie schon aus der Analysis 1 wissen: $(0, 0)$ ist ein stationärer Punkt von $(x, y) \mapsto x^2 - y^2$, aber kein lokales Extremum. Beachten Sie, dass die Voraussetzung, dass U offen ist, wichtig ist: Andernfalls kann es am Rand von U weitere lokale Extrema geben. Zum Beispiel hat $[0, 1] \ni x \mapsto x$ ein Minimum bei 0, aber dort verschwindet die Ableitung nicht.
Beweis zu Lemma 2.47: Kontraposition. Gegeben sei $x \in U$ mit $df_x \neq 0$. Dann gibt es ein $h \in \mathbb{R}^n$ mit $df_x(h) \neq 0$ sagen wir (evtl. nach einem Vorzeichenwechsel von h) $df_x(h) > 0$. Für alle $t \in \mathbb{R}$ genügend nahe bei 0 ist dann $f(x + th)$ definiert, und es gilt

$$f(x + th) - f(x) = t df_x(h) + o(t) \text{ für } t \to 0.$$

was für $t > 0$ nahe bei 0 positiv und für $t < 0$ nahe bei 0 negativ ist. Der Punkt x ist also kein lokales Extremum.

□

Aus der Analysis 1 wissen Sie für eine Nullstelle x der Ableitung $f'(x) = 0$ einer zweifach differenzierbaren Funktionen $f : \mathbb{R} \to \mathbb{R}$:

$$\begin{aligned} f''(x) > 0 &\Rightarrow x \text{ ist lokales Minimum von } f, \\ f''(x) < 0 &\Rightarrow x \text{ ist lokales Maximum von } f, \end{aligned}$$

während im Fall $f''(x) = 0$ alle drei Fälle (lokales Minimum, lokales Maximum, kein lokales Extremum) möglich sind.
Etwas Analoges gilt auch multidimensional. Allerdings ist die 2. Ableitung $D^2 f$ nun eine *Matrix*, die Hessematrix. Wir müssen daher zunächst die richtige Verallgemeinerung von $f''(x) > 0$ für Matrizen definieren:

Definition 2.48 (Positiv/Negativ definite Matrizen) *Es sei $A = A^t \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix. Weiter sei*

$$Q_A : \mathbb{R}^n \to \mathbb{R}, \quad Q_A(h) = h^t A h$$

die zugehörige quadratische Form. (Hierbei wird h als Spaltenvektor, also die Transponierte h^t als Zeilenvektor aufgefasst.) Die Matrix A wird positiv definit genannt, in Zeichen $A > 0$, wenn für alle $h \in \mathbb{R}^n \setminus \{0\}$ gilt: $Q_A(h) > 0$. Analog wird sie negativ definit genannt, in Zeichen $A < 0$, wenn für alle $h \in \mathbb{R}^n \setminus \{0\}$ gilt: $Q_A(h) < 0$.
Sie wird indefinit genannt, wenn es $h \in \mathbb{R}^n$ und $k \in \mathbb{R}^n$ mit $Q_A(h) > 0$ und $Q_A(k) < 0$ gibt.
A wird positiv semidefinit genannt, in Zeichen $A \geq 0$, wenn für alle $h \in \mathbb{R}^n$ gilt: $Q_A(h) \geq 0$. Analog heißt sie negativ semidefinit, in Zeichen $A \leq 0$, wenn für alle $h \in \mathbb{R}^n$ gilt: $Q_A(h) \leq 0$.

Beispiel 2.49 Ist $A = \operatorname{diag}(a_1, \dots, a_n)$ eine Diagonalmatrix, so gilt

$$Q_A(h) = \sum_{j=1}^{n} a_j h_j^2$$

für $h = (h_1, \dots, h_n)^t \in \mathbb{R}^n$. Die Diagonalmatrix A ist positiv definit, falls alle Diagonaleinträge positiv sind, negativ definit, falls alle Diagonaleinträge negativ sind, indefinit, falls sie sowohl positive als auch negative Diagonaleinträge besitzt, positiv semidefinit, falls alle Diagonaleinträge ≥ 0 sind, und negativ definit, falls alle Diagonaleinträge ≤ 0 sind.

Ist $A = A^t \in \mathbb{R}^{n \times n}$ symmetrisch und $T \in \mathbb{R}^{n \times n}$, so ist auch $T^t A T$ wieder symmetrisch, und es gilt für alle $h \in \mathbb{R}^n$:

$$Q_{T^t A T}(h) = h^t T^t A T h = (Th)^t A (Th) = Q_A(Th).$$

Ist zusätzlich T invertierbar, so folgt: Q_A ist genau dann positiv definit, wenn auch Q_{T^tAT} positiv definit ist. Analoge Aussagen gelten für negativ definite, indefinite bzw. semidefinite Matrizen.
In der Linearen Algebra wird gezeigt, dass jede symmetrische Matrix $A = A^t \in \mathbb{R}^{n\times n}$ mit einer *orthogonalen* Matrix $T = (T^t)^{-1} \in \mathbb{R}^{n\times n}$ diagonalisiert werden kann: $D = T^{-1}AT = T^tAT$ mit einer Diagonalmatrix D, deren Diagonaleinträge genau die Eigenwerte von A auflistet.
Damit haben wir folgendes Kriterium:
Es sei $A \in \mathbb{R}^{n\times n}$ mit den Eigenwerten $\lambda_1, \dots, \lambda_n$ (mit Vielfachheit aufgelistet). Dann gilt:

$$\begin{aligned}
A \text{ ist positiv definit} &\Leftrightarrow \text{Alle Eigenwerte von } A \text{ sind positiv},\\
A \text{ ist negativ definit} &\Leftrightarrow \text{Alle Eigenwerte von } A \text{ sind negativ},\\
A \text{ ist indefinit} &\Leftrightarrow \text{Manche Eigenwerte von } A \text{ sind positiv, andere negativ},\\
A \text{ ist positiv semidefinit} &\Leftrightarrow \text{Alle Eigenwerte von } A \text{ sind } \geq 0,\\
A \text{ ist negativ semidefinit} &\Leftrightarrow \text{Alle Eigenwerte von } A \text{ sind } \leq 0.
\end{aligned}$$

Warnung: Zwar sind die Diagonaleinträge einer positiv definiten Matrix positiv und die Diagonaleinträge einer negativ definiten Matrix negativ. Dennoch sollte man positive Definitheit *nicht* mit der Positivität aller Matrixeinträge verwechseln! Zum Beispiel hat

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

die Eigenwerte 1 und 3, ist also trotz teils negativer Einträge positiv definit. Es ist die Hessematrix von

$$\frac{1}{2}Q_A(x,y) = x^2 + y^2 - xy = \frac{1}{4}[(x+y)^2 + 3(x-y)^2].$$

Andererseits hat

$$B = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

die Eigenwerte 3 und -1, ist also trotz positiver Einträge indefinit. Es ist die Hessematrix von

$$\frac{1}{2}Q_B(x,y) = \frac{1}{2}(x^2 + y^2 + 4xy) = \frac{1}{4}[3(x+y)^2 - (x-y)^2].$$

Um zu entscheiden, ob eine gegebene symmetrische Matrix $A = A^t \in \mathbb{R}^{n\times n}$ positiv definit ist, ist es nicht nötig, alle Eigenwerte auszurechnen, nur um zu entscheiden, ob sie alle positiv sind. Ein anderes Kriterium liefert (hier unbewiesen, da der Beweis eher zur Linearen Algebra gehört):

Bemerkung 2.50 (Determinantenkriterium für positive Definitheit) *Für eine symmetrische Matrix $A = A^t = (a_{i,j})_{i,j=1,\dots,n} \in \mathbb{R}^{n\times n}$ sind äquivalent:*

1. *A ist positiv definit.*

2. *$-A$ ist negativ definit.*

3. *Alle Unterdeterminanten der Gestalt* $\det(a_{i,j})_{i,j=1,\ldots,k}$*, mit* $k = 1, \ldots, n$ *sind positiv, also alle Unterdeterminanten "in der linken oberen Ecke".*

4. *Alle diagonalen Unterdeterminanten* $\det(a_{i_j,i_k})_{j,k=1,\ldots,l}$ *mit* $l = 1, \ldots, n$ *und* $1 \leq i_1 < i_2 < \ldots < i_l \leq n$ *sind positiv.*

Zum Beispiel ist der $1,1$-Eintrag 2 der Matrix

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

positiv und $\det A = 3 > 0$, also ist A positiv definit. Andererseits ist

$$\det B = \det \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} < 0,$$

also B nicht positiv definit.
Ein alternatives Kriterium zur Entscheidung über positive Definitheit liefert:

Verfahren 2.51 (Quadratische Ergänzung, Cholesky-Zerlegung) Gegeben eine symmetrische Matrix $A = A^t = (a_{i,j})_{i,j=1,\ldots,n} \in \mathbb{R}^{n\times n}$ und die zugehörige quadratische Form

$$Q_A(x_1, \ldots, x_n) = (x_1, \ldots, x_n) A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i a_{i,j} x_j = \sum_{i=1}^{n} a_{i,i} x_i^2 + \sum_{i=2}^{n} \sum_{j=1}^{i-1} 2a_{i,j} x_i x_j$$

kann man so entscheiden, ob A positiv definit ist:

1. Falls $a_{1,1} \leq 0$, ist A nicht positiv definit.

2. Falls $a_{1,1} > 0$ und $n = 1$, ist $A = (a_{1,1})$ positiv definit.

3. Andernfalls, also im Fall $a_{1,1} > 0$ und $n > 1$, ergänzt man alle Terme, in denen die Variable x_1 vorkommt, quadratisch, schreibt also

   $$Q_A(x_1, \ldots, x_n) = a_{1,1} \left(x_1 + \sum_{j=2}^{n} \frac{a_{1,j}}{a_{1,1}} x_j \right)^2 + Q_B(x_2, \ldots, x_n)$$

   mit der $(n-1) \times (n-1)$-Matrix $B = (b_{i,j})_{i,j=2,\ldots,n} \in \mathbb{R}^{(n-1)\times(n-1)}$ mit den Einträgen

   $$b_{i,j} = a_{i,j} - \frac{a_{1,i} a_{1,j}}{a_{1,1}}.$$

   Die Matrix A ist positiv definit genau dann, wenn $a_{1,1} > 0$ gilt und B positiv definit ist. Rekursiv setzt man dann das gleiche Verfahren für B "in einer Dimension niedriger" fort, um zu entscheiden, ob B positiv definit ist.

Mit diesem Verfahrens erhält man für positiv definite Matrixen A eine Darstellung

$$Q_A(x_1, \ldots, x_n) = \sum_{i=1}^{n} d_i \cdot \left(x_i + \sum_{j=i+1}^{n} t_{i,j} x_j \right)^2$$

mit positiven $d_1, \ldots d_n$. Das Gleiche in Matrixnotation:

$$A = T^t D T$$

mit der oberen Dreiecksmatrix $T = (t_{i,j})_{i,j=1,\ldots,n}$ mit Einsen in der Diagonalen und der Diagonalmatrix D mit positiven Diagonaleinträgen $d_1, \ldots d_n$. Diese Zerlegung positiv definiter Matrizen heißt *Cholesky-Zerlegung*. Dieses Verfahren eignet sich auch zur Implementierung auf dem Computer.

Beispiel 2.52 *Entscheiden wir mit quadratischer Ergänzung, ob die Matrix*

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 7 & -1 \\ 1 & -1 & 5 \end{pmatrix}$$

positiv definit ist. Wir rechnen:

$$\begin{aligned} Q_A(x, y, z) &= x^2 + 4xy + 2xz + 7y^2 - 2yz + 5z^2 \\ &= (x + 2y + z)^2 + 3y^2 - 6yz + 4z^2 \\ &= (x + 2y + z)^2 + 3(y - z)^2 + z^2, \end{aligned}$$

also ist A positiv definit. Das Gleiche in Matrixnotation:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

Übung 2.53 Entscheiden Sie für folgende Matrizen

a) $A = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 5 & 1 \\ 1 & 1 & 3 \end{pmatrix}$

b) $B = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

ob sie positiv definit sind

(i) indem Sie das charakteristische Polynom ausrechnen und entscheiden, ob alle Nullstellen davon positiv sind,

(ii) mit Vorzeichen von Unterdeterminanten,

(iii) mit quadratischer Ergänzung.

Übung 2.54 *Beweisen Sie, dass die Matrix*

$$A = \begin{pmatrix} 1 & -2 & 1 & -1 \\ -2 & 8 & -4 & 4 \\ 1 & -4 & 3 & -3 \\ -1 & 4 & -3 & 4 \end{pmatrix}$$

positiv definit ist, indem Sie ihre Cholesky-Zerlegung berechnen.

Hier nun die Klassifizierung stationärer Punkte mit der zweiten Ableitung:

Lemma 2.55 (Hinreichendes Kriterium für lokale Extrema) Es sei $f : U \to \mathbb{R}$ zweimal stetig differenzierbar, $U \subseteq \mathbb{R}^n$ offen und $x \in U$ ein stationärer Punkt von f. Dann gilt:

1. Ist die Hessematrix $D^2 f(x)$ *positiv definit*, so ist x ein *lokales Minimum* von f.

2. Ist die Hessematrix $D^2 f(x)$ *negativ definit*, so ist x ein *lokales Maximum* von f.

3. Ist die Hessematrix $D^2 f(x)$ *indefinit*, so ist x kein lokales Extremum von f. In diesem Fall nennt man x einen *Sattelpunkt* von f. Die Funktion f fällt dann in manchen Richtungen in der Nähe von x ab und steigt in anderen Richtungen an. Genauer gesagt: Es gibt dann $h, k \in \mathbb{R}^2$, so dass für genügend kleine $\epsilon > 0$ die Abbildung $]-\epsilon, \epsilon[\ \ni s \mapsto f(x + sh)$ ein Minimum bei 0 und $]-\epsilon, \epsilon[\ \ni s \mapsto f(x + sk)$ ein Maximum bei 0 besitzt.

4. Ist die Hessematrix $D^2 f(x)$ *positiv semidefinit*, aber nicht gleich 0, so ist x *entweder ein lokales Minimum* von f *oder kein lokales Extremum*. Die 2. Ordnung Taylorentwicklung reicht in diesem Fall nicht aus, zu entscheiden, welcher der beiden Fälle eintritt.

5. Ist die Hessematrix $D^2 f(x)$ *negativ semidefinit*, aber nicht gleich 0, so ist x *entweder ein lokales Maximum* von f *oder kein lokales Extremum*. Die 2. Ordnung Taylorentwicklung reicht auch in diesem Fall nicht aus, zu entscheiden, welcher der beiden Fälle eintritt.

6. Verschwindet die Hessematrix: $D^2 f(x) = 0$, so sind *alle Fälle möglich*: lokales Minimum, lokales Maximum, kein lokales Extremum. Die 2. Ordnung Taylorentwicklung reicht dann nicht zur Entscheidung aus.

Beweis. Wir verwenden die Taylorentwicklung 2. Ordnung bei x. Wegen $df_x = 0$ gilt:

$$f(x+h) - f(x) = \frac{1}{2} h^t \cdot D^2 f(x) \cdot h + o(\|h\|^2) \text{ für } h \to 0.$$

1. Ist $D^2 f(x)$ positiv definit, so wird durch

$$\langle \cdot, \cdot \rangle_{D^2 f(x)} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \quad \langle u, v \rangle_{D^2 f(x)} = u^t \cdot D^2 f(x) \cdot v$$

ein Skalarprodukt auf $\mathbb{R}^n$ definiert. Weil nach Satz 1.169 *alle* Normen auf $\mathbb{R}^n$ äquivalent sind, kann auch die zugehörige Norm

$$\mathbb{R}^n \ni h \mapsto \|h\|_{D^2 f(x)} = \sqrt{\langle h, h \rangle_{D^2 f(x)}}$$

nach unten mit der gegebenen Norm $\|\cdot\|$ abgeschätzt werden:

$$\exists c > 0 \; \forall h \in \mathbb{R}^n : \; \|h\|_{D^2 f(x)} \geq c\|h\|.$$

Es folgt:

$$f(x+h) - f(x) = \frac{1}{2}\|h\|^2_{D^2 f(x)} + o(\|h\|^2) \geq \frac{1}{2} c^2 \|h\|^2 + o(\|h\|^2) > 0$$

für alle $h \neq 0$ ein einer genügend kleinen Umgebung von $0 \in \mathbb{R}^n$.
2. Die Aussage 2. folgt aus 1., indem man $-f$ statt f betrachtet.
Für die übrigen Fälle beobachten wir: Ist $h \in \mathbb{R}^n$ mit $h^t D^2 f(x) h \neq 0$, so gilt für $\mathbb{R} \ni s \to 0$:

$$f(x+sh) - f(x) = \frac{1}{2} h^t D^2 f(x) h \, s^2 + o(s^2),$$

also haben $f(x+sh) - f(x)$ und $h^t D^2 f(x) h$ für s nahe bei 0 das gleiche Vorzeichen.
Es folgt die Aussage 3. und der erste Satz in 4. und 5.
Dass die Taylorentwicklung 2. Ordnung in den Fällen 4. und 5. keine vollständige Entscheidung erlaubt, zeigen folgende Beispiele: Ist $A \in \mathbb{R}^{n \times n}$ symmetrisch, so besitzt

$$f(x+h) := \frac{1}{2} h^t A h + c\|h\|_2^4$$

die Hessematrix $D^2 f(x) = A$. Ist A positiv (bzw. negativ) semidefinit, aber nicht positiv (bzw. negativ) definit, so ist x ein Minimum (bzw. Maximum) von f für $c > 0$ (bzw. $c < 0$), aber kein lokales Extremum für $c < 0$ (bzw. $c > 0$).
6. Die Beispiele $f(x+h) = f(x) + h_1^4$, $f(x+h) = f(x) - h_1^4$ und $f(x+h) = f(x) + h_1^3$ zeigen, dass die Taylorentwicklung 2. Ordnung nicht zur Entscheidung ausreicht, wenn $Df(x) = 0$ und $D^2 f(x) = 0$ gilt.

□

Übung 2.56 Finden Sie alle stationären Punkte der Funktionen

(a) $f : \mathbb{R}^2 \to \mathbb{R},\ f(x,y) = ((x+y)^2 - 1)^2 + (x-y)^2$

(b) $f : \mathbb{R}^2 \to \mathbb{R},\ f(x,y) = 2(x+y)^3 - 3(x+y)^2 + y^2$

(c) $f : \mathbb{R}^2 \to \mathbb{R},\ f(x,y) = \operatorname{Re}\cos(x+iy)$ und $g : \mathbb{R}^2 \to \mathbb{R},\ f(x,y) = \operatorname{Im}\cos(x+iy)$.

und entscheiden Sie jeweils, ob es sich um lokale Minima, lokale Maxima oder um Sattelpunkte handelt. Skizzieren Sie jeweils ein Niveaulinienbild der Funktionen (bei (c) für f und g in ein gemeinsames Bild).

2.8 Die Räume C_b^1

In diesem Abschnitt seien $(V, \|\cdot\|)$ ein normierter Raum, $(U, \|\cdot\|)$ ein Banachraum und $V_1 \subseteq V$ offen. Wer will, kann sich $V = \mathbb{R}^m$, $U = \mathbb{R}^n$ vorstellen.

Wir versehen den Raum $\mathcal{B}(V,U)$ aller stetigen linearen Abbildungen $L : V \to U$ stets mit der Operatornorm
$$\|L\|_{V\to U} = \sup_{\substack{x\in V \\ \|x\|\le 1}} \|L(x)\|.$$

Übung 2.57 *Zeigen Sie, dass der Raum $(\mathcal{B}(V,U), \|\cdot\|_{V\to U})$ vollständig ist.*

Auf dem Raum $C(V_1, \mathcal{B}(V,U))$ aller bezüglich der Operatornorm $\|\cdot\|_{V\to U}$ stetigen Funktionen $G : V_1 \to \mathcal{B}(V,U)$ definieren wir
$$\|\cdot\|_\infty : C(V_1, \mathcal{B}(V,U)) \to \mathbb{R} \cup \{\infty\}, \quad \|G\|_\infty := \sup\{\|G(y)\|_{V\to U} \mid y \in V_1\}.$$
Dies steht in Analogie zu
$$\|g\|_\infty := \sup\{\|g(y)\| \mid y \in V_1\}$$
für Funktionen $g : V_1 \to U$. Der Raum aller stetigen, *beschränkten* Funktionen von V_1 mit Werten in $\mathcal{B}(V,U)$ wird mit
$$C_b(V_1, \mathcal{B}(V,U)) := \{G \in C(V_1, \mathcal{B}(V,U)) \mid \|G\|_\infty < \infty\}$$
bezeichnet. Nach Satz 1.107 ist der Raum $(C_b(V_1, \mathcal{B}(V,U)), \|\cdot\|_\infty)$ vollständig.
Schließlich definieren wir den *Raum der beschränkten stetig differenzierbaren Funktionen auf V_1 mit Werten in U mit beschränkter Ableitung*:
$$C_b^1(V_1, U) := \{g \in C_b(V_1, U) \mid g \text{ ist stetig differenzierbar mit } \|dg\|_\infty < \infty\}$$
und versehen ihn mit der Norm
$$\|\cdot\|_{C^1} : C_b^1(V_1, U) \to \mathbb{R}, \quad \|g\|_{C^1} := \max\{\|g\|_\infty, \|dg\|_\infty\}.$$

Lemma 2.58 *Der normierte Raum $(C_b^1(V_1, U), \|\cdot\|_{C^1})$ ist vollständig.*

Beweis: Es sei $(g_k)_{k\in\mathbb{N}}$ eine Cauchyfolge in $(C_b^1(V_1,U),\|\cdot\|_{C^1})$. Dann ist $(g_k)_{k\in\mathbb{N}}$ auch eine Cauchyfolge im vollständigen Raum $(C_b(V_1,U),\|\cdot\|_\infty)$, also konvergent gegen ein $g \in C_b(V_1,U)$. Ebenso ist $(dg_k)_{k\in\mathbb{N}}$ eine Cauchyfolge im vollständigen Raum $(C_b(V_1,\mathcal{B}(V,U)),\|\cdot\|_\infty)$, also konvergent gegen ein $G \in C_b(V_1,\mathcal{B}(V,U))$. Wir müssen nur noch zeigen, dass g differenzierbar mit Ableitung $dg = G$ ist, denn dann folgt $g \in C_b^1(V_1,U)$ und

$$\|g_k - g\|_{C^1} = \max\{\|g_k - g\|_\infty, \|dg_k - dg\|_\infty\} \xrightarrow{n\to\infty} 0.$$

Wir beweisen das zuerst im Fall $V = \mathbb{R}^m$, $U = \mathbb{R}^n$. Hierzu seien $g^1,\ldots,g^n$ die Komponenten von g, analog $g_k^1,\ldots,g_k^n$ die Komponenten von g_k, sowie $(G_{j,i} : V_1 \to \mathbb{R})_{\substack{i=1,\ldots,n\\ j=1,\ldots,m}}$ die Einträge der Matrix, die G darstellt. Es genügt nach Lemma 2.21 zu zeigen, dass $D_i g^j = G_{j,i}$ für alle i,j gilt. Nach dem Hauptsatz der Differential- und Integralrechnung diese Behauptung äquivalent zu

$$\int_0^s G_{j,i}(x+te_i)\,dt = g^j(x+se_i) - g^j(x)$$

für alle $x \in V_1$ und $s \in \mathbb{R}$ mit $[x, x+se_i] \subseteq V_1$. Nochmals nach dem Haupsatz haben wir

$$\int_0^s D_i g_k^j(x+te_i)\,dt = g_k^j(x+se_i) - g_k^j(x)$$

für die gleichen x und s und alle $k \in \mathbb{N}$. Da $\|g_k - g\|_\infty \xrightarrow{n\to\infty} 0$ wissen wir

$$g_k^j(x+se_i) - g_k^j(x) \xrightarrow{k\to\infty} g^j(x+se_i) - g^j(x).$$

Weiter folgt

$$\sup_{t\in[0,s]} |D_i g_k^j(x+te_i) - G_{j,i}(x+te_i)| \xrightarrow{k\to\infty} 0$$

wegen $\|dg_k - G\|_\infty \xrightarrow{k\to\infty} 0$. Aus dieser *gleichmäßigen* Konvergenz erhalten wir

$$\int_0^s D_i g_k^j(x+te_i)\,dt \xrightarrow{k\to\infty} \int_0^s G_{j,i}(x+te_i)\,dt,$$

also zusammen die Behauptung.

Nun zeigen wir die Behauptung $dg = G$ im allgemeinen Fall. Wir müssen also zeigen:

$$\forall x \in V_1\ \forall \epsilon > 0\ \exists \delta > 0\ \forall y \in U_\delta^{\|\cdot\|}(x):\ y \in V_1 \text{ und } \|g(y) - g(x) - G_x(y-x)\| \le \epsilon\|y-x\|.$$

Hierzu seien $x \in V_1$ und $\epsilon > 0$ gegeben. Wir wählen $\delta > 0$ so klein, dass $U_\delta^{\|\cdot\|}(x) \subseteq V_1$ (möglich, da V_1 offen) und dass für alle $z \in U_\delta^{\|\cdot\|}(x)$ gilt: $\|G_z - G_x\|_{V\to U} \le \epsilon$. Nun sei $y \in U_\delta^{\|\cdot\|}(x)$ gegeben.
Wir verwenden folgende Darstellung der Norm:

Dualitätsprinzip: *Ist $U^* := \mathcal{B}(U,\mathbb{R})$ der topologische Dualraum des Banachraums $(U,\|\cdot\|)$, also die Menge aller stetigen Linearformen $\ell : U \to \mathbb{R}$, versehen mit der Operatornorm $\|\cdot\|_{U\to\mathbb{R}}$, so gilt für alle $x \in U$:*

$$\|x\| = \max\{\ell(x) | \ \ell \in U^*,\ \|\ell\|_{U\to\mathbb{R}} \le 1\} \tag{42}$$

Das Dualitätsprinzip wird in der Funktionalanalysis bewiesen; wir beweisen es hier nicht allgemein. Einen wichtigen Spezialfall kennen Sie jedoch schon: $U = \ell^p(I)$, siehe Korollar 1.22 und Übung 1.74. Für Hilberträume U wird das das Dualitätsprinzip besonders einfach: Ist nämlich $x \in U \setminus \{0\}$, so wird das Maximum in (42) für $\ell : U \to \mathbb{R}$, $\ell(v) = \langle x, v\rangle / \|x\|$ angenommen. In der Tat ist $\|\ell\|_{U\to\mathbb{R}} = 1$, denn für alle $v \in U$ gilt $\ell(v) = \langle x, v\rangle / \|x\| \le \|v\|$ nach der Cauchy-Schwarz-Ungleichung.

Wir wählen mit dem Dualitätsprinzip eine stetige Linearform[17] $\ell \in U^*$ mit $\|\ell\|_{U\to\mathbb{R}} \le 1$ und

$$\|g(y) - g(x) - G_x(y - x)\| = \ell(g(y) - g(x) - G_x(y - x)).$$

Nach dem Hauptsatz der Differential- und Integralrechnung und der Kettenregel erhalten wir für alle $k \in \mathbb{N}$ wegen $\|g_k - g\|_\infty \overset{k\to\infty}{\longrightarrow} 0$ und $\|Dg_k - G\|_\infty \overset{k\to\infty}{\longrightarrow} 0$ die Abschätzung

$$\begin{aligned}
&\|g(y) - g(x) - G_x(y - x)\| = \ell(g(y) - g(x) - G_x(y - x)) \\
&= \lim_{k\to\infty} \ell(g_k(y) - g_k(x) - (dg_k)_x(y - x)) \\
&= \lim_{k\to\infty} \int_0^1 \frac{d}{ds}\ell(g_k(x + s(y - x)) - g_k(x) - s(dg_k)_x(y - x))\, ds \\
&= \lim_{k\to\infty} \int_0^1 \ell((dg_k)_{x+s(y-x)}(y - x) - (dg_k)_x(y - x))\, ds \\
&\le \limsup_{k\to\infty} \int_0^1 \|\ell\|_{U\to R}\|(dg_k)_{x+s(y-x)} - (dg_k)_x\|_{V\to U}\|y - x\|\, ds \\
&\le \limsup_{k\to\infty} \sup_{0\le s\le 1} \|(dg_k)_{x+s(y-x)} - (dg_k)_x\|_{V\to U}\|y - x\| \\
&\le \limsup_{k\to\infty} \sup_{z\in U_\delta^{\|\cdot\|}(x)} \|(dg_k)_z - (dg_k)_x\|_{V\to U}\|y - x\|.
\end{aligned} \tag{43}$$

Für $z \in U_\delta^{\|\cdot\|}(x)$ schätzen wir ab:

$$\begin{aligned}
&\|(dg_k)_z - (dg_k)_x\|_{V\to U} \le \\
&\|(dg_k)_z - G_z\|_{V\to U} + \|G_z - G_x\|_{V\to U} + \|G_x - (dg_k)_x\|_{V\to U} \le \|G_z - G_x\|_{V\to U} + 2\|G - dg_k\|_\infty,
\end{aligned}$$

also

$$\begin{aligned}
&\sup_{z\in U_\delta^{\|\cdot\|}(x)} \|(dg_k)_z - (dg_k)_x\|_{V\to U} \\
&\le \sup_{z\in U_\delta^{\|\cdot\|}(x)} \|G_z - G_x\|_{V\to U} + 2\|G - dg_k\|_\infty \le \epsilon + 2\|G - dg_k\|_\infty \overset{k\to\infty}{\longrightarrow} \epsilon.
\end{aligned}$$

Eingesetzt in (43) folgt hieraus die Behauptung:

$$\|g(y) - g(x) - G_x(y - x)\| \le \epsilon\|x - y\|.$$

□

Wir besprechen nun eine Operatorenversion der geometrischen Reihe. Für $L \in \mathcal{B}(U, U)$ definieren wir die Potenzen von L rekursiv: $L^0 = \mathrm{id}_U$, $L^{j+1} = L \circ L^j$ für $j \in \mathbb{N}_0$.

[17] Die Linearform ℓ spielt hier im allgemeinen Fall eine ähnliche Rolle wie das Bilden der j-ten Komponente oben im Spezialfall: Sie reduziert den multidimensionalen Fall auf eindimensionale Werte.

Lemma 2.59 (geometrische Reihe, Neumann-Reihe) *Sind $(U, \|\cdot\|)$ ein Banachraum und $L \in \mathcal{B}(U,U)$ mit $\|L\|_{U\to U} < 1$, so besitzt $\mathrm{id}_U - L : U \to U$ die Inverse*

$$(\mathrm{id}_U - L)^{-1} = \sum_{j=0}^{\infty} L^j \in \mathcal{B}(U,U),$$

wobei die Konvergenz der Reihe bezüglich der Operatornorm $\|\cdot\|_{U\to U}$ gemeint ist. Genauer gesagt gilt:

$$\left\| \sum_{j=0}^{N} L^j - (\mathrm{id}_U - L)^{-1} \right\|_{U\to U} \le \frac{\|L\|_{U\to U}^N}{1 - \|L\|_{U\to U}} \overset{N\to\infty}{\longrightarrow} 0.$$

Die Abbildung

$$\{L \in \mathcal{B}(U,U) |\ \|L\|_{U\to U} < 1\} \to \mathcal{B}(U,U), \quad L \mapsto (\mathrm{id}_U - L)^{-1}$$

ist stetig bezüglich der Operatornorm $\|\cdot\|_{U\to U}$.

Beweis: Für $N, M \in \mathbb{N}_0$ gilt wegen der Submultiplikativität der Operatornorm:

$$\begin{aligned}
&\left\| \sum_{j=0}^{N+M} L^j - \sum_{j=0}^{N} L^j \right\|_{U\to U} = \left\| \sum_{j=N+1}^{N+M} L^j \right\|_{U\to U} \le \sum_{j=N+1}^{N+M} \|L^j\|_{U\to U} \\
&\le \sum_{j=N+1}^{N+M} \|L\|_{U\to U}^j = \|L\|_{U\to U}^{N+1} \frac{1 - \|L\|_{U\to U}^M}{1 - \|L\|_{U\to U}} \le \frac{\|L\|_{U\to U}^{N+1}}{1 - \|L\|_{U\to U}}
\end{aligned}$$

und daher für gegebenes $0 < k < 1$:

$$\sup_{M\in\mathbb{N}} \sup_{\substack{L\in\mathcal{B}(U,U) \\ \|L\|_{U\to U}<k}} \left\| \sum_{j=0}^{N+M} L^j - \sum_{j=0}^{N} L^j \right\|_{U\to U} \le \frac{k^{N+1}}{1-k} \overset{N\to\infty}{\longrightarrow} 0.$$

Die Folge

$$\left(L \mapsto \sum_{j=0}^{N} L^j \right)_{N\in\mathbb{N}}$$

ist also eine Cauchyfolge im vollständigen Raum $(C_b(U_k^{\|\cdot\|_{U\to U}}(0), \mathcal{B}(U,U)), \|\cdot\|_\infty)$ und daher konvergent. Insbesondere ist die Grenzfunktion $L \mapsto \sum_{j=0}^\infty L^j$ stetig auf $U_k^{\|\cdot\|_{U\to U}}(0)$ für alle $0 < k < 1$ und daher auch stetig auf $U_1^{\|\cdot\|_{U\to U}}(0)$. Für den Limes der Folge (bezüglich $\|\cdot\|_{U\to U}$ und gleichmäßig auf $U_k^{\|\cdot\|_{U\to U}}(0)$) gilt:

$$(\mathrm{id}_U - L) \circ \sum_{j=0}^{\infty} L^j = \lim_{N\to\infty} (\mathrm{id}_U - L) \circ \sum_{j=0}^{N} L^j = \lim_{N\to\infty} (\mathrm{id}_U - L^{N+1}) = \mathrm{id}_U$$

und analog

$$\left(\sum_{j=0}^{\infty} L^j\right) \circ (\mathrm{id}_U - L) = \mathrm{id}_U,$$

für $\|L\|_{U\to U} < 1$. (Hier verwenden wir die Stetigkeit der Kompositionen von links und von rechts mit $\mathrm{id}_U - L$, die aus der Submultiplikativität der Operatornorm folgt.) Es folgt die Behauptung.

□

Korollar 2.60 *Es seien $(U, \|\cdot\|)$ und $(V, \|\cdot\|)$ zwei Banachräume. Dann ist die Menge*

$$\mathrm{GL}(V, U) := \{L \in \mathcal{B}(V, U) |\ L : V \to U \text{ ist bijektiv mit } L^{-1} \in \mathcal{B}(U, V)\} \subseteq \mathcal{B}(V, U)$$

offen, und die Inversenbildung

$$\cdot^{-1} : \mathrm{GL}(V, U) \to \mathrm{GL}(U, V)$$

ist stetig bezüglich der Operatornormen $\|\cdot\|_{U\to V}$ und $\|\cdot\|_{V\to U}$.

Beweis Für jedes $L_0 \in GL(V, U)$ ist

$$U_L := \{L \in GL(V, U) |\ \|L_0^{-1} \circ L - \mathrm{id}_V\|_{V\to V} < 1\} \subseteq GL(V, U)$$

eine offene Umgebung von L_0, da Komposition mit L_0^{-1} stetig ist. Da $L^{-1} = (L_0^{-1} \circ L)^{-1} \circ L_0^{-1}$ für $L \in U_L$ gilt, folgt die Behauptung aus Lemma 2.59.

□

Übung 2.61 (Ableitung der Matrix-Exponentialfunktion) *Es sei $n \in \mathbb{N}$.*

1. *Zeigen Sie für gegebenes $j \in \mathbb{N}$, dass die Matrix-Potenzfunktion $f : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$, $f(A) = A^j$ die Ableitung $df_A(B) = \sum_{k=1}^{j} A^{k-1} B A^{j-k}$ für $A, B \in \mathbb{R}^{n\times n}$ besitzt.*

2. *Es sei $U \subseteq \mathbb{R}^{n\times n}$ offen und beschränkt. Zeigen Sie, dass die Matrix-Exponentialreihe*

$$U \ni A \mapsto \exp(A) = \sum_{j=1}^{\infty} \frac{1}{j} A^j$$

   in $(C_b^1(U, \mathbb{R}^{n\times n}), \|\cdot\|_{C^1})$ konvergiert.

3. *Folgern Sie für $A, B \in \mathbb{R}^{n\times n}$:*

$$d\exp_A(B) = \int_0^1 e^{tA} B e^{(1-t)A}\, dt.$$

   Insbesondere gilt $d\exp_{\mathrm{Id}} = \mathrm{id}_{\mathbb{R}^{n\times n}}$.

Übung 2.62 (Funktionale Version der Ableitungsregel $d\frac{1}{1-x} = \frac{dx}{(1-x)^2}$) *Es sei $(U, \|\cdot\|)$ ein Banachraum. Wir versehen $\mathcal{B}(U,U)$ wie immer mit der Operatornorm $\|\cdot\|_{U\to U}$.*

1. *Zeigen Sie für gegebenes $j \in \mathbb{N}$, dass die Operator-Potenzfunktion $f : \mathcal{B}(U,U) \to \mathcal{B}(U,U)$,*
$$f(A) = A^j = \underbrace{A \circ \ldots \circ A}_{j\text{-mal}}$$
die Ableitung $df_A(B) = \sum_{k=1}^{j} A^{k-1} \circ B \circ A^{j-k}$ für $A, B \in \mathcal{B}(U,U)$ besitzt.

2. *Es sei $0 < k < 1$ und $U_k := \{L \in \mathcal{B}(U,U) |\ \|L\|_{U\to U} < k\}$. Zeigen Sie, dass die geometrische Reihe*
$$f : U_k \ni A \mapsto (\mathrm{id}_U - A)^{-1} = \sum_{j=1}^{\infty} A^j$$
in $(C_b^1(U_k, U), \|\cdot\|_{C^1})$ konvergiert.

3. *Folgern Sie für $A \in U_k$ und $B \in \mathcal{B}(U,U)$:*
$$df_A(B) = f(A) \circ B \circ f(A).$$

2.9 Der lokale Umkehrsatz

Eine differenzierbare Funktion f hat nahe bei einer Stelle x sehr ähnliche Eigenschaften wie ihre Linearisierung $f(x) + df_x$ bei x.

Diese vage Aussage bildet das Leitprinzip der Differentialrechnung. Sie lässt sich auf viele Weisen zu präzisen Aussagen ausformen. Eine erste solche Präzisierung kennen Sie schon: die Kettenregel.
Wir besprechen nun einen weiteren Satz in dieser Philosophie: f ist *nahe bei x* invertierbar, wenn dort ihre Linearisierung invertierbar ist. Genauer gesagt gilt:

Satz 2.63 (Lokaler Umkehrsatz) *Es seien $(U, \|\cdot\|)$ und $(V, \|\cdot\|)$ Banachräume, $U_0 \subseteq U$ offen, $f : U_0 \to V$ stetig differenzierbar und $x_0 \in U_0$. Weiter sei die Ableitung $df_{x_0} : U \to V$ bijektiv mit einer stetigen Inversen ${df_{x_0}}^{-1} : V \to U$. Dann gibt es eine offene Umgebung $U_1 \subseteq U_0$ von x_0 und eine offene Umgebung $V_1 \subseteq V$ von $y_0 := f(x_0)$ mit folgenden Eigenschaften:*

1. *Die Einschränkung $f|_{U_1}$ von f ist eine Bijektion $f|_{U_1} : U_1 \to V_1$ mit stetiger Umkehrabbildung $h := (f|_{U_1})^{-1} : V_1 \to U_1$.*

2. *Die Umkehrabbildung $h : V_1 \to U_1$ ist stetig differenzierbar. Es gilt für alle $x \in U_1$ und $y \in V_1$ mit $y = f(x)$:*
$$dh_y = {df_x}^{-1}$$

3. *Ist f sogar p-fach stetig differenzierbar, $p \in \mathbb{N}$, so ist auch die Umkehrabbildung $h : V_1 \to U_1$ p-fach stetig differenzierbar.*

In der Definition 2.14 der Ableitung haben wir gefordert, dass df_{x_0} eine *stetige* lineare Abbildung ist. In der Funktionalanalysis wird gezeigt, dass die Inverse einer stetigen linearen Bijektion zwischen zwei Banachräumen stets stetig ist. Daher ist es eigentlich überflüssig vorauszusetzen, dass die Inverse $(df_{x_0})^{-1} : V \to U$ stetig sein soll, weil das unter den übrigen Voraussetzungen automatisch der Fall ist.
Der **wichtigste Fall** ist $U = V = \mathbb{R}^n$. In diesem Fall ist die Ableitung $df_{x_0} : U \to V$ genau dann bijektiv, wenn die Jacobimatrix $Df(x_0)$ invertierbar ist. In diesem Fall gilt

$$Dh(y) = Df(h(y))^{-1} \text{ für } y \in V_1.$$

Wer will, kann sich im folgenden Beweis diesen Fall $U = \mathbb{R}^n$ vorstellen, der etwas anschaulicher ist.

Beweis von Teil 1 des Satzes 2.63:
Um den Beweis möglichst durchsichtig und die Notation möglichst einfach zu halten, beweisen wir zunächst einen Spezialfall. Hierzu nehmen wir an:

$$x_0 = 0,\ f(0) = 0,\ U = V \text{ und } df_{x_0} = \mathrm{id}_U\,.$$

Der allgemeine Fall wird anschließend auf diesen Spezialfall zurückgeführt. Wer will, kann sich den Spezialfall $U = \mathbb{R}^n$ vorstellen, der etwas anschaulicher ist.
In diesem Fall ist $\mathrm{id}_U(x) = x$ die Linearisierung von f bei $x_0 = 0$. Insbesondere gilt:

$$f(x) = f(x) - f(0) = df_0(x) + o(\|x\|) = x + o(\|x\|) \text{ für } x \to 0.$$

Der Beweis beruht auf dem Banachschen Fixpunktsatz. Für die gesuchte Inverse $h : V_1 \to U$ von $f|_{U_1}$ mit geeigneten Mengen U_1, V_1 setzen wir die Bestimmungsgleichung

$$f(h(y)) = y \text{ für } y \in V_1 \tag{44}$$

an. Für die Anwendung des Banachschen Fixpunktsatzes schreiben wir sie in Fixpunktform:

$$h(y) = h(y) - f(h(y)) + y \tag{45}$$

Zur Motivation dieser Wahl der Fixpunktform folgendes heuristisches Argument: Wegen $df_0 = \mathrm{id}_U$ wird f in der Nähe von 0 gut durch die Identität approximiert. Die rechte Seite $h(y) - f(h(y)) + y$ wird also nur schwach von h abhängen, so dass wir gute Kontraktionseigenschaften erwarten dürfen.

Präzisieren wir das Argument: Wir fixieren ein $k \in \mathbb{R}$ mit $0 < k < 1$, z.B. $k = 1/2$. Die Zahl k spielt unten die Rolle der Kontraktionskonstanten im Banachschen Fixpunktsatz. Wir kürzen ab: $r : U_0 \to U$, $r(x) = f(x) - x$. Insbesondere gilt $dr_0 = df_0 - \mathrm{id}_U = 0$. Wegen der Stetigkeit der Ableitungsfunktion $dr : U_0 \to \mathcal{B}(U, U)$ in $0 \in U_0$ können wir $\epsilon > 0$ so klein wählen, dass

$$U_2 := \{x \in U |\ \|x\| < \epsilon\} \subseteq U_0$$

und

$$\|dr_x\|_{U\to U} \leq k \text{ für alle } x \in U_2$$

gilt. (Man beachte, dass $df_x = dr_x + \mathrm{id}_U$ für alle $x \in U_2$ nach Lemma 2.59 (geom. Reihe) invertierbar mit stetiger Inverser ist. Das wird später noch nützlich sein.) Die ϵ-Umgebung $U_2 \subseteq U$ ist offen. Wir wählen $\delta > 0$ mit $\delta < (1-k)\epsilon$ und erhalten

$$B := \{x \in U | \, \|x\| \leq \delta/(1-k)\} \subseteq U_2.$$

Man beachte, dass B in U abgeschlossen ist. Wir setzen

$$V_1 := \{y \in V = U | \, \|y\| < \delta\}.$$

Wir wenden den Banachschen Fixpunktsatz in der Teilmenge

$$M := \{g \in C(V_1, B) | \, g(0) = 0\} = \{g \in C_b(V_1, U) | \, \|g\|_\infty \leq \delta/(1-k), \ g(0) = 0\} \tag{46}$$

des Raums $C_b(V_1, U)$ an, versehen mit der Supremumsnorm

$$\|\cdot\|_\infty : C_b(V_1, U) \to \mathbb{R}, \quad \|g\|_\infty := \sup_{x \in V_1} \|g(x)\|.$$

Man beachte, dass $M \subseteq C_b(V_1, U)$ abgeschlossen ist, denn die Normabbildung $\|\cdot\|_\infty : (C_b(V_1, U), \|\cdot\|_\infty) \to \mathbb{R}$ und die Auswertungsabbildung $\delta_0 : (C_b(V_1, U), \|\cdot\|_\infty) \to U$, $g \mapsto g(0)$ sind stetig. Satz 1.107 zeigt, dass $(C_b(V_1, U), \|\cdot\|_\infty)$ vollständig ist. Also ist nach Lemma 1.106 auch die abgeschlossene Teilmenge $M \subseteq C_b(V_1, U)$ vollständig.
Motiviert durch die Fixpunktgleichung (45) definieren wir für beliebiges $g \in M$:

$$\Phi(g) : V_1 \to U, \quad \Phi(g)(y) := g(y) - f(g(y)) + y = y - r(g(y)). \tag{47}$$

In Kurznotation bedeutet das: $\Phi(g) = \mathrm{id}_{V_1} - r \circ g$. Man beachte, dass $\Phi(g)$ stetig ist, denn g und r sind stetig. Wir müssen nun die Kontraktionseigenschaft

$$\forall g_1, g_2 \in M : \|\Phi(g_1) - \Phi(g_2)\|_\infty \leq k\|g_1 - g_2\|_\infty \tag{48}$$

und

$$\Phi[M] \subseteq M \tag{49}$$

zeigen. Hierzu betrachten wir $x_1, x_2 \in U_2$. Weil die abgeschlossene Kugel U_2 konvex[18] ist (Dreiecksungleichung!), gilt für die Verbindungsstrecke:

$$[x_1, x_2] = \{(1-s)x_1 + sx_2 | \, 0 \leq s \leq 1\} \subseteq B.$$

Wir zeigen nun

$$\|r(x_2) - r(x_1)\| \leq k\|x_1 - x_2\|. \tag{50}$$

[18] Eine Teilmenge $\tilde{U}$ eines $\mathbb{R}$-Vektorraums wird *konvex* genannt, wenn für alle $x, y \in \tilde{U}$ gilt: $[x, y] \subseteq \tilde{U}$.

Im Spezialfall $U = \mathbb{R}^n$ erhalten wir aus dem Hauptsatz der Differential- und Integralrechung und der Kettenregel:

$$\begin{aligned} r(x_2) - r(x_1) &= \int_0^1 \frac{d}{ds} r((1-s)x_1 + sx_2)\, ds \\ &= \int_0^1 dr_{(1-s)x_1+sx_2}(x_2 - x_1)\, ds \end{aligned}$$

also

$$\begin{aligned} \|r(x_2) - r(x_1)\| &\le \int_0^1 \|dr_{(1-s)x_1+sx_2}(x_2 - x_1)\|\, ds \\ &\le \int_0^1 \|dr_{(1-s)x_1+sx_2}\|_{U\to U}\|x_2 - x_1\|\, ds \le k\|x_2 - x_1\|. \end{aligned} \tag{51}$$

Dieses Argument lässt sich auch auf Banachräume U erweitern, wenn man den Riemannschen Integralbegriff und den Hauptsatz auch für stetige Abbildungen $[a, b] \to U$ formuliert. Einen alternativen Weg, der keine Integrale von Abbildungen mit Werten in Banachräumen benötigt, erhalten wir mit dem Dualitätsprinzip (42): Wählen wir ein $\ell \in U^*$ mit $\|\ell\|_{U\to\mathbb{R}} \le 1$ und

$$\|r(x_2) - r(x_1)\| = \ell(r(x_2) - r(x_1)),$$

so folgt aus dem *eindimensionalen* Hauptsatz der Differential- und Integralrechnung und der Kettenregel folgender alternativer Beweis der Ungleichung (50):

$$\begin{aligned} \|r(x_2) - r(x_1)\| &= \ell(r(x_2) - r(x_1)) \\ &= \int_0^1 \frac{d}{ds}\ell(r((1-s)x_1 + sx_2))\, ds \\ &= \int_0^1 \ell(dr_{(1-s)x_1+sx_2}(x_2 - x_1))\, ds \le \int_0^1 |\ell(dr_{(1-s)x_1+sx_2}(x_2 - x_1))|\, ds \\ &\le \int_0^1 \|\ell\|_{U\to\mathbb{R}}\|dr_{(1-s)x_1+sx_2}\|_{U\to U}\|x_2 - x_1\|\, ds \le k\|x_2 - x_1\|. \end{aligned} \tag{52}$$

Man beachte hierbei $d\ell_x = \ell$ für $x \in U$, da ℓ stetig und *linear* ist.

Setzen wir nun für $g_1, g_2 \in M$ und $y \in V_1$ die Werte $x_1 := g_1(y) \in B \subseteq U_2$ und $x_2 := g_2(y) \in B \subseteq U_2$ in (50) ein, folgt

$$\|\Phi(g_1)(y) - \Phi(g_2)(y)\| = \|[y - r(x_1)] - [y - r(x_2)]\| = \|r(x_2) - r(x_1)\| \le k\|g_1(y) - g_2(y)\|$$

und daher die Kontraktionseigenschaft (48). Setzen wir andererseits für $g \in M$ und $y \in V_1$ die Terme $x_2 := g(y)$ und $x_1 = 0$ in (51) ein, folgt

$$\begin{aligned} &\|r(g(y))\| = \|r(x_2)\| = \|r(x_2) - r(0)\| \\ &\le k\|x_2 - 0\| = k\|g(y)\| \le \frac{k\delta}{1-k} \end{aligned}$$

und daher

$$\|\Phi(g)(y)\| = \|y - r(g(y))\| \leq \|y\| + \|r(g(y))\| \leq \delta + \frac{k\delta}{1-k} = \frac{\delta}{1-k}.$$

Dies zeigt $\|\Phi(g)\|_\infty \leq \delta/(1-k)$. Weiter beobachten wir $\Phi(g)(0) = 0 - r(g(0)) = -r(0) = 0$ für $g \in M$. Zusammen folgt $\Phi[M] \subseteq M$.
Also ist der Banachsche Fixpunktsatz auf $\Phi : M \to M$ anwendbar. Er liefert uns, dass es ein $h \in M$ mit $\Phi(h) = h$ gibt. Insbesondere ist $h : V_1 \to B \subseteq U_2 \subseteq U$ stetig, und es gilt die Fixpunktgleichung (45) für alle $y \in V_1$, also auch die äquivalente Form (44), d.h. $f \circ h = \mathrm{id}_{V_1}$. Wir setzen $U_1 := f^{-1}[V_1] \cap U_2$. Wegen der Stetigkeit von f und wegen $f(0) = 0$ ist U_1 eine offene Umgebung von 0, da $V_1 = U_\delta^{\|\cdot\|}(0)$ und $U_2 = U_\epsilon^{\|\cdot\|}(0)$ offene Umgebungen von 0 sind. Für alle $y \in V_1$ gilt $h(y) \in f^{-1}[V_1] \cap U_2 = U_1$ wegen $f(h(y)) = y \in V_1$, es gilt also $h : V_1 \to U_1$. Nach Definition von U_1 gilt $f[U_1] \subseteq V_1$. Es gilt sogar $f[U_1] = V_1$, denn für alle $y \in V_1$ folgt $h(y) \in U_1$, also $y = f(h(y)) \in f[U_1]$. Die Einschränkung $f|_{U_1} : U_1 \to V_1$ ist also surjektiv.
Weiter ist die $f|_{U_1}$ injektiv, denn für alle $x_1, x_2 \in U_1 \subseteq U_2$ mit $x_1 \neq x_2$ gilt wegen (50):

$$\begin{aligned}\|f(x_2) - f(x_1)\| &= \|x_2 - x_1 + r(x_2) - r(x_1)\| \geq \|x_2 - x_1\| - \|r(x_2) - r(x_1)\| \\ &\geq \|x_2 - x_1\| - k\|x_2 - x_1\| = (1-k)\|x_2 - x_1\| > 0.\end{aligned}$$

Also ist $f|_{U_1} : U_1 \to V_1$ eine Bijektion. Wegen $f \circ h = \mathrm{id}_{V_1}$ ist $h : V_1 \to U_1$ die Inverse von $f|_{U_1} : U_1 \to V_1$.

Nun beweisen wir, dass die Funktion h differenzierbar in 0 mit der Ableitung $dh_0 = \mathrm{id}_U$ ist. Wir müssen also zeigen:

$$\forall \epsilon_1 > 0\ \exists \delta_1 > 0\ \forall y \in V_1 :\ \|y\| < \delta_1 \Rightarrow \|h(y) - y\| \leq \epsilon_1 \|y\|. \tag{53}$$

Hierzu sei $\epsilon_1 > 0$ gegeben. Weil die Funktion f differenzierbar in 0 mit der Ableitung $df_0 = \mathrm{id}_U$ ist, finden wir ein $\delta_2 > 0$, so dass

$$\|r(x)\| = \|f(x) - x\| \leq (1-k)\epsilon_1 \|x\|$$

für alle $x \in U_{\delta_2}^{\|\cdot\|}(0)$ gilt; man beachte $k < 1$. Wir setzen $\delta_1 = (1-k)\min\{\epsilon_1, \delta_2\} > 0$. Nun sei $y \in V_1$ mit $\|y\| < \delta_1$ gegeben. Setzen wir $x = h(y) \in U_1 \subseteq U_2$, also $f(x) = y$, so folgt wegen (50):

$$\|x\| - \|y\| \leq \|y - x\| = \|f(x) - x\| = \|r(x)\| = \|r(x) - r(0)\| \leq k\|x - 0\| = k\|x\|$$

und daher wegen $k < 1$

$$\begin{aligned}&(1-k)\|x\| = \|x\| - k\|x\| \leq \|y\|, \\ &\|x\| \leq \frac{\|y\|}{1-k} < \frac{\delta_1}{1-k} \leq \delta_2, \\ &\|h(y) - y\| = \|x - f(x)\| = \|r(x)\| \leq (1-k)\epsilon_1\|x\| \leq \epsilon_1\|y\|.\end{aligned}$$

Damit ist (53), also $dh_0 = \mathrm{id}_U$ gezeigt.

Nun führen wir den allgemeinen Fall von Teil 1 des lokalen Umkehrsatzes auf den eben bewiesenen Spezialfall zurück.
Gegeben $f : U_0 \to V$, $x_0 \in U_0$ und $y_0 = f(x_0)$ mit $L := {df_{x_0}}^{-1} \in \mathcal{B}(V, U)$, definieren wir folgende "*in den Nullpunkt verschobene und linear reskalierte*" Version:

$$\tilde{f} : U_0 - x_0 \to U, \quad \tilde{f}(x - x_0) = {df_{x_0}}^{-1}(f(x) - y_0) = L(f(x) - y_0).$$

Natürlich können wir f aus $\tilde{f}$ mittels

$$f(x) = df_{x_0}(\tilde{f}(x - x_0)) + y_0$$

für $x \in U_0$ rekonstruieren, und es gilt $\tilde{f}(0) = 0$. Da L linear und stetig ist liefert die Kettenregel: $\tilde{f}$ ist stetig differenzierbar der Ableitung

$$d\tilde{f}_{x-x_0} = L \circ df_x, \quad x \in U_0$$

Der schon bewiesene Spezialfall ist daher auf $\tilde{f}$ anwendbar und liefert uns, dass $\tilde{f}$ eine geeignete offene Umgebung $\tilde{U}_1$ von $0 \in U$ bijektiv auf eine offene Umgebung $\tilde{V}_1$ von $0 \in U$ abbildet, und dass die Inverse $\tilde{h} = (\tilde{f}|_{V_1})^{-1} : \tilde{U}_1 \to \tilde{V}_1$ in 0 die Ableitung $d\tilde{h}_0 = \mathrm{id}_U$ besitzt. Zusätzlich war $\tilde{V}_1$ so gewählt, dass $d\tilde{f}_{x-x_0} : U \to U$ für alle $x - x_0 \in \tilde{U}_1$ invertierbar ist. Da $df_{x_0} : U \to V$ ein *Homöomorphismus* ist, also eine stetige Bijektion mit stetiger Inversen, bildet df_{x_0} offene Mengen bijektiv auf offene Mengen ab. Zusammengesetzt folgt, dass auch f die offene Umgebung $U_1 := x_0 + \tilde{U}_1$ von $x_0 \in U$ bijektiv auf die offene Umgebung $V_1 := y_0 + df_{x_0}[\tilde{V}_1]$ von $y_0 \in V$ abbildet. Die Inverse wird durch

$$h := (f|_{V_1})^{-1} : V_1 \to U_1, \quad h(y) = \tilde{h}(L(y - y_0)) + x_0$$

gegeben.
Insbesondere gilt nach der Kettenregel

$$dh_{y_0} = d\tilde{h}_0 \circ L = \mathrm{id}_U \circ {df_{x_0}}^{-1} = {df_{x_0}}^{-1}.$$

Wir wissen auch $df_x = df_{x_0} \circ d\tilde{f}_{x-x_0} \in \mathrm{GL}(U, V)$ für alle $x \in U_1$. Damit ist Teil 1 des Lemmas (und ein wenig von Teil 2) bewiesen.

Bemerkung 2.64 Die Fixpunktversion(45) für die Gleichung $\tilde{y} = \tilde{f}(\tilde{x})$, also

$$\tilde{x} = \tilde{x} - \tilde{f}(\tilde{x}) + \tilde{y},$$

lässt sich mit der eben betrachteten "Verschiebung und Reskalierung" $\tilde{x} = x - x_0$, $\tilde{y} = {df_{x_0}}^{-1}(y - y_0)$, $\tilde{f}(\tilde{x}) = {df_{x_0}}^{-1}(f(x) - y_0)$ auch in der Form

$$x = x - {df_{x_0}}^{-1}(f(x) - y)$$

schreiben. So übersetzt liefert die Anwendung des Banachschen Fixpunktsatzes aus dem Beweis des lokalen Umkehrsatzes das Iterationsverfahren

$$\boxed{\begin{array}{c}\textbf{Vereinfachtes Newtonverfahren}\\[1em] x^{(n+1)} = x^{(n)} - {df_{x_0}}^{-1}(f(x^{(n)}) - y), \quad n \in \mathbb{N}_0\end{array}}$$

zur Näherungslösung der Gleichung $f(x) = y$ für ein gegebenes y genügend nahe bei $y_0 = f(x_0)$ und Startwerte $x^{(0)}$ genügend nahe bei x_0, z.B. $x^{(0)} = x_0$.
Eine Variante davon, das *Newtonverfahren*, erhält man durch Verwendung der Ableitung $df_{x^{(n)}}$ statt df_{x_0}:

$$\boxed{\begin{array}{c}\textbf{Newtonverfahren}\\[1em] x^{(n+1)} = x^{(n)} - {df_{x^{(n)}}}^{-1}(f(x^{(n)}) - y), \quad n \in \mathbb{N}_0\end{array}}$$

Den eindimensionalen Spezialfall des Newtonverfahrens kennen Sie schon aus der Analysis 1. Das vereinfachte Newtonverfahren hat Vorteil, dass die Ableitung df_{x_0} immer an der gleichen Stelle x_0 verwendet werden kann, so dass (für $U = \mathbb{R}^n$) nur eine einzige Matrix invertiert werden muss bzw. in jedem Schritt ein lineares Gleichungssystem mit immer der gleichen Matrix $Df(x_0)$ gelöst werden muss. Für unsere Zwecke im Beweis hat es auch den Vorteil, dass wir keine zweite Ableitung für die Analyse brauchten, da x_0 festgehalten wird. Dagegen hat das Newtonverfahren den Vorteil einer viel höheren Konvergenzgeschwindigkeit. Genaueres darüber können Sie in der Numerischen Mathematik lernen.

Übung 2.65 Lösen Sie das Gleichungssystem

$$\begin{aligned} r\cos\phi &= x \\ r\sin\phi &= y \end{aligned}$$

für die rechte Seite $(x, y) = (1.1,\ 0.1)$ näherungsweise numerisch mit dem Taschenrechner (oder, noch besser, mit einem Computer)

(a) mit dem vereinfachten Newtonverfahren zum Start- und Referenzpunkt $(r_0, \phi_0) = (1, 0)$,

(b) mit dem Newtonverfahren zum gleichen Startpunkt.

Führen Sie jeweils einige Iterationen, z.B. drei, aus. Beobachten Sie, wie schnell sich die berechneten Näherungen an die exakte Lösung $(r, \phi) = (\sqrt{x^2 + y^2}, \arctan(y/x))$ annähern, aber auch, welcher Rechenaufwand jeweils nötig ist.

Beweis von Satz 2.63, Teil 2. Wir fixieren die Bijektion $h : V_1 \to U_1$ aus dem Teil 1 des Beweises, lassen aber nun $y \in V_1$ variieren. Insbesondere wissen wir $dh_{y_0} = {df_{x_0}}^{-1}$. Weil $df_x \in \mathrm{GL}(U, V)$ für alle $x \in U_1$, können wir ebensogut ein beliebiges $x \in U_1$ statt x_0 und $y = f(x) \in V_1$ statt $y_0 = f(x_0)$ nehmen und erhalten die Differenzierbarkeit von h

bei y und $dh_y = {df_x}^{-1}$. Die Stetigkeit der Ableitung $dh : V_1 \to \mathcal{B}(V,U)$ folgt nun durch Komposition der stetigen Abbildungen $h : V_1 \to U_1$, von $df : U_1 \to \mathrm{GL}(U,V)$ und der Inversenbildung $\cdot^{-1} : \mathrm{GL}(U,V) \to \mathrm{GL}(V,U)$; siehe Korollar 2.60.

Beweis von Satz 2.63, Teil 3. Wir zeigen induktiv, dass h sogar p-fach stetig differenzierbar ist, wenn f p-fach stetig differenzierbar ist. Der Beweis ist etwas abstrakt. Der Fall $p = 1$ ist mit Teil 2 des Satzes schon gezeigt. Gegeben $p \in \mathbb{N}$, nehmen wir als Induktionsvoraussetzung an, dass die Behauptung 3 des Satzes für dieses p und *alle* Banachräume U und V und *alle* Abbildungen $f : U_0 \to V$, die die Voraussetzungen des lokalen Umkehrsatzes erfüllen, gelte, nicht nur für ein fixiertes $f : U_0 \to V$.
Gegeben ein $(p+1)$-mal differenzierbares $f : U_0 \to V$ wie in der Voraussetzung des Satzes, erhalten wir aus dem schon Gezeigten nahe bei einer Stelle eine lokale Bijektion $f|_{U_1} : U_1 \to V_1$ und deren Inverse $h : V_1 \to U_1$. Für $y \in V_1$ betrachten wir das Gleichungssystem

$$\begin{aligned} f(x) &= y \\ df_x \circ X &= \mathrm{id}_V \end{aligned} \tag{54}$$

für ein gesuchtes $x \in U_1$ und ein gesuchtes $X \in \mathrm{GL}(V,U)$. Natürlich kennen wir die eindeutige Lösung $x = h(y)$ und $X = dh_y = {df_x}^{-1}$ dieses Gleichungssystems.
Wir können aber das Gleichungssystem auch so auffassen: Betrachten wir die Abbildung

$$F : U_1 \times \mathrm{GL}(V,U) \to V_1 \times \mathrm{GL}(V,V), \quad F(x,X) = (f(x), df_x \circ X).$$

wobei wir

$$U_1 \times \mathrm{GL}(V,U) \subseteq U \times \mathcal{B}(V,U)$$

mit der Norm $\|(x,X)\| = \max\{\|x\|, \|X\|_{V\to U}\}$ und

$$V_1 \times \mathrm{GL}(V,V) \subseteq V \times \mathcal{B}(V,V)$$

mit der Norm $\|(y,Y)\| = \max\{\|y\|, \|Y\|_{V\to V}\}$ versehen. Nun ist f eine $(p+1)$-mal differenzierbare Abbildung, also df eine p-mal differenzierbare Abbildung, also F eine p-mal differenzierbare Abbildung.[19] Weiter gilt $F(x, dh_x) = (y, \mathrm{id}_V)$. Mit der Kettenregel und Beispiel 2.22 erhalten wir die Ableitung (in Matrixnotation)

$$dF_{(x,dh_x)} \begin{pmatrix} x' \\ X' \end{pmatrix} = \begin{pmatrix} df_x(x') \\ d^2 f_x(x') \circ dh_x + df_x \circ X' \end{pmatrix} = \begin{pmatrix} df_x & 0 \\ d^2 f_x(\cdot) \circ dh_x & df_x \circ \cdot \end{pmatrix} \begin{pmatrix} x' \\ X' \end{pmatrix}.$$

Insbesondere ist $dF_{(x,dh_x)}$ invertierbar mit einer stetigen Inversen, da $d^2 f_x(\cdot) \circ dh_x$ stetig ist und da df_x und $df \circ \cdot$ invertierbar mit stetigen Inversen sind. Der lokale Umkehrsatz Teil 1 und 2 und die Induktionsvoraussetzung sind daher auf F anwendbar. Wir erhalten die p-fache Differenzierbarkeit der Lösung $y \mapsto X = dh_y$ des Gleichungssystems (54), also die behauptete $(p+1)$-fache Differenzierbarkeit von h.

□

Übung 2.66 (Quadratwurzel von Matrizen und Operatoren) *Bearbeiten Sie entweder Teil (a) oder Teil (b) der folgenden Aufgabe:*

(a) Zeigen Sie, dass die Abbildung

$$Q : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}, \quad Q(A) = A^2$$

[19] Hier verwenden wir, dass die Kompositionsabbildung glatt ist, vgl. Bsp. 2.22.

stetig differenzierbar mit der Ableitung $dQ_A(B) = A \cdot B + B \cdot A$ für $A, B \in \mathbb{R}^{n\times n}$ ist. Folgern Sie, dass es eine stetig differenzierbare, in einer offenen Umgebung V_1 von $\mathrm{Id} \in \mathbb{R}^{n\times n}$ definierte Abbildung $\mathrm{sqrt} : V_1 \to \mathbb{R}^{n\times n}$, $A \mapsto \sqrt{A}$, mit $\sqrt{\mathrm{Id}} = \mathrm{Id}$ gibt, die Q genügend nahe bei Id invertiert. Zeigen Sie $d\,\mathrm{sqrt}_{\mathrm{Id}}(A) = \frac{1}{2}A$ für $A \in \mathbb{R}^{n\times n}$.

(b) Es sei $(U, \|\cdot\|)$ ein Banachraum. Zeigen Sie, dass die Abbildung

$$Q : (\mathcal{B}(U,U), \|\cdot\|_{U\to U}) \to (\mathcal{B}(U,U), \|\cdot\|_{U\to U}), \quad Q(A) = A \circ A$$

stetig differenzierbar mit der Ableitung $dQ_A(B) = A \circ B + B \circ A$ für $A, B \in \mathcal{B}(U,U)$ ist. Folgern Sie, dass es eine stetig differenzierbare, in einer offenen Umgebung V_1 von $\mathrm{id}_U \in \mathcal{B}(U,U)$ definierte Abbildung $\mathrm{sqrt} : V_1 \to \mathcal{B}(U,U)$, $A \mapsto \sqrt{A}$, mit $\sqrt{\mathrm{id}_U} = \mathrm{id}_U$ gibt, die Q genügend nahe bei id_U invertiert. Zeigen Sie $d\,\mathrm{sqrt}_{\mathrm{id}_U}(A) = \frac{1}{2}A$ für $A \in \mathcal{B}(U,U)$.

Übung 2.67 (Matrix-Logarithmus) Zeigen Sie, dass es eine offene Umgebung U von $0 \in \mathbb{R}^{n\times n}$ gibt, so dass die Matrix-Exponentialfunktion $\exp|_U : U \to V := \exp[U] \subseteq \mathbb{R}^{n\times n}$ ein Homöomorphismus mit einer stetig differenzierbaren Umkehrung $\log : V \to U$ ist. Berechnen Sie $d\log_{\mathrm{Id}}$.

Nun besprechen wir eine globale Variante des lokalen Umkehrsatzes:

Definition 2.68 (Diffeomorphismen) *Es seien $V, U \subseteq \mathbb{R}^n$ offen und $p \in \mathbb{N} \cup \{\infty\}$. Eine Abbildung $f : U \to V$ wird C^p-Diffeomorphismus genannt, wenn sie bijektiv ist und f und f^{-1} p-fach stetig differenzierbar sind.*

Diffeomorphismen sind also sozusagen die "Isomorphismen der Differentialrechnung". Wir erhalten:

Korollar 2.69 (Offenheitssatz und globaler Umkehrsatz) *Es sei $U \subseteq \mathbb{R}^n$ offen und $f \in C^p(U, \mathbb{R}^n)$. Für alle $x \in U$ sei $df_x : \mathbb{R}^n \to \mathbb{R}^n$ invertierbar. Dann gilt:*

1. *Für jedes offene $\tilde{U} \subseteq U$ ist $f[\tilde{U}]$ offen.*[20]

2. *Ist zusätzlich $f : U \to V$ bijektiv, so ist f ein C^p-Diffeomorphismus.*

Beweis:

1. Gegeben $\tilde{U}$ wie in der Voraussetzung, gibt es für jedes $x \in \tilde{U}$ nach dem lokalen Umkehrsatz offene Umgebungen $U_1 \subseteq \tilde{U}$ von x und $V_1 \subseteq f[\tilde{U}]$ von $f(x)$, so dass $f|_{U_1} : U_1 \to V_1$ bijektiv ist. Also ist $f[\tilde{U}]$ offen.

2. $f^{-1} : V \to U$ ist stetig, da $(f^{-1})^{-1}[\tilde{U}] = f[\tilde{U}]$ für alle offenen $\tilde{U} \subseteq U$ offen ist. Zudem ist nach dem lokalen Umkehrsatz f^{-1} p-fach stetig differenzierbar. Es folgt die Behauptung.

[20] Hierfür verwendet man auch die Sprechweise: "f ist offen".

□

Übung 2.70 (Ein hinreichendes Kriterium für Diffeomorphie) Es sei $U \subseteq \mathbb{R}^n$ offen und konvex, $p \in \mathbb{N} \cup \{\infty\}$ und $f \in C^p(U, \mathbb{R}^n)$. Es gelte $\langle y, df_x(y) \rangle > 0$ (mit dem euklidischen Skalarprodukt) für alle $x \in U$ und $y \in \mathbb{R}^n$. Zeigen Sie, dass $f : U \to f[U]$ ein C^p-Diffeomorphismus ist.

Beispiel 2.71 (Singuläre elliptische Kurve mit Selbstschnitt) Das folgende Gegenbeispiel zeigt, dass die Aussage des Offenheitssatzes sich nicht auf Abbildungen $f : \mathbb{R}^m \to \mathbb{R}^n$ mit $m < n$ übertragen lässt. Es sei

$$f : \mathbb{R} \to \mathbb{R}^2, \quad f(t) = (t^2 - 1, t^3 - t)) = ((t+1)(t-1), t(t+1)(t-1))$$

Die Abbildung f parametrisiert eine Kurve in der Ebene $\mathbb{R}^2$, nämlich den Lösungsraum der Gleichung

$$x^2(x+1) = y^2.$$

Allerdings hat diese Kurve einen "Selbstschnitt", nämlich bei $f(-1) = f(1) = (0,0)$. Wir betrachten $g := f|_{]-1,\infty[}$. Es gilt zwar: $df_t : \mathbb{R} \to \mathbb{R}^2$ ist für alle $t \in \mathbb{R}$ injektiv, d.h. die Kurve wird überall mit positiver Geschwindigkeit durchlaufen, und $g :]-1,\infty[\to g[]-1,\infty[]$ ist bijektiv, aber dennoch ist die Inverse $g^{-1} : g[]-1,\infty[] \to]-1,\infty[$ unstetig in $(0,0)$, denn das Urbild $g[]-1/2, 1/2[]$ der offenen Menge $]-1/2, 1/2[$ unter g^{-1} ist *keine* Umgebung von $(0,0)$ in $g[]-1,\infty[]$.

2.10 Implizit definierte Funktionen

Betrachten wir ein System von n linearen Gleichungen

$$\begin{aligned}
a_{11}x_1 + \ldots + a_{1m}x_m + b_{11}y_1 + \ldots + b_{1n}y_n &= c_1, \\
a_{21}x_1 + \ldots + a_{2m}x_m + b_{21}y_1 + \ldots + b_{2n}y_n &= c_2, \\
\vdots \qquad\qquad\qquad\qquad \vdots \qquad & \quad \vdots \\
a_{n1}x_1 + \ldots + a_{nm}x_m + b_{n1}y_1 + \ldots + b_{nn}y_n &= c_n
\end{aligned} \tag{55}$$

für n Unbekannte $y_1, \ldots, y_n$, wobei wir uns die Konstanten a_{ij} und b_{ik} sowie die rechten Seiten c_i als fest gegeben vorstellen. Auch die Variablen $x_1, \ldots, x_m$ stellen wir uns als gegeben, aber variabel vor. Unter welchen Umständen kann man dieses Gleichungssystem nach den $y_1, \ldots, y_n$ auflösen, also eindeutige affin-lineare Funktionen

$$\begin{aligned}
y_1 &= g_1(x_1, \ldots, x_m) = g_{11}x_1 + \ldots + g_{1m}x_m + d_1, \\
y_2 &= g_2(x_1, \ldots, x_m) = g_{21}x_1 + \ldots + g_{2m}x_m + d_2, \\
\vdots \quad & \qquad \vdots \qquad\qquad \vdots \\
y_n &= g_n(x_1, \ldots, x_m) = g_{n1}x_1 + \ldots + g_{nm}x_m + d_n
\end{aligned} \tag{56}$$

angeben, die das aufgelöste lineare Gleichungsystem beschreiben?

Solche Funktionen $g_1, \dots g_n : \mathbb{R}^m \to \mathbb{R}$, falls sie existieren und eindeutig sind, nennen wir die durch das Gleichungssystem (55) *implizit definierten Funktionen* oder kurz *implizite Funktionen.*

Schreiben wir das Gleichungssystem (55) in Matrixform. Setzen wir dazu $A = (a_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,m}} \in \mathbb{R}^{n\times m}$, $B = (b_{ik})_{i,k=1,\dots,n} \in \mathbb{R}^{n\times n}$, $c = (c_1, \dots, c_n)^t \in \mathbb{R}^n$, $y = (y_1, \dots, y_n)^t \in \mathbb{R}^n$, $x = (x_1, \dots, x_n)^t \in \mathbb{R}^m$ und[21]

$$f : \mathbb{R}^n \oplus \mathbb{R}^m \to \mathbb{R}^n, \quad f(x, y) = Ax + By$$

so können wir das Gleichungssystem (55) auch in der Form $f(x, y) = c$ schreiben. Wir suchen also eine implizit definierte Funktion $g = (g_1, \dots, g_n)^t : \mathbb{R}^m \to \mathbb{R}^n$ mit

$$f(x, g(x)) = c.$$

Die lineare Algebra gibt uns eine einfache Antwort auf obige Frage: Eine eindeutige implizit definierte Funktion $g : \mathbb{R}^m \to \mathbb{R}^n$ existiert genau dann, wenn die Matrix B invertierbar ist. In diesem Fall beschreibt

$$y = g(x) = B^{-1} \cdot (c - Ax) = B^{-1}c - B^{-1}Ax \tag{57}$$

diese eindeutige implizit definierte Funktion.
Wie verhält es sich nun mit einem nichtlinearen Gleichungssystem $f(x, y) = c$? Unter welchen Voraussetzungen kann man ein nichtlineares Gleichungssystem $f(x, y) = c$ "nahe bei einem Referenzpunkt" (x_0, y_0) mit $f(x_0, y_0) = c$ eindeutig durch implizit definierte Funktionen $y = g(x)$ nach y auflösen?
Die Leitidee hinter der Differentialrechnung, dass eine differenzierbare Funktion f "nahe bei einer Stelle" (x_0, y_0) sehr ähnliche Eigenschaften wie ihre Linearisierung $(x, y) \mapsto df_{(x_0,y_0)}(x - x_0, y - y_0) + f(x_0, y_0)$ haben sollte, gibt auch hier die richtige Antwort. Folgende Notation ist zweckmäßig: $d_1f_{(x,y)}$ soll die Ableitung von $f(\cdot, y) : x \mapsto f(x, y)$ bei festgehaltenem y beschreiben, während $d_2f_{(x,y)}$ die Ableitung von $f(x, \cdot) : y \mapsto f(x, y)$ bei festgehaltenem x bedeuten soll. Es gilt also

$$df_{(x_0,y_0)}(x - x_0, y - y_0) = d_1f_{(x_0,y_0)}(x - x_0) + d_2f_{(x_0,y_0)}(y - y_0).$$

$D_1f(x, y)$ und $D_2f(x, y)$ bezeichnen die zugehörigen Jacobimatrizen. $A := D_1f(x_0, y_0)$ entspricht also linearisiert der Matrix A von oben, während $B := D_2f(x_0, y_0)$ der Matrix B von oben entspricht.

Satz 2.72 (Satz von den impliziten Funktionen in $\mathbb{R}^n$) *Es seien $m, n \in \mathbb{N}$, $\mathcal{O} \subseteq \mathbb{R}^m \oplus \mathbb{R}^n$ eine offene Menge, $(x_0, y_0) \in \mathcal{O}$ und $f : \mathcal{O} \to \mathbb{R}^n$ stetig differenzierbar. Wir setzen $c = f(x_0, y_0)$. Weiter sei die Matrix $D_2f_{(x_0,y_0)} \in \mathbb{R}^{n\times n}$ invertierbar. Dann gibt es eine offene Umgebung U_1 von x_0 in $\mathbb{R}^m$ und eine offene Umgebung $\mathcal{O}_1 \subseteq \mathcal{O}$ von (x_0, y_0),*

[21]Die *direkte Summe* $\mathbb{R}^n \oplus \mathbb{R}^m$ ist einfach $\mathbb{R}^n \times \mathbb{R}^m$ mit komponentenweisen Vektoroperationen.

für die Folgendes gilt: Es gibt eine eindeutig bestimmte Funktion $g : U_1 \to \mathbb{R}^n$, so dass für alle $x \in U_1$ gilt: $(x, g(x)) \in \mathcal{O}_1$ und

$$f(x, g(x)) = c. \tag{58}$$

Weiter ist g stetig differenzierbar mit der Ableitung

$$dg_x = -D_2 f(x, g(x))^{-1} \cdot D_1 f(x, g(x)). \tag{59}$$

Ist die Funktion f sogar p-fach stetig differenzierbar, so ist auch die Funktion g p-fach stetig differenzierbar.

Die Funktion g wird die durch das Gleichungssystem $f(x, g(x)) = c$ *implizit definerte Funktion* (kurz auch: "*implizite Funktion*") bei (x_0, y_0) genannt.
Insbesondere wird garantiert, dass $D_2 f_{(x,g(x))} \in \mathbb{R}^{n \times n}$ für alle $x \in U_1$ invertierbar ist. Man beachte die Ähnlichkeit der Formeln (59) und (57): Beide enthalten

$$Dg(x_0) = -B^{-1} A.$$

Der Satz lässt sich wie folgt auf Banachräume verallgemeinern. Die Formulierung des Spezialfalls in $\mathbb{R}^n$ und der Verallgemeinerung auf Banachräume sind sehr ähnlich:

Satz 2.73 (Satz von den impliziten Funktionen in Banachräumen) *Es seien $(U, \|\cdot\|)$, $(V, \|\cdot\|)$ und $(W, \|\cdot\|)$ drei Banachräume über $\mathbb{R}$. Die direkte Summe $U \oplus V$ sei mit der Norm $\|(u,v)\| := \max\{\|u\|, \|v\|\}$ versehen. Es seien $\mathcal{O} \subseteq U \oplus V$ eine offene Menge, $(x_0, y_0) \in \mathcal{O}$ und $f : \mathcal{O} \to W$ stetig differenzierbar. Wir setzen $c = f(x_0, y_0)$. Weiter sei die lineare Abbildung $d_2 f_{(x_0,y_0)} : V \to W$ invertierbar mit einer stetigen Inversen. Dann gibt es eine offene Umgebung U_1 von x_0 in U und eine offene Umgebung $\mathcal{O}_1 \subseteq \mathcal{O}$ von (x_0, y_0), für die Folgendes gilt: Es gibt eine eindeutig bestimmte $g : U_1 \to V$, so dass für alle $x \in U_1$ gilt: $(x, g(x)) \in \mathcal{O}_1$ und*

$$f(x, g(x)) = c. \tag{60}$$

Weiter ist g stetig differenzierbar mit der Ableitung

$$dg_x = -(d_2 f_{(x,g(x))})^{-1} \circ d_1 f_{(x,g(x))}. \tag{61}$$

Ist die Funktion f sogar p-fach stetig differenzierbar, so ist auch die Funktion g p-fach stetig differenzierbar.

Die Bedingung $d_2 f_{(x_0,y_0)} \in \mathrm{GL}(V, W)$ spielt hier die Rolle einer *Regularitätsbedingung*, die singuläre Fälle ausschließt.
Beweis der Banachraumversion. Die Idee des Beweises besteht in einer "Erweiterung" der Funktion f zu

$$F : \mathcal{O} \to U \oplus W, \quad F(x, y) = (x, f(x, y)).$$

Dabei wird der Raum $U \oplus W$ analog zu $U \oplus V$ mit der Norm $\|(u, w)\| = \max\{\|u\|, \|w\|\}$ versehen. Man beachte, dass $U \oplus W$ und $U \oplus V$ mit ihren Normen wieder Banachräume sind. Die Funktion F ist, ebenso wie f, stetig differenzierbar mit der Ableitung (in Matrixform geschrieben)

$$dF_{(x,y)} = \begin{pmatrix} \mathrm{id}_U & 0 \\ d_1 f_{(x,y)} & d_2 f_{(x,y)} \end{pmatrix} : U \oplus V \to U \oplus W.$$

Zudem ist F p-fach stetig differenzierbar, falls auch f p-fach stetig differenzierbar ist. An allen Stellen $(x, y) \in \mathcal{O}$, an denen $d_2 f_{(x,y)} : V \to W$ eine stetige lineare Inverse besitzt, besitzt auch $dF_{(x,y)}$ eine stetige lineare Inverse, nämlich

$$dF_{(x,y)}{}^{-1} = \begin{pmatrix} \mathrm{id}_U & 0 \\ -(d_2 f_{(x,y)})^{-1} \circ d_1 f_{(x,y)} & (d_2 f_{(x,y)})^{-1} \end{pmatrix} : U \oplus W \to U \oplus V. \tag{62}$$

Nach Voraussetzung ist das insbesondere für $(x, y) = (x_0, y_0)$ der Fall, also auch in einer offenen Umgebung von (x_0, y_0), da $\mathrm{GL}(U \oplus V, U \oplus W)$ nach Korollar 2.60 offen ist. Nach dem lokalen Umkehrsatz gibt es eine offene Umgebung $\mathcal{O}_1 \subseteq \mathcal{O}$ von (x_0, y_0) und eine offene Umgebung $\mathcal{O}'_1 \subseteq U \oplus W$ von $F(x_0, y_0) = (x_0, c)$, so dass $F|\mathcal{O}_1 : \mathcal{O}_1 \to \mathcal{O}'_1$ bijektiv mit einer stetig differenzierbaren Umkehrung $H := (F|\mathcal{O}_1)^{-1} : \mathcal{O}'_1 \to \mathcal{O}_1$ ist (p-fach stetig differenzierbar, falls auch F p-fach stetig differenzierbar ist). Zudem gilt

$$dH_{(x,z)} = dF_{(x,y)}{}^{-1}$$

für $(x, z) \in \mathcal{O}'_1$ und $(x, y) = H(x, z)$. Da $F(x, y) = (x, f(x, y))$ die erste Komponente x festlässt, muss auch H die erste Komponente festlassen:

$$H(x, z) = (x, G(x, z))$$

für alle $(x, z) \in \mathcal{O}'_1$ mit einer (p-fach) stetig differenzierbaren Funktion $G : \mathcal{O}'_1 \to V$. Insbesondere gilt $G(x_0, c) = y_0$ wegen $H(x_0, c) = (x_0, y_0)$, und (in Zeilenvektornotation)

$$dG(x, z) = (-(d_2 f_{(x,y)})^{-1} \circ d_1 f_{(x,y)}, (d_2 f_{(x,y)})^{-1}) \tag{63}$$

für $(x, y) = H(x, z)$ wegen der zweiten Zeile der Matrix in (62). Wir setzen

$$U_1 = \{x \in U | \ (x, c) \in \mathcal{O}'_1\}$$

Die Menge U_1 ist offen, da $\mathcal{O}'_1$ offen ist. Aufgrund der Bestimmungsgleichung $f(x, g(x)) = c$ für die unbekannte Funktion g, also $F(x, g(x)) = (x, c)$, also $H(x, c) = (x, g(x))$ bleibt uns nur die einzige Möglichkeit, $g : U_1 \to V$ zu definieren:

$$g(x) := G(x, c).$$

Insbesondere besitzt g die Ableitung

$$dg(x) = -(d_2 f_{(x,y)})^{-1} \circ d_1 f_{(x,y)}$$

für $y = g(x)$, wie man aus dem ersten Eintrag im Zeilenvektor (63) abliest. Die Funktion g leistet das Gewünschte.

$\square$

Beispiel 2.74 (Nord- und Südhalbkugel als Graph) Betrachten wir die Abbildung

$$f : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}, f(x_1, x_2, y) = x_1^2 + x_2^2 + y^2$$

und einen Punkt $(x_{0,1}, x_{0,2}, y_0)$ auf der Einheitsphäre S^2, also zum Niveau $f(x_{0,1}, x_{0,2}, y_0) = c = 1$. Die Regularitätsbedingung

$$\frac{\partial}{\partial y} f(x_1, x_2, y) = 2y \neq 0$$

gilt für $y \neq 0$. Sie ist also am Äquator verletzt. Diese Stellen hatten wir schon früher anschaulich-geometrisch als Beschreibungssingularitäten identifiziert. Ist die Regularitätsbedingung bei $(x_{0,1}, x_{0,2}, y_0)$ erfüllt, gilt also $x_{0,1}^2 + x_{0,2}^2 < 1$, erhalten wir aus dem Satz von den impliziten Funktionen eine in einer Umgebung U_1 von $(x_{0,1}, x_{0,2})$ definierte implizite Funktion $g : U_1 \to \mathbb{R}$ mit $f(x, y, g(x, y)) = c$ und

$$\begin{aligned} dg(x_1, x_2) &= -\left(\frac{\partial f}{\partial y}(x_1, x_2, y)\right)^{-1} \left(\frac{\partial f}{\partial x_1}(x_1, x_2, y)\, dx_1 + \frac{\partial f}{\partial x_1}(x_1, x_2, y)\, dx_2\right)\Bigg|_{y=g(x_1,x_2)} \\ &= -\frac{x_1\, dx_1 + x_2\, dx_2}{g(x_1, x_2)}. \end{aligned}$$

Natürlich kennen wir diese implizit definierte Funktion auch explizit:

$$g(x_1, x_2) = \pm\sqrt{1 - x_1^2 - x_2^2},$$

je nachdem, ob $y_0 > 0$ oder $y_0 < 0$ gilt.

Beispiel 2.75 (Ableitung der Matrixinvertierung) Es sei

$$\mathrm{GL}_n(\mathbb{R}) = \{X \in \mathbb{R}^{n\times n} |\ \det X \neq 0\}$$

und

$$\mathrm{inv} : \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^{n\times n}, \quad \mathrm{inv}(X) = X^{-1}.$$

Diese Abbildung wird auch implizit durch die Gleichung

$$X \cdot \mathrm{inv}(X) = \mathrm{Id}$$

definiert. Betrachten wir also die Abbildung

$$f : \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}, \quad f(X, Y) = XY.$$

Es gilt

$$df_{(X,Y)}(X', Y') = X'Y + XY',$$

also

$$d_1 f_{(X,Y)}(X') = X'Y,$$
$$d_2 f_{(X,Y)}(Y') = XY'.$$

Die Regularitätsbedingung besagt, dass $Y' \mapsto d_2 f_{X,Y}(Y') = XY'$ invertierbar sein soll. Sie ist genau dann erfüllt, wenn X invertierbar ist. Für die implizit definierte Funktion inv erhalten wir also für $X \in \mathrm{GL}_n(\mathbb{R})$, $Y = X^{-1}$ und $X' \in \mathbb{R}^{n\times n}$ die Ableitung

$$d\,\mathrm{inv}_X(X') = -(d_2 f_{(X,Y)})^{-1} \circ d_1 f_{(X,Y)}(X') = -X^{-1}(X'Y)$$

also

$$\boxed{d\,\mathrm{inv}_X(X') = -X^{-1}X'X^{-1}} \tag{64}$$

Eine Variante dieser Formel ist uns aus Übung 2.62 bekannt. Die Formel (64) verallgemeinert die eindimensionale Ableitungsregel $d(x^{-1}) = -x^{-2}\,dx$.

Übung 2.76 Zeigen Sie mit dem Banachschen Fixpunktsatz und dem Satz von den impliziten Funktionen, dass durch das Gleichungssystem

$$g_1(x_1,x_2) = \frac{1}{10}\arctan(g_1(x_1,x_2) + g_2(x_1,x_2) + x_1)$$
$$g_2(x_1,x_2) = \frac{1}{20}\arctan(g_1(x_1,x_2) - g_2(x_1,x_2) + x_2)$$

eine eindeutige glatte Funktion $g = (g_1, g_2) : \mathbb{R}^2 \to \mathbb{R}^2$ *global* implizit definiert wird. Berechnen Sie $dg_1(0,0)$ und $dg_2(0,0)$.

2.11 Untermannigfaltigkeiten

Wir definieren nun singularitätenfreie m-dimensionale Gebilde im n-dimensionalen Raum, wobei $m \le n$.

Definition 2.77 (Untermannigfaltigkeiten) *Es seien U ein n-dimensionaler $\mathbb{R}$-Vektorraum, $n \in \mathbb{N}$, und $m \in \mathbb{N}_0$ mit $0 \le m \le n$ sowie $p \in \mathbb{N} \cup \{\infty\}$. Eine Teilmenge $M \subseteq U$ heißt m-dimensionale C^p-Untermannigfaltigkeit von U, wenn es für alle $x \in M$ eine offene*[22] *Umgebung U_1 von x und eine Abbildung $f \in C^p(U_1, V)$ in einen $(n-m)$-dimensionalen $\mathbb{R}$-Vektorraum V gibt, so dass*

$$M \cap U_1 = f^{-1}[f(x)]$$

gilt und $df_x : U \to V$ surjektiv ist. C^∞-Untermannigfaltigkeiten heißen auch glatte Untermannigfaltigkeiten. Wir nennen $n - m = \dim_{\mathbb{R}} V$ auch die Kodimension der Untermannigfaltigkeit M in U.[23]

[22] U sei mit der Standardtopologie und mit einer Norm versehen. Auf die Wahl der Norm kommt es nicht an, da alle Normen auf U äquivalent sind.

[23] Mit Ausnahme des Trivialfalls $M = \emptyset$ sind Dimension und Kodimension einer Untermannigfaltigkeit M wohldefiniert, also schon durch die Menge M und den Vektorraum U alleine bestimmt, unabhängig von der Wahl der Funktionen f.

Wenn wir einfach von "Untermannigfaltigkeiten" (ohne Zusatz) reden, sind C^1-Untermannigfaltigkeiten gemeint. Im Hinblick auf Anwendungen, z.B. in Matrizenräumen, haben wir in der Definition einen beliebigen endlichdimensionalen Vektorraum U statt nur $U = \mathbb{R}^n$ zugelassen. Wir hätten in der Definition auch einfach äquivalent $V = \mathbb{R}^{n-m}$ fordern können, doch das ist nicht immer praktisch.[24]
Die *Regularitätsbedingung* $df_x[U] = V$ kann man wegen $\dim_{\mathbb{R}} V = n - m$ auch in der Form $\operatorname{Rang} df_x = n - m$ schreiben. Sie dient zum Ausschluss von Singularitäten. Für $U = \mathbb{R}^n$, $V = \mathbb{R}^{m-n}$ bedeutet sie, dass die Jacobimatrix $Df(x)$ den Rang $n - m$ besitzt.
Zum Beispiel sind 1-dimensionale Untermannigfaltigkeiten von $\mathbb{R}^2$ bzw. $\mathbb{R}^3$ singularitätenfreie Kurven in der Ebene bzw. im Raum, und 2-dimensionale Untermannigfaltigkeiten von $\mathbb{R}^3$ sind Flächen im Raum. Es gibt noch zwei Trivialfälle: n-dimensionale Untermannigfaltigkeiten von $\mathbb{R}^n$ sind nichts anderes als offene Mengen, und 0-dimensionale Untermannigfaltigkeiten sind einfach Mengen von isolierten Punkten.

Beispiel 2.78 (Singuläre elliptische Kurven)

1. Das Lösungsgebilde M der Gleichung
$$x^3 - y^2 = 0$$
in der Ebene $\mathbb{R}^2$ ist eine Kurve mit einer Spitze im Nullpunkt $(0,0)$. Es ist wegen dieser Singularität *keine* Untermannigfaltigkeit, und in der Tat ist die Ableitung
$$d(x^3 - y^2) = 3x^2\,dx - 2y\,dy : \mathbb{R}^2 \to \mathbb{R}$$
genau an der Stelle $x = y = 0$ *nicht* surjektiv, also die Regularitätsbedingung dort (und nur dort) verletzt. Entfernen wir jedoch diesen singulären Punkt, so wird $M \setminus \{(0,0)\}$ eine 1-dimensionale Untermannigfaltigkeit der Ebene.

2. Das Lösungsgebilde M der Gleichung
$$x^2(x+1) - y^2 = 0$$
in der Ebene $\mathbb{R}^2$ ist zwar eine glatte Kurve, doch auch sie ist *keine* Untermannigfaltigkeit, da sie im Nullpunkt ebenfalls eine Singularität besitzt, und zwar einen *Selbstschnitt*, vgl. Beispiel 2.71. In der Tat ist die Ableitung
$$d(x^2(x+1) - y^2) = (3x^2 + 2x)\,dx - 2y\,dy : \mathbb{R}^2 \to \mathbb{R}$$
für $(x,y) = (0,0)$ *nicht* surjektiv, aber in allen anderen Punkten $(x,y) \in M \setminus \{(0,0)\}$ ist sie surjektiv, so dass auch die Kurve $M \setminus \{(0,0)\}$, die durch Herausnehmen des Selbstschnitts entsteht, eine 1-dimensionale Untermannigfaltigkeit der Ebene bildet.

[24] Einen allgemeineren Begriff "*Banachuntermannigfaltigkeiten*" bekommt man, wenn man statt U und V beliebige Banachräume zulässt. Diesen allgemeineren Begriff besprechen wir in dieser Vorlesung nicht, obwohl die allgemeinere Theorie sehr ähnlich zum hier besprochenen endlichdimensionalen Fall ist.

Beispiel 2.79 (Sphären) Für $n \in \mathbb{N}$ und $r > 0$ ist das Niveaugebilde

$$rS^n := \{x \in \mathbb{R}^n |\ \|x\|_2^2 = r^2\}$$

des euklidischen Normquadrats eine $(n-1)$-dimensionale C^∞-Untermannigfaltigkeit von $\mathbb{R}^n$, synonym: eine glatte *Hyperfläche*. ("*Hyper-*" meint hier "1-kodimensional", $n - m = 1$.) In der Tat ist

$$d(\|x\|^2) = 2\langle x, \cdot\rangle : \mathbb{R}^n \to \mathbb{R}$$

für $x \neq 0$ surjektiv, für $x = 0$ jedoch nicht. Geometrisch passt das damit zusammen, dass das Niveau-0-Gebilde des Normquadrats kein $(n-1)$-dimensionales Gebilde mehr ist, sondern nur mehr aus dem Nullpunkt besteht.

Beispiel 2.80 Der Lösungsraum M in $\mathbb{R}^3$ des Gleichungssystems

$$\begin{aligned} &x^2 + y^2 + z^2 = 3, \\ &x + y + z = 0 \end{aligned}$$

also des Schnitts einer Kugeloberfläche mit einer Ebene durch den Nullpunkt, ist eine eindimensionale Untermannigfaltigkeit von $\mathbb{R}^3$, nämlich eine Kreislinie. Sie hat die Kodimension $2 = 3 - 1$. Die Regularitätsbedingung für die Jacobimatrix

$$\operatorname{Rang}\begin{pmatrix} 2x & 2y & 2z \\ 1 & 1 & 1 \end{pmatrix} = 2$$

ist in allen Punkten $(x, y, z) \in M$ in der Tat erfüllt.
Dagegen besteht der Lösungsraum $\{P\}$ in $\mathbb{R}^3$ des Gleichungssystems

$$\begin{aligned} &x^2 + y^2 + z^2 = 3, \\ &x + y + z = 3 \end{aligned}$$

nicht aus einer eindimensionalen Linie, sondern nur nur aus dem Punkt $P = (1, 1, 1)$, denn es handelt sich um den Schnitt einer Kugeloberfläche mit ihrer Tangentialebene in P. In der Tat ist die Regularitätsbedingung in P verletzt, denn die Jacobimatrix zu der linken Seite des Gleichungssystems hat in P hat nicht den Rang 2, sondern nur den Rang 1:

$$\operatorname{Rang}\begin{pmatrix} 2x & 2y & 2z \\ 1 & 1 & 1 \end{pmatrix}\Bigg|_{(x,y,z)=(1,1,1)} = 1.$$

Beispiel 2.81 (Die orthogonale Gruppe als Untermannigfaltigkeit) In der linearen Algebra wird für $n \in \mathbb{N}$ die "*orthogonale Gruppe*"

$$\mathrm{O}(n) = \{U \in \mathbb{R}^{n\times n} |\ U^t U = \mathrm{Id}\}$$

definiert. Ihre Elemente heißen "*orthogonale Matrizen*; sie beschreiben genau die isometrischen linearen Abbildungen (z.B. Drehungen und Spiegelungen):

$$\mathrm{O}(n) = \{U \in \mathbb{R}^{n\times n} | \forall x \in \mathbb{R}^n :\ \|Ux\|_2 = \|x\|_2\}$$

Wir zeigen nun, dass $\mathrm{O}(n)$ eine glatte Untermannigfaltigkeit von $\mathbb{R}^{n\times n}$ der Dimension $n(n-1)/2$ ist. Hierzu sei $V = \{A \in \mathbb{R}^{n\times n} | A = A^t\}$ der Raum der symmetrischen reellen $n \times n$-Matrizen. Dieser Raum ist $n(n+1)/2$-dimensional. (Übung zur linearen Algebra: Geben Sie eine Basis von V an.) Gegeben $X_0 \in O(n)$, betrachten wir $f : \mathbb{R}^{n\times n} \to V$, $f(X) = X_0 X^t X X_0^t$. Man beachte, dass in der Tat $f(X)^t = f(X)$, also $f(X) \in V$ gilt. Da X_0 invertierbar ist, folgt

$$\mathrm{O}(n) = f^{-1}[\{\mathrm{Id}\}] = f^{-1}[\{f(X_0)\}]$$

denn für $X \in \mathbb{R}^{n\times n}$ gilt

$$f(X) = \mathrm{Id} \Leftrightarrow X_0^t X_0 X^t X X_0^t X_0 = X_0^t \,\mathrm{Id}\, X_0 \Leftrightarrow X^t X = \mathrm{Id} \Leftrightarrow X \in \mathrm{O}(n)$$

Offensichtlich ist f glatt. Wir berechnen die Ableitung von f in $X \in \mathbb{R}^{n\times n}$ an der Stelle $X' \in \mathbb{R}^{n\times n}$ mit der Produktregel:

$$df_X(X') = X_0 X^t X' X_0^t + X_0 X'^t X X_0^t$$

und damit speziell für $X = X_0 \in \mathrm{O}(n)$:

$$df_{X_0}(X') = X_0 X_0^t X' X_0^t + X_0 X'^t X_0 X_0^t = \mathrm{Id}\, X' X_0^t + X_0 X'^t \,\mathrm{Id} = X' X_0^t + X_0 X'^t = X' X_0^t + (X' X_0^t)^t$$

Insbesondere ist $df_{X_0} : \mathbb{R}^{n\times n} \to V$ surjektiv, denn für $Y \in V$ gilt

$$df_{X_0}(\frac{1}{2} Y X_0) = \frac{1}{2}(Y + Y^t) = Y.$$

Dieses Beispiel zeigt auch, dass es praktisch ist, in der Definition 2.77 mit endlichdimensionalen Vektorräumen statt nur mit $\mathbb{R}^n$ und $\mathbb{R}^{n-m}$ zu arbeiten.

Übung 2.82 (Die unitäre Gruppe und die Lorentzgruppe als Untermannigfaltigkeiten) Zeigen Sie analog zum Beispiel 2.81 Folgendes:

1. Die "*unitäre Gruppe*" zu $n \in \mathbb{N}$, definiert durch

   $$\mathrm{U}(n) := \{A \in \mathbb{C}^{n\times n} |\ A^* A = \mathrm{Id}\} = \{A \in \mathbb{C}^{n\times n} |\ \forall x \in \mathbb{C}^n : \|Ax\|_2 = \|x\|_2\}$$

   bildet eine n^2-dimensionale Untermannigfaltigkeit des $2n^2$-dimensionalen $\mathbb{R}$-Vektorraums $\mathbb{C}^{n\times n}$. Hierbei bedeutet $A^* = \overline{A}^t$ die hermitesch konjugierte Matrix zu A, also die elementweise konjugiert komplex und transponiert genommene Matrix zu A.
   Hinweis: Betrachten Sie den n^2-dimensionalen $\mathbb{R}$-Vektorraum

   $$V = \{A \in \mathbb{C}^{n\times n} |\ A = A^*\}$$

   der *hermiteschen Matrizen* und für eine *unitäre Matrix* $X_0 \in \mathrm{U}(n)$ die Abbildung

   $$f : \mathbb{C}^{n\times n} \to V, \quad f(X) = X_0 X^* X X_0^*.$$

2. Es sei $G = \operatorname{diag}(1, -1, -1, -1) \in \mathbb{R}^{4\times4}$. Die "*Lorentzgruppe*"

$$\mathrm{O}(1,3) := \{A \in \mathbb{R}^{4\times4} |\ A^t G A = G\}$$

bildet eine 6-dimensionale Untermannigfaltigkeit von $\mathbb{R}^{4\times4}$.
Hinweis: Betrachten Sie den 10-dimensionalen $\mathbb{R}$-Vektorraum

$$V = \{A \in \mathbb{R}^{4\times4} |\ A^t = A\}$$

der symmetrischen reellen 4×4-Matrizen und für eine Matrix $X_0 \in \mathrm{O}(1,3)$ die Abbildung

$$f : \mathbb{R}^{4\times4} \to V, \quad f(X) = (X_0^{-1})^t X^t G X X_0^{-1}.$$

Die linearen Abbildungen $\mathbb{R}^4 \to \mathbb{R}^4$, $x \mapsto Lx$ mit $L \in \mathrm{O}(1,3)$ werden *Lorentztransformationen* genannt und spielen in der speziellen Relativitätstheorie von Albert Einstein eine zentrale Rolle.

Übung 2.83 (Die spezielle lineare Gruppe) Zeigen Sie: Die reelle "*spezielle lineare Gruppe*"

$$\mathrm{SL}_n(\mathbb{R}) := \{A \in \mathbb{R}^{n\times n} |\ \det A = 1\}$$

bildet eine Hyperfläche (also eine 1-kodimensionale Untermannigfaltigkeit) in $\mathbb{R}^{n\times n}$.
Zeigen Sie ebenso, dass die komplexe spezielle lineare Gruppe

$$\mathrm{SL}_n(\mathbb{C}) := \{A \in \mathbb{C}^{n\times n} |\ \det A = 1\}$$

eine reell 2-kodimensionale Untermannigfaltigkeit des $2n^2$-dimensionalen $\mathbb{R}$-Vektorraums $\mathbb{C}^{n\times n}$ bildet.

Wir zeigen nun, dass eine m-dimensionale Untermannigfaltigkeit eines n-dimensionalen Raums lokal "bis auf Diffeomorphie" wie ein m-dimensionaler Untervektorraum eines n-dimensionalen $\mathbb{R}$-Vektorraums aussieht:

Lemma 2.84 (lokale Gestalt von Untermannigfaltigkeiten) *Es sei U ein n-dimensionaler $\mathbb{R}$-Vektorraum, $n \in \mathbb{N}$, und $M \subseteq U$ eine Teilmenge. Weiter seien $m \in \mathbb{N}$ mit $0 \le m \le n$ und $p \in \mathbb{N} \cup \{\infty\}$. Dann sind folgende Aussagen äquivalent:*

1. *M ist eine m-dimensionale C^p-Untermannigfaltigkeit von U,*

2. *Für jedes $x \in M$ gibt es einen n-dimensionalen $\mathbb{R}$-Vektorraum W, einen m-dimensionalen Untervektorraum $T \subseteq W$, eine offene Umgebung $U_1 \subseteq U$ von x in U, eine offene Umgebung $W_1 \subseteq W$ von 0 in W und einen C^p-Diffeomorphismus $F : U_1 \to W_1$ mit $F(x) = 0$ und*

$$F[M \cap U_1] = T \cap W_1.$$

Zusatz zu "1.⇒2.": *Ist $M \cap U_1 = f^{-1}[\{f(x)\}]$ eine lokale Beschreibung von M nahe bei $x \in M$ mit $f \in C^p(U_1, V)$ und $df_x[U] = V$ wie in der Definition 2.77 von Untermannigfaltigkeiten, so kann man nach eventueller Verkleinerung von U_1*

$$T = \mathrm{Ker}(df_x) \subseteq T \oplus V =: W$$

und[25]

$$F : U_1 \to W, \quad F(y) = (g(y) - g(x), f(y) - f(x)) \tag{65}$$

mit irgendeiner Fortsetzung[26] *der Identität $\mathrm{id}_T : T \to T$ zu einer linearen Abbildung $g : U \to T$ nehmen.*

Manchmal nennt man eine Abbildung F wie im Lemma einen lokalen *Flachmacher* von M bei x. Wenn man will, kann man auch einfach $W = \mathbb{R}^n$ und $V = \mathrm{span}_{\mathbb{R}}\{e_1, \ldots, e_m\} \subseteq \mathbb{R}^n$ nehmen, indem man zu einer Darstellung von Vektoren in W bezüglich einer Basis $b_, \ldots, b_n$ übergeht, die eine Basis $b_1, \ldots, b_m$ von T umfasst.
Beweis des Lemmas. "1.⇒2." Definieren wir T, V, W und F wie im Zusatz angegeben. Insbesondere ist $F(x) = 0$ sowie

$$\dim T = \dim \mathrm{Ker}(df_x) = \dim U - \dim(df_x[U]) = n - (n - m) = m.$$

und

$$\dim W = \dim T + \dim V = m + (n - m) = n = \dim U.$$

Weiter gilt $F \in C^p(U_1, W)$ mit

$$dF_x(y) = (dg_x(y), df_x(y)) = (g(y), df_x(y)) \text{ für } y \in U,$$

da wir g linear gewählt haben. Die Abbildung $dF_x : U \to W$ ist bijektiv. Um dies zu sehen, genügt es wegen $\dim U = \dim W$ zu zeigen, dass dF_x injektiv ist, also dass $\mathrm{Ker}(dF_x) = \{0\}$ gilt. In der Tat: Ist $dF_x(y) = 0$, also $g(y) = 0$ und $df_x(y) = 0$, also $y \in \mathrm{Ker}(df_x) = T$ und damit $0 = g(y) = \mathrm{id}_T(y) = y$. Nach dem lokalen Umkehrsatz gibt es eine offene Umgebung $\tilde{U}_1 \subseteq U_1$ von x und eine offene Umgebung W_1 von $F(x)$ in W, so dass $F|_{\tilde{U}_1} : \tilde{U}_1 \to W_1$ ein C^p-Diffeomorphismus ist. Nach eventueller Verkleinerung von U_1 dürfen wir o.B.d.A. $\tilde{U}_1 = U_1$ annehmen. Wir erhalten für beliebige $z = F(y) \in W_1$, $y \in U_1$ die Äquivalenz der folgenden Aussagen:

$$\begin{aligned} & z \in T \cap W_1 \\ \Leftrightarrow \quad & f(y) - f(x) = 0 \\ \Leftrightarrow \quad & y \in f^{-1}[\{f(x)\}] = M \cap U_1. \end{aligned}$$

[25] Mit der Identifikation von $t \in T$ mit $(t, 0) \in T \oplus V$ wird T als Untervektorraum $T \oplus 0$ von $T \oplus V = T \times V$ aufgefasst.

[26] Aus der Linearen Algebra verwenden wir, dass sich jede lineare Abbildung auf einem Untervektorraum eines endlichdimensionalen Vektorraums zu einer linearen Abbildung auf dem ganzen Raum fortsetzen lässt. Das Analogon dieser Aussage für *abgeschlossene* Unterräume von Banachräumen und *stetige* lineare Abbildungen ist übrigens manchmal falsch, in Hilberträumen jedoch wahr. Beim Aufbau der Theorie der "*Banachuntermannigfaltigkeiten*" muss man hierauf achten.

Damit ist $F[M \cap U_1] = T \cap W_1$ gezeigt.

"2.$\Rightarrow$1." Es gelte 2. Da $\dim T = m$ und $\dim W = n$, können wir einen $(n-m)$-dimensionalen $\mathbb{R}$-Vektorraum V und eine surjektive lineare Abbildung $h : W \to V$ mit $\operatorname{Ker} h = T$ wählen. Eine mögliche Wahl hierfür ist der Quotientenraum $V = W/T$ mit der kanonischen Abbildung $h : z \mapsto z + T$. Wir setzen $f := h \circ F \in C^p(U_1, V)$. Es gilt nach der Kettenregel $df_x = dh_0 \circ dF_x = h \circ dF_x$, da $F(x) = 0$ gilt und da h linear ist. Nun ist $h : W \to V$ surjektiv und $dF_x : U \to W$ bijektiv; also ist auch $df_x : U \to V$ surjektiv. Weiter gilt für alle $y \in U_1$, $z = F(y) \in W_1$ die Äquivalenz der folgenden Aussagen:

$$\begin{aligned} & y \in M \cap U_1 \\ \Leftrightarrow\quad & F[y] \in T \cap W_1 \subseteq \operatorname{Ker} h \\ \Leftrightarrow\quad & f(y) = h(F(y)) = 0 = h(0) = h(F(0)) = f(x) \\ \Leftrightarrow\quad & y \in f^{-1}[\{f(x)\}] \end{aligned}$$

Damit ist $M \cap U_1 = f^{-1}[\{f(x)\}]$ gezeigt.

□

Korollar 2.85 (Tangentialraum) Sind M eine m-dimensionale Untermannigfaltigkeit des n-dimensionalen Raums U und $x \in M$, so erhält man für alle lokalen Beschreibungen

$$M \cap U_1 = f^{-1}[\{f(x)\}]$$

von M mit Funktionen $f \in C^1(U_1, V)$, $df_x[U] = V$ wie in Definition 2.77 den *gleichen* m-dimensionalen Untervektorraum $\operatorname{Ker} df_x$ von U. Er wird *Tangentialraum* von M in x genannt und mit

$$T_x M := \operatorname{Ker}(df_x)$$

bezeichnet.

Beweis: Gegeben seien zwei Abbildungen $f : U_1 \to V$, $\tilde{f} : \tilde{U}_1 \to \tilde{V}$, die zwei lokale Beschreibungen von M nahe bei x wie in Definition 2.77 liefern. Wir dürfen o.B.d.A. $U_1 = \tilde{U}_1$ annehmen. Nach eventueller Verkleinerung von U_1 wählen wir C^p-Diffeomorphismen

$$\begin{aligned} F : U_1 \to W_1 \subseteq W = \operatorname{Ker}(df_x) \oplus V, &\quad F = (g - g(x), f - f(x)) \\ \tilde{F} : U_1 \to \tilde{W}_1 \subseteq \tilde{W} = \operatorname{Ker}(d\tilde{f}_x) \oplus \tilde{V}, &\quad \tilde{F} = (\tilde{g} - \tilde{g}(x), \tilde{f} - \tilde{f}(x)) \end{aligned}$$

wie in Lemma 2.84 Teil 2. inkl. Zusatz. Mit der Inklusionsabbildung $\iota : \operatorname{Ker}(df_x) \to \operatorname{Ker}(df_x) \oplus V = W$, $\iota(v) = (v, 0)$ und der Projektion $\pi : \operatorname{Ker}(d\tilde{f}_x) \oplus \tilde{V} \to \tilde{V}$, $\pi(a, b) = b$, betrachten wir die Komposition N der Abbildungen

$$\operatorname{Ker}(df_x) \supseteq \iota^{-1}[W_1] \xrightarrow{\iota} W_1 \xrightarrow{F^{-1}} U_1 \xrightarrow{\tilde{F}} \tilde{W}_1 \xrightarrow{\pi} \tilde{V},$$

also $N(v) = \pi(\tilde{F}(F^{-1}(\iota(v))))$ für $v \in \iota^{-1}[W_1]$. Man beachte, dass $\iota^{-1}[W_1]$ eine offene Umgebung von 0 in $\operatorname{Ker}(df_x)$ ist. Für $v \in \iota^{-1}[W_1]$ erhalten wir

$$\iota(v) = (v, 0) \in (\operatorname{Ker}(df_x) \oplus 0) \cap W_1 = F[M \cap U_1],$$

also

$$F^{-1}(\iota(v)) \in M \cap U_1,$$

also

$$\tilde{F}(F^{-1}(\iota(v))) \in \tilde{F}[M \cap U_1] \subseteq \mathrm{Ker}(d\tilde{f}_x) \oplus 0$$

und daher

$$N(v) = \pi(\tilde{F}(F^{-1}(\iota(v)))) = 0.$$

Die Abbildung $N : \iota^{-1}[W_1] \to \tilde{V}$ ist also die Nullabbildung. Es folgt mit der Kettenregel, dem lokalen Umkehrsatz und mit $\tilde{f} - \tilde{f}(x) = \pi \circ \tilde{F}$ für $y \in \mathrm{Ker}(df_x)$:

$$0 = dN_0(y) = d\pi_0 \circ d\tilde{F}_x \circ (dF_x)^{-1} \circ d\iota_0(y) = d\tilde{f}_x \circ (dF_x)^{-1}(y, 0)$$

da ι und π linear sind. Das bedeutet: $(dF_x)^{-1}(y, 0) \in \mathrm{Ker}(d\tilde{f}_x)$. Es gibt also ein $z \in \mathrm{Ker}(d\tilde{f}_x)$ mit

$$(g(z), df_x(z)) = (dg_x(z), df_x(z)) = dF_x(z) = (y, 0),$$

also $g(z) = y$ und $df_x(z) = 0$, also $z \in \mathrm{Ker}(df_x)$. Weil g eingeschränkt auf $\mathrm{Ker}(df_x)$ die Identität ist, schließen wir $z = g(z) = y$ und daher $y \in \mathrm{Ker}(d\tilde{f}_x)$. Damit ist gezeigt: $\mathrm{Ker}(df_x) \subseteq \mathrm{Ker}(d\tilde{f}_x)$. Mit vertauschten Rollen von f und $\tilde{f}$ folgt ebenso $\mathrm{Ker}(d\tilde{f}_x) \subseteq \mathrm{Ker}(df_x)$, also zusammen $\mathrm{Ker}(d\tilde{f}_x) = \mathrm{Ker}(df_x)$. Also ist der Tangentialraum T_xM wohldefiniert. Die Dimension des Tangentialraums erhalten wir aus

$$\dim T_xM = \dim \mathrm{Ker}\, df_x = \dim U - \dim df_x[U] = m - (n - m) = m.$$

$\square$

Übung 2.86 Zeigen Sie in der Situation von Lemma 2.84, Teil 2, dass $dF_{\tilde{x}}[T_{\tilde{x}}M] = T$ für alle $\tilde{x} \in M \cap U_1$ gilt.

Übung 2.87

1. Es sei $M \subset \mathbb{R}^3$ das durch die Gleichung

$$x^2 + 4xy + 2xz + 7y^2 - 2yz + 6z^2 = 18$$

beschriebene Ellipsoid. Berechnen Sie den "geometrischen Tangentialraum" $p + T_pM$ von M im Punkt $p = (1, 1, 1)$. Schreiben Sie $p + T_pM$ als ein geeignetes Niveaugebilde.

2. Berechnen Sie die Tangente im gleichen Punkt p der Ellipse, die durch den Schnitt von M mit der durch die Gleichung

$$2x - y + z = 0$$

beschriebenen Ebene entsteht. Geben Sie die Tangente durch eine Parametrisierung mit der ersten Koordinate x an.

Übung 2.88 (C^p-Abbildungen zwischen Untermannigfaltigkeiten)

1. Es seien U und U' zwei endlichdimensionale $\mathbb{R}$-Vektorräume, $M \subseteq U$ und $M' \subseteq U'$ zwei C^p-Untermannigfaltigkeiten, $p \in \mathbb{N} \cup \{\infty\}$. Weiter sei $f : M \to M'$ eine Abbildung. Wir nennen die Abbildung f *p-fach stetig differenzierbar*, in Zeichen $f \in C^p(M, M')$, wenn es für jedes $x \in M$ eine offene Umgebung U_1 von x in U und eine C^p-Fortsetzung $F : U_1 \to U'$ von $f|_{U_1 \cap M}$ gibt. Zeigen Sie, dass dann die Abbildung $dF_x|_{T_x M} : T_x M \to U'$ den Tangentialraum $T_x M$ in den Tangentialraum $T_{f(x)} M'$ abbildet und dass die Abbildung

$$df_x := dF_x|_{T_x M} : T_x M \to T_{f(x)} M'$$

nicht von der Wahl von F abhängt. Wir nennen df_x die *Linearisierung* oder auch *Ableitung* von f bei x.

2. Es seien $f : M \to M'$ und $g : M' \to M''$ p-fach stetig differenzierbare Abbildungen zwischen Untermannigfaltigkeiten endlichdimensionaler Vektorräume. Zeigen Sie folgende Variante der Kettenregel für $x \in M$:

$$d(g \circ f)_x = dg_{f(x)} \circ df_x.$$

Übung 2.89 (Tangentialbündel) Es sei M eine m-dimensionale C^p-Untermannigfaltigkeit eines n-dimensionalen $\mathbb{R}$-Vektorraums U, wobei $p \geq 2$. Zeigen Sie, dass

$$TM := \{(x, y) \in U \oplus U |\; y \in T_x M\}$$

eine $2m$-dimensionale C^{p-1}-Untermannigfaltigkeit von $U \oplus U$ ist. *Hinweis:* Betrachten Sie die Abbildung $F : U_1 \times U \to V \oplus V$, $F(x, y) = (f(x), df_x(y))$, wenn M wie in der Definition von Untermannigfaltigkeiten lokal als Niveaugebilde von $f : U \to V$ gegeben wird. TM wird das *Tangentialbündel* von M genannt.

Übung 2.90 (Halbstetigkeit des Ranges)

(a) Es seien $m, n, k \in \mathbb{N}$ und $A \in \mathbb{R}^{n \times m}$ mit $\text{Rang}(A) = k$. Zeigen Sie: Es gibt eine Umgebung $U \subseteq \mathbb{R}^{n\times m}$ von A, so dass für alle $B \in U$ gilt: $\text{Rang}(B) \geq k$.

(b) Zeigen Sie in der Situation von (a): Ist speziell $k = \min\{m, n\}$, so gilt für alle $B \in U$ sogar $\text{Rang}(B) = k$.

Neben der lokalen Beschreibung von Untermannigfaltigkeiten als *Niveaugebilde* gibt es auch alternative lokale Darstellungen als *Graphen* oder durch *Parametrisierungen.* Bei der Erdkugel kennen Sie solche Beschreibungen ja schon.

Lemma 2.91 (Lokale Parametrisierungen und Karten) Es sei M eine Teilmenge eines n-dimensionalen $\mathbb{R}$-Vektorraums U und $0 \leq m \leq n$, $p \in \mathbb{N} \cup \{\infty\}$. Dann sind äquivalent:

1. M ist eine m-dimensionale C^p-Untermannigfaltigkeit von U

2. Für alle $x \in M$ gibt es eine offene Umgebung U_1 von x in U, eine offene Teilmenge T_1 eines m-dimensionalen $\mathbb{R}$-Vektorraums T, eine Abbildung $\pi \in C^p(T_1, U)$ mit $\dim d\pi_{\tilde{y}}[T] = m$ für alle $\tilde{y} \in T_1$, so dass $\pi[T_1] = M \cap U_1$ gilt und $\pi : T_1 \to M \cap U_1$ ein Homöomorphismus ist.

Zusatz: In diesem Fall gilt $T_{\tilde{x}}M = d\pi_{\tilde{y}}[T]$ für alle $\tilde{y} \in T_1$ und $\tilde{x} \in M \cap U_1$ mit $\tilde{x} = \pi(\tilde{y})$. Es sei $g : U \to T_xM$ irgendeine fixierte lineare Abbildung, die die Identität $\mathrm{id}_{T_xM} : T_xM \to T_xM$ fortsetzt. Dann kann man in 2. zusätzlich fordern: $T = T_xM$ und $g \circ \pi(\tilde{y}) - g \circ \pi(y) = \tilde{y} - y$ für $y = \pi^{-1}(x)$ und alle $\tilde{y} \in T_1$.

Jede Abbildung $\pi : T_1 \to M \cap U_1$ wie in 2. (ohne Zusatz) heißt eine *lokale Parametrisierung* von M bei x. Ihre Umkehrung $\pi^{-1} : M \cap U_1 \to T_1$ heißt *(C^p-)Karte* bei x von M. Eine Familie von (C^p-)Karten, deren Definitionsbereiche eine Überdeckung von M bilden, heißt *(C^p-)Atlas*.

Die Sprechweise stammt natürlich aus der Geographie: Eine Karte zur Erdkugel bildet ein Stück $M \cap U_1$ der Erdoberfläche $M = S^2$ auf ein 2-dimensionales Papierstück T_1 ab, und ein Weltatlas ist eine Sammlung von Karten, bei der jedes Stück der Erdoberfläche irgendwo abgebildet ist.

Beweis des Lemmas “1.⇒2.” Es sei $F : U_1 \to W_1 \subseteq W = T \oplus V$ ein Flachmacher von M bei $x \in M$ mit $F[M \cap U_1] = T \cap W_1 =: T_1$ und $dF_{\tilde{x}}[T_{\tilde{x}}M] = T$ für alle $\tilde{x} \in U_1 \cap M$ wie in Lemma 2.84 und Übung 2.86. Dann leistet die Abbildung

$$\pi := (F|_{T \cap W_1})^{-1} : T \cap W_1 \to M \cap U_1$$

das Gewünschte: Setzen wir $y = F(x)$. Da $F : W_1 \to U_1$ ein Diffeomorphismus ist, ist π ein Homöomorphismus, und es gilt $d\pi_{\tilde{y}}[T] = (dF_{\tilde{x}})^{-1}[T] = T_{\tilde{x}}M$ für $\tilde{x} = \pi(\tilde{y})$, $\tilde{y} \in T_1$, und daher $\dim d\pi_{\tilde{y}}[T] = \dim T_{\tilde{x}}M = m$.

Beweis des Zusatzes: Es sei $g : U \to T_xM$ eine lineare Abbildung wie in der Voraussetzung. Wählt man dazu den Flachmacher F von oben wie im Zusatz des Lemma 2.84, insbesondere $y = F(x) = 0$, so folgt für $\tilde{y} \in T_1$ mit Formel (65) aus Lemma 2.84:

$$g(\pi(\tilde{y})) - g(\pi(y)) = g(F^{-1}(\tilde{y})) - g(x) = \tilde{y} = \tilde{y} - y.$$

Man beachte, dass hier $y \in T$ mit $(y, 0) \in T \oplus V = W$ identifiziert ist.

“2.⇒1.” Es sei $\pi : T_1 \to M \cap U_1$, $T_1 \subseteq T$, eine lokale Parametrisierung von M bei $x = \pi(y)$ mit $\dim T = \dim d\pi_y[T] = m$ und einer offenen Umgebung U_1 in U von x. O.B.d.A. dürfen wir $y = 0$ annehmen; sonst verschieben wir um y, d.h. wir ersetzen wir π durch $T_1 - y \ni y' \mapsto \pi(y + y')$. Wir wählen einen Komplementärraum V von $d\pi_y[T]$ in U, also einen Untervektorraum $V \subseteq U$ mit $d\pi_y[T] + V = U$ und $d\pi_y[T] \cap V = \{0\}$. Insbesondere ist $\dim V = \dim U - \dim d\pi_y[T] = n - m$. Wir definieren $G : T_1 \times V \to U$, $G(\tilde{y}, v) = \pi(\tilde{y}) + v$. Dann ist $dG_{(0,0)}(y', v') = d\pi_0(y') + v'$, also $dG_{(0,0)} : T \oplus V \to U$ surjektiv. Wegen $\dim(T \oplus V) = n = \dim U$ ist dann die lineare Abbildung $dG_{(0,0)}$ sogar bijektiv. Nach dem lokalen Umkehrsatz und (wenn nötig) Verkleinern von U_1 ist dann die Einschränkung von G auf eine geeignete Umgebung W_1 von $(0,0) \in W := T \oplus V$ ein Diffeomorphismus $G : W_1 \to U_1$ Die Umkehrung $F := G^{-1} : U_1 \to W_1$ ist dann ein Flachmacher von M nahe bei x mit $F[M \cap U_1] = T \cap W_1$. Die Behauptung folgt dann aus Teil 2.$\Rightarrow$ 1. von Lemma 2.84.

$\square$

Übung 2.92 1. **(Fortsetzung von Karten zu Flachmachern)** Es seien M eine m-dimensionale Untermannigfaltigkeit eines n-dimensionalen $\mathbb{R}$-Vektorraums U, T ein m-dimensionaler Unterraum eines n-dimensionalen $\mathbb{R}$-Vektorraums W, $x \in M$ und M_1 eine offene Umgebung von x in M. Weiter sei $\phi : M_1 \to T_1 \subseteq T$ eine Karte von M bei x mit $\phi(x) = 0$. Zeigen Sie, dass es eine offene Umgebung U_1 von x in U und einen Flachmacher $\Phi : U_1 \to W_1 \subseteq W$ mit $\Phi|_{M_1 \cap U_1} = \phi|_{M_1 \cap U_1}$ gibt.

2. **(Kartenwechsel)** Nun sei $\tilde{\phi} : \tilde{M}_1 \to \tilde{T}_1 \subseteq \tilde{T}$ eine weitere solche Karte bei x. Zeigen Sie, dass
$$\tilde{\phi} \circ \phi^{-1} : \phi[M_1 \cap \tilde{M}_1] \to \tilde{\phi}[M_1 \cap \tilde{M}_1]$$
ein Diffeomorphismus ist.

Übung 2.93 (Lokale Darstellung von Mannigfaltigkeiten als Graph) a) Es sei M eine m-dimensionale Untermannigfaltigkeit eines n-dimensionalen Vektorraums U, $x \in U$ und $U = T \oplus V$ eine Zerlegung von U als direkte Summe eines m-dimensionalen Unterraums $T \subseteq U$ und eines dazu komplementären $(m-n)$-dimensionalen Unterraums V mit den zugehörigen Projektionen $\pi_T : U \to T$, $\pi_V : U \to V$, $\pi_T(t+v) = t$ und $\pi_V(t+v) = v$ für $t \in T$ und $v \in V$. Es gelte $\pi_T[T_xM] = T$. Zeigen Sie, dass man eine offene Umgebung T_1 von $\pi_T(x)$ in T, eine offene Umgebung V_1 von $\pi_V(x)$ in V und eine stetig differenzierbare Abbildung $f : T_1 \to V_1$ mit
$$M \cap (T_1 + V_1) = \{t + f(t) |\, t \in T_1\}$$
finden kann.

b) Nun sei speziell $U = \mathbb{R}^n$. Für $I \subseteq \{1, \ldots, n\}$ fassen wir $\mathbb{R}^I$ als Teilraum von $\mathbb{R}^n$ auf, indem wir jedes $(x_i)_{i \in I}$ mit Nullen zu einem Element von $\mathbb{R}^n$ fortsetzen. Weiter sei $I^c = \{1, \ldots, n\} \setminus I$. Zeigen Sie, dass man in a) stets ein m-elementiges $I \subseteq \{1, \ldots, n\}$ (abhängig von x und M) sowie $T = \mathbb{R}^I$, $V = \mathbb{R}^{I^c}$ wählen kann.
Hinweis: Verwenden Sie, dass jede $n \times m$-Matrix vom Rang m eine $m \times m$-Untermatrix vom Rang m besitzt. Richtig interpretiert ist also
$$\{(y, f(y)) |\, y \in T_1\} \subseteq \mathbb{R}^I \times \mathbb{R}^{I^c} = \mathbb{R}^n$$
mit offenem $T_1 \subseteq \mathbb{R}^I$ Darstellung von M nahe bei x also Graph.

Übung 2.94 (Stereographische Projektionen) 1. Zeigen Sie, dass die beiden stereographischen Projektionen
$$\phi_\pm : S^2 \setminus \{(0,0,\mp 1)\} \to \mathbb{R}^2, \quad \phi(x,y,z) = \frac{1}{1 \pm z}(x,y)$$
zusammen einen Atlas der Einheitssphäre bilden.

2. Zeigen Sie, dass die Linearisierungen dieser beiden stereographischen Projektionen Ähnlichkeitsabbildungen sind, also dass die Riemannsche Metrik

$$D(\phi_\pm^{-1})(x,y)^t \cdot D(\phi_\pm^{-1})(x,y)$$

für $(x,y) \in \mathbb{R}^2$ ein Vielfaches $c(x,y)\,\mathrm{Id}$ der Einheitsmatrix Id ist. Geometrisch gesprochen bedeutet das, dass die stereographischen Projektionen $\phi_\pm$ winkeltreue Abbildungen sind.

Übung 2.95 (Das Möbiusband) 1. Zeigen Sie, dass das "Möbiusband"

$$M = \{(\cos(2x), \sin(2x), y\cos x, y\sin x)|\ x, y \in \mathbb{R}\}$$

eine zweidimensionale Untermannigfaltigkeit von $\mathbb{R}^4$ bildet.

2. Basteln Sie sich mit Papier, Schere und Klebstoff ein Papiermodell des Möbiusbands.

2.12 Stationäre Punkte unter Nebenbedingungen

Beispiel 2.96 Für welchen Punkt (x,y) auf der Einheitskreislinie, beschrieben durch die Gleichung $x^2 + y^2 = 1$, nimmt $x + y$ den größten bzw. den kleinsten Wert an? In diesem Beispiel kann man das noch geometrisch-anschaulich ohne Rechnung sehen: $(x,y) = \pm(1/\sqrt{2}, 1/\sqrt{2})$. Dort verlaufen nämlich die Niveaugeraden der "Zielfunktion" $(x,y) \mapsto x+y$ tangential zum Einheitskreis. Das Minimum bzw. Maximum, das es wegen der Kompaktheit des Einheitskreises geben muss, wird dort angenommen.

Verallgemeinern die Fragestellung: Gegeben sei eine stetig differenzierbare "*Zielfunktion*" $f : U \to \mathbb{R}$, $U \subseteq \mathbb{R}^n$ offen, und $k \le n$ Gleichungen, die "*Nebenbedingungen*":

$$\begin{aligned} h_1(x_1,\dots,x_n) &= c_1, \\ h_2(x_1,\dots,x_n) &= c_2, \\ \vdots \quad & \quad \vdots \\ h_k(x_1,\dots,x_n) &= c_k. \end{aligned} \tag{66}$$

mit stetig differenzierbaren Funktionen $h_j : U \to \mathbb{R}$, $j = 1,\dots,k$. Das Gleiche in Kurznotation:

$$h(x) = c$$

mit $h = (h_1,\dots,h_k) \in C^1(U, \mathbb{R}^k)$, $x = (x_1,\dots,x_n) \in U$ und $c = (c_1,\dots,c_k) \in \mathbb{R}^k$. Gesucht sind lokale Extrema von f unter den Nebenbedingungen (66). Damit sind lokale Extrema der Einschränkung $f|_M$ von f auf den Lösungsraum $M = \{x \in U|\ h(x) = c\} \subseteq U$ des Gleichungssystems (66) gemeint. Machen wir nun die Regularitätsannahme, dass die Linearisierungen $dh_j(x)$, $j = 1,\dots,k$ an jeder Stelle $x \in M$ voneinander linear unabhängig sind. Anders gesagt: Rang $Dh(x) = k$ für alle $x \in M$. Das bedeutet anschaulich, dass in linearisierter Approximation die k Nebenbedingungen voneinander unabhängig sind. Diese Regularitätsannahme garantiert, dass M eine k-kodimensionale Untermannigfaltigkeit von $U \subseteq \mathbb{R}^n$ ist. Ähnlich wie früher sind Punkte mit verschwindender Ableitung auf M Kandidaten für lokale Extrema:

Definition 2.97 (Stationäre Punkte auf Untermannigfaltigkeiten) Es sei M eine Untermannigfaltigkeit von $\mathbb{R}^n$ und $g \in C^1(M, \mathbb{R})$. Ein Punkt $x \in M$ heißt *stationärer Punkt* der Funktion g, wenn $dg_x : T_xM \to \mathbb{R}$ die Nullabbildung ist.

Wird $g = f|_M$ durch Einschränkung einer Funktion $f \in C^1(U, \mathbb{R})$ ($U \subseteq \mathbb{R}^n$ offen) gegeben, so können wir die Stationaritätsbedingung $dg_x = 0$ äquivalent durch

$$T_xM \subseteq \operatorname{Ker} df_x \tag{67}$$

beschreiben, da dg_x die Einschränkung von df_x auf den Tangentialraum T_xM ist. Äquivalent dazu können wir diesen Begriff der Stationarität auf Untermannigfaltigkeiten auf den Stationaritätsbegriff aus Definition 2.46 zurückführen: Ist $\pi : T_1 \to M$ eine lokale Parametrisierung von M bei $x = \pi(y) \in M$ mit einem offenen Definitionsbereich $T_1 \subseteq T$ in einem endlichdimensionalen Vektorraum T, so ist x genau dann ein stationärer Punkt von g, wenn y ein stationärer Punkt von $g \circ \pi$ ist. In der Tat: $dg_x = 0 \Leftrightarrow d(g \circ \pi)_y = dg_x \circ d\pi_y = 0$, da $d\pi_y : T \to T_xM$ ein Isomorphismus ist.

Lemma 2.98 (Lokale Extrema auf Untermannigfaltigkeiten sind stationär.) *Seien $M \subseteq \mathbb{R}^n$ eine $\mathcal{C}^1$-Untermannigfaltigkeit und $g \in C^1(M, \mathbb{R})$. Dann ist jedes lokale Extremum $x \in M$ von g ein stationärer Punkt.*

Beweis: Die folgt unmittelbar mittels einer lokalen Parametrisierung $\pi : T_1 \to M$ von M bei $x = \pi(y) \in M$: Ist $x = \pi(y)$ ein lokales Extremum von g, so ist y ein lokales Extremum von $g \circ \pi$, da π ein Homöomorphismus von T_1 auf eine Umgebung von x in M ist. Nach Lemma 2.47 ist dann y ein stationärer Punkt von $g \circ \pi$, also $x = \pi(y)$ ein stationärer Punkt von g.

$\square$

Korollar 2.99 (Charakterisierung stationärer Punkte von Einschränkungen) *Seien $M \subseteq \mathbb{R}^n$ eine $\mathcal{C}^1$-Untermannigfaltigkeit, $x \in M$, $U \subseteq \mathbb{R}^n$ eine offene Umgebung von x in $\mathbb{R}^n$ und $f \in \mathcal{C}^1(U)$. Hat die Einschränkung von f auf $U \cap M$ bei x ein lokales Extremum, so gilt*

$$T_xM \subseteq \operatorname{Ker} df_x.$$

Beweis: Die folgt unmittelbar aus dem vorhergehenden Lemma und der Charakterisierung (67) stationärer Punkte.

$\square$

Erinnerung an den Begriff der dualen Abbildung aus der Linearen Algebra; siehe Definition 2.27:

Es seien V und W $\mathbb{K}$-Vektorräume und $L : V \to W$ eine lineare Abbildung. Dann wird

$$L^* : W' \to V', \quad L^*(\mu) = \mu \circ L$$

die duale Abbildung von L genannt.

Satz 2.100 (Lagrange-Multiplikatoren) *Es seien $M \subseteq \mathbb{R}^n$ eine k-kodimensionale $\mathcal{C}^1$-Untermannigfaltigkeit, $x \in M$, $U \subseteq \mathbb{R}^n$ eine offene Umgebung von x in $\mathbb{R}^n$ und $f \in \mathcal{C}^1(U, \mathbb{R})$. Weiter sei $U \cap M = \{y \in U \mid h(y) = c\}$ eine lokale Beschreibung der Untermannigfaltigkeit als Niveaugebilde, wobei $h \in \mathcal{C}^1(U, \mathbb{R}^k)$, so dass $dh_x : \mathbb{R}^n \to \mathbb{R}^k$ surjektiv ist. Dann sind äquivalent:*

a) $T_x M \subseteq \operatorname{Ker} df_x$*, d.h. x ist ein stationärer Punkt der Einschränkung $f|_{U\cap M}$.*

b) Es gibt $\mu \in (\mathbb{R}^k)'$ mit

$$\mu \circ dh_x = df_x.$$

Anders gesagt: $df_x \in (dh_x)^*[(\mathbb{R}^k)']$.

Hierbei bezeichnet $(\mathbb{R}^k)' = \operatorname{Hom}_{\mathbb{R}}(\mathbb{R}^k, \mathbb{R})$ den Dualraum von $\mathbb{R}^k$.

Beweis des Satzes: Die behauptete Äquivalenz ist der Spezialfall $V = \mathbb{R}^n$, $W = \mathbb{R}^k$, $\varphi = df_x$ und $L = dh_x$ des folgenden Lemmas aus der Linearen Algebra.

$\square$

.

Lemma 2.101 *Seien V und W Vektorräume über einem Körper $\mathbb{K}$, $L : V \to W$ eine lineare Abbildung und $\varphi \in V'$. Folgende Aussagen sind äquivalent:*

a) $\operatorname{Ker} L \subseteq \operatorname{Ker} \varphi$,

b) Es gibt $\mu \in W'$ mit $\mu \circ L = \varphi$.
Anders gesagt: $\varphi \in L^*[W']$.

Beweis: "b)⇒a)": Es sei $\mu \circ L = \varphi$ und $v \in \operatorname{Ker} L$, d.h. $L(v) = 0$. Dann folgt $\varphi(v) = \mu(L(v)) = 0$, also $v \in \operatorname{Ker} \varphi$.
"a)⇒b)": Wir definieren zunächst die Einschränkung $\kappa : L[V] \to \mathbb{K}$ der gesuchten linearen Abbildung $\mu : W \to \mathbb{K}$ auf $L[V]$. Sei hierzu $w \in L[V]$. Wir wählen ein beliebiges $v \in V$ mit $L(v) = w$ und setzen $\kappa(w) := \varphi(v)$. Diese Definition hängt nicht von der Wahl von v ab. Ist nämlich $v' \in V$ ein anderer Vektor mit $L(v') = w$, so folgt $L(v - v') = L(v) - L(v') = 0$, also $v - v' \in \operatorname{Ker} L$ und daher $v - v' \in \operatorname{Ker} \varphi$ wegen unserer Annahme a). Hieraus folgt $\varphi(v) = \varphi(v')$, also die Wohldefiniertheit von $\kappa(w)$.
Die Abbildung $\kappa : L[V] \to \mathbb{K}$ ist linear (Übung!), und es gilt $\kappa \circ L = \varphi$. Wir wählen eine beliebige lineare Fortsetzung[27] $\mu : W \to \mathbb{K}$ von $\kappa : L[V] \to \mathbb{K}$ und erhalten auch $\mu \circ L = \varphi$.

$\square$

[27] Eine solche lineare Fortsetzung gibt es stets, wie in der Linearen Algebra gezeigt wird. In unserer Anwendung ist $L = dh_x : \mathbb{R}^n \to \mathbb{R}^k$ surjektiv, so dass es in der Anwendung nicht nötig ist, eine Forsetzung zu suchen: $\mu = \kappa$.

Etwas abstrakter können wir damit die Aussage b) des Satzes 2.100 auch in der Form $df'_x \in (dh_x)^*[(\mathbb{R}^k)']$ schreiben. In Matrixnotation ausgedrückt besagt Aussage b) des Satzes folgendes:

Das lineare Gleichungssystem

$$(\mu_1, \ldots, \mu_k) \cdot Dh(x) = Df(x)$$

besitzt eine Lösung $(\mu_1, \ldots, \mu_k) \in \mathbb{R}^{1\times k}$.

Die Koeffizienten $\mu_1, \ldots, \mu_k$ einer solchen Lösung werden *Lagrange-Multiplikatoren* genannt. Stationäre Punkte $x = (x_1, \ldots, x_n) \in M \cap U$ von $f|_M$ werden also genau durch Lösungen $(x_1, \ldots, x_n, \mu_1, \ldots, \mu_k) \in U \times \mathbb{R}^k$ des Gleichungssystems

$$\begin{aligned} \sum_{j=1}^{k} \mu_j \frac{\partial h_j}{\partial x_i}(x_1, \ldots, x_n) = \frac{\partial f}{\partial x_i}(x_1, \ldots, x_n) \text{ für } i = 1, \ldots, n \\ h_j(x) = c_j \text{ für } j = 1, \ldots, k \end{aligned} \tag{68}$$

beschrieben. Dieses Gleichungssystem besitzt $n + k$ Gleichungen und ebensoviele Unbekannte $x_1, \ldots, x_n, \mu_1, \ldots, \mu_k$.

Übung 2.102 *Zeigen Sie unter den Voraussetzungen des Satzes 2.100, dass der "Lagrange-Multiplikator" μ für jeden stationären Punkt x eindeutig bestimmt ist.*

Beispiel 2.103 (Fortsetzung von Beispiel 2.96) Stationäre Punkte von

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

unter der Nebenbedingung $x^2 + y^2 = 1$ werden durch das Gleichungssystem

$$\begin{aligned} 2x\mu &= 1 \\ 2y\mu &= 1 \\ x^2 + y^2 &= 1 \end{aligned}$$

bestimmt. Es besitzt die beiden Lösungen

$$(x, y, \mu) \in \{(1/\sqrt{2}, 1/\sqrt{2}, 1/\sqrt{2}), (-1/\sqrt{2}, -1/\sqrt{2}, -1/\sqrt{2})\}.$$

Weil die Einheitskreislinie kompakt ist (sie ist nämlich abgeschlossen als Lösungsgebilde von $x^2 + y^2 = 1$ und beschränkt), muss der die eine Lösung (die erste) ein globales Maximum beschreiben, die andere Lösung ein globales Minimum.

Übung 2.104 a) Warum gibt es (mindestens) einen Punkt auf der Parabel P mit der Gleichung

$$x^2 + y = 0,$$

welcher zum Punkt $(0, -4)$ minimalen Abstand hat?

b) Bestimmen Sie alle solchen Punkte mit Hilfe von Lagrange-Multiplikatoren.

Übung 2.105 Gegeben seien zwei Punkte P, Q im Raum $\mathbb{R}^3$ mit Abstand $\|P-Q\|_2 > 2$. sowie die Einheitssphären K, L mit Zentrum P bzw. Q. Welche Punkte $X \in K$ und $Y \in L$ haben minimalen Abstand voneinander, welche Punkte $X' \in K$, $Y' \in L$ haben maximalen Abstand voneinander? Berechnen Sie dies mit Hilfe von Lagrange-Multiplikatoren und interpretieren Sie das Ergebnis geometrisch.

Übung 2.106 Die Gleichungen

$$\begin{aligned} x^2 + 4xy + 2xz + 7y^2 - 2yz + 5z^2 &= 17, \\ x + y + z &= 3 \end{aligned}$$

beschreiben eine Ellipse im Raum $\mathbb{R}^3$. Für welchen Punkt auf dieser Ellipse nimmt $x + 5y + 2z$ den größten bzw. den kleinsten Wert an?

2.13 1-Formen und Kurvenintegrale

1-Formen sind uns schon als Ableitungen df von Funktionen vom Typ $\mathbb{R}^n \to \mathbb{R}$ begegnet.

Definition 2.107 (1-Formen) *Eine 1-Form (synonym: Pfaffsche Form, Differentialform 1. Grades[28]) auf einer offenen Menge $U \subseteq \mathbb{R}^n$ ist eine Abbildung $\omega : U \to (\mathbb{R}^n)'$.*

Anders gesagt ist eine 1-Form eine Linearkombination

$$\omega : U \to (\mathbb{R}^n)', \quad \omega_x = \sum_{j=1}^{n} f_j(x)\, dx_j$$

für $x = (x_1, \ldots, x_n) \in U$ mit Koeffizienten $f_j : U \to R$. In Matrixdarstellung wird solch eine 1-Formen ω durch den *Zeilenvektor* $(f_1, \ldots, f_n)$ repräsentiert:

$$\omega_x(h) = (f_1(x), \ldots, f_n(x)) \cdot \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = \sum_{j=1}^{n} f_j(x) h_j$$

für $x \in U$ und $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$. Besonders wichtig sind glatte 1-Formen $\omega \in C^\infty(U, (\mathbb{R}^n)')$. Allerdings ist nicht jede glatte 1-Form die Ableitung einer stetig differenzierbaren Funktion. Zum Beispiel ist die folgende Differentialform ω auf $\mathbb{R}^2$

$$\omega_{(x,y)} = y\, dx - x\, dy$$

keine Ableitung, lässt sich also *nicht* in der Form

$$\omega = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$$

[28] Differentialformen höheren Grades behandeln wir erst in der Analysis 3.

schreiben. Es folgte nämlich

$$1 = \frac{\partial y}{\partial y} = \frac{\partial}{\partial y}\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}\frac{\partial f}{\partial y} = \frac{\partial(-x)}{\partial x} = -1,$$

ein Widerspruch. Wir verallgemeinern nun diese Beweisidee. Hierzu einige Notationen. Um eine Buchhaltung mit Differenzierbarkeitsstufen zu vermeiden, beschränken wir uns oft auf glatte 1-Formen.

Definition 2.108 (geschlossene und exakte 1-Formen.) Es sei $U \subseteq \mathbb{R}^n$ offen.

a) Eine glatte 1-Form $\omega \in C^\infty(U, (\mathbb{R}^n)')$ heißt *exakt*,[29] wenn sie eine Ableitung ist, d.h. wenn es eine glatte Funktion $f \in C^\infty(U, \mathbb{R})$ mit Ableitung $df = \omega$ gibt. Mit

$$B^1(U) = \{df | \, f \in C^\infty(U, \mathbb{R})\}$$

bezeichnen wir den Raum aller exakten 1-Formen auf U.

b) Eine glatte 1-Form $\omega \in C^\infty(U, (\mathbb{R}^n)')$, geschrieben als Linearkombination der kanonischen Projektionen dx_j in der Form

$$\omega = \sum_{j=1}^{n} f_j \, dx_j \in C^\infty(U, (\mathbb{R}^n)'),$$

heißt *geschlossen*,[30], wenn gilt:

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i} \text{ für } i, j = 1, \ldots, n.$$

Mit

$$Z^1(U) = \{\omega \in C^\infty(U, (\mathbb{R}^n)') | \, \omega \text{ ist geschlossen } \}$$

bezeichnen wir den Raum aller geschlossenen 1-Formen über U.

Lemma 2.109 (exakt $\Rightarrow$ geschlossen) *Jede exakte 1-Form ist geschlossen. Es gilt also* $B^1(U) \subseteq Z^1(U)$.

Beweis: Ist $\omega = df \in B^1(U)$ mit $f \in C^\infty(U, \mathbb{R})$, also

$$\omega = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} \, dx_j$$

so folgt für $i, j = 1, \ldots, n$:

$$\frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j} = \frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i}$$

nach dem Lemma von Schwarz. Also ist $\omega \in Z^1(U)$.

[29] Synonym: "*de Rham 1-Korand*"

[30] Synonym: "*de Rham 1-Kozykel*". Was "1-Kozykeln" mit "1-Zykeln" (aus geschlossenen Kurven zusammengesetzten Gebilden) und Koränder mit "Rändern" zu tun haben, wird zu einem Teil schon unten, genauer aber erst später in der Analysis 3 oder in der Algebraischen Topologie klar werden.

$\square$

Ob die Umkehrung des Lemmas gilt, also ob jede geschlossene Form auf U auch exakt ist, hängt von der Gestalt des Gebiets U ab. Dies wird unten noch genauer analysiert.

Übung 2.110 a) Es sei $U \subseteq \mathbb{R}^n$ offen und $f \in C^1(U, \mathbb{R})$ mit $df = 0$. Zeigen Sie, dass f *lokal konstant* ist, d.h. dass es für alle $x \in U$ eine Umgebung $V \subseteq U$ von x gibt, so dass die Einschränkung $f|_V$ von f auf V konstant ist.

b) Folgern Sie, dass f konstant ist, falls U *zusammenhängend* ist, d.h. falls es keine Zerlegung $U = U_1 \cup U_2$, $U_1 \cap U_1 = \emptyset$ von U in zwei nichtleere, disjunkte offene Mengen U_1 und U_2 gibt.

c) Zeigen Sie, dass U zusammenhängend ist, wenn je zwei Punkte $x, y \in U$ durch eine stetige Kurve $k : [0, 1] \to U$, $k(0) = x$, $k(1) = y$, verbunden werden können.

Der Rückzug f^*dg von Ableitungen, den Sie im Abschnitt über die Kettenregel schon als "Einsetzoperation" oder "Substitutionsoperation" kennengelernt haben lässt sich wörtlich auch auf beliebige 1-Formen verallgemeinern:

Definition 2.111 (Rückzug) Es seien $U \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^m$ offen ($m, n \in \mathbb{N}$), $\omega : U \to (\mathbb{R}^n)'$ eine 1-Form und $f : V \to U$ differenzierbar. Der *Rückzug* (engl. *pullback*) von ω mit f wird so definiert:

$$f^*\omega : V \to (\mathbb{R}^m)', \quad (f^*\omega)_x = (df_x)^*\omega_{f(x)} = \omega_{f(x)} \circ df_x,$$

also

$$(f^*\omega)_x(h) = \omega_{f(x)}(df_x(h))$$

für $x \in V$, $h \in \mathbb{R}^m$.
Analog wird für eine Funktion $g : U \to \mathbb{R}$ auch die Komposition $g \circ f : V \to \mathbb{R}$ manchmal *Rückzug der Funktion* g mit f genannt und mit $f^*g = g \circ f$ bezeichnet.

Hat $f = (f_1, \ldots, f_n) : V \to U$ die Komponenten $f_j : V \to \mathbb{R}$, $j = 1, \ldots, n$ und ist

$$\omega_y = \sum_{j=1}^{n} \alpha_j(y)\, dy_j$$

die Komponentendarstellung von ω, so können wir das auch so schreiben:

$$(f^*\omega)_x = \sum_{j=1}^{n} \alpha_j(f(x))\, (df_j)_x = \sum_{j=1}^{n} \alpha_j(f(x)) \sum_{i=1}^{m} \frac{\partial f_j}{\partial x_i}(x)\, dx_i,$$

für $x \in V$, also

$$f^*\omega = \sum_{i=1}^{m} \beta_i\, dx_i \quad \text{mit} \quad \beta_i(x) = \sum_{j=1}^{n} \alpha_j(f(x)) \frac{\partial f_j}{\partial x_i}(x) \tag{69}$$

Der Rückzug $f^*\omega$ entsteht also aus $\omega = \sum_{j=1}^n \alpha_j\, dy_j$ durch Einsetzen (Substituieren) von $f(x)$ für y und df_j für dy_j.
Das Gleiche in Matrixnotation: Stellt der Zeilenvektor

$$(\alpha_1, \ldots, \alpha_n) : U \to \mathbb{R}^{1\times n}$$

die Form ω dar, so stellt der Zeilenvektor

$$(\alpha_1(f(x)), \ldots, \alpha_n(f(x))) \cdot Df(x) \in \mathbb{R}^{1\times m}$$

den Rückzug $(f^*\omega)_x$ an der Stelle $x \in V$ dar. Der Rückzug von 1-Formen wird also durch die Multiplikation mit der Jacobimatrix *von rechts* dargestellt.
Gelegentlich wird es nützlich sein, in der Definition 2.111 auf die Voraussetzung zu verzichten, dass U und V offen sind, z.B. für Intervalle $V = [a, b] \subseteq \mathbb{R}$ oder für Halbräume $U = [0, \infty[\times\mathbb{R}^{n-1}$. Wichtig zur Wohldefiniertheit ist dann nur, dass die Ableitung df_x auch noch an Randpunkten $x \in V \cap \partial V$ existiert, z.B. als einseitige Ableitung.

Übung 2.112 (Rückzug ist funktoriell) Es seien $l, m, n \in \mathbb{N}$, $U \subseteq \mathbb{R}^l$, $V \subseteq \mathbb{R}^m$ und $W \subseteq \mathbb{R}^n$ offene Mengen und $f : U \to V$, $g : V \to W$ differenzierbare Abbildungen, sowie $\omega : W \to (\mathbb{R}^n)'$ eine 1-Form. Zeigen Sie:

1. $f^*(g^*\omega) = (g \circ f)^*\omega$

2. $\mathrm{id}_W^*\, \omega = \omega$

Von früher kennen Sie die Schreibweise der Kettenregel mit Hilfe des Rückzugs:

Lemma 2.113 (Kettenregel: Rückzug vertauscht mit Ableitung) *Sind $U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^m$ offen und sind $f : V \to U$ und $g : U \to \mathbb{R}$ differenzierbar, so gilt*

$$f^*dg = d(f \circ g)$$

anders gesagt:

$$f^*dg = d(f^*g).$$

Beweis: Dies ist in der Tat nur eine andere Schreibweise der Kettenregel:

$$(f^*dg)_x = dg_{f(x)} \circ df_x = d(g \circ f)_x = (d(f^*g))_x$$

für $x \in V$.

□

Lemma 2.114 (Geschlossenheit und Exaktheit bleiben unter Rückzug erhalten.) *Es seien $U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^m$ offen, $f : V \to U$ glatt und $\omega \in C^\infty(U, (\mathbb{R}^n)')$ eine glatte 1-Form.*

1. *Ist ω geschlossen, so ist auch $f^*\omega$ geschlossen. Anders gesagt:*

$$f^*[Z^1(U)] \subseteq Z^1(V).$$

2. *Ist ω exakt, so ist auch $f^*\omega$ exakt. Anders gesagt:*

$$f^*[B^1(U)] \subseteq B^1(V).$$

Beweis. Die Behauptung 2. folgt unmittelbar aus dem vorhergehenden Lemma.
Zu Behauptung 1.: Es sei

$$\omega = \sum_{j=1}^{n} \alpha_j \, dy_j$$

geschlossen, also

$$\frac{\partial \alpha_j}{\partial y_k} = \frac{\partial \alpha_k}{\partial y_j} \tag{70}$$

für $j, k = 1, \ldots, n$. Setzen wir $y_j = f_j(x)$ für die Koordinaten ein, so erhalten aus Formel (69):

$$f^*\omega = \sum_{i=1}^{m} \beta_i \, dx_i \quad \text{mit} \quad \beta_i(x) = \sum_{j=1}^{n} \alpha_j(f(x)) \frac{\partial f_j}{\partial x_i}(x) \tag{71}$$

und daher für $i, l = 1, \ldots, m$ mit der Kettenregel:

$$\begin{aligned}
\frac{\partial \beta_i}{\partial x_l}(x) &= \sum_{j=1}^{n} \frac{\partial}{\partial x_l} \left[\alpha_j(f(x)) \frac{\partial f_j}{\partial x_i}(x) \right] \\
&= \sum_{j=1}^{n} \left[\frac{\partial \alpha_j(f(x))}{\partial x_l} \frac{\partial f_j(x)}{\partial x_i} + \alpha_j(f(x)) \frac{\partial^2 f_j}{\partial x_l \, \partial x_i}(x) \right] \\
&= \sum_{j=1}^{n} \sum_{k=1}^{n} \left. \frac{\partial \alpha_j(y)}{\partial y_k} \right|_{y=f(x)} \frac{\partial f_k(x)}{\partial x_l} \frac{\partial f_j(x)}{\partial x_i} + \sum_{j=1}^{n} \alpha_j(f(x)) \frac{\partial^2 f_j}{\partial x_l \, \partial x_i}(x)
\end{aligned}$$

Mit vertauschten Rollen von i und l und einer Umbenennung $j \leftrightarrow k$ lautet das:

$$\frac{\partial \beta_l}{\partial x_i}(x) = \sum_{k=1}^{n} \sum_{j=1}^{n} \left. \frac{\partial \alpha_k(y)}{\partial y_j} \right|_{y=f(x)} \frac{\partial f_j(x)}{\partial x_i} \frac{\partial f_k(x)}{\partial x_l} + \sum_{j=1}^{n} \alpha_j(f(x)) \frac{\partial^2 f_j}{\partial x_i \, \partial x_l}(x)$$

Wegen

$$\frac{\partial^2 f_j}{\partial x_i \, \partial x_l}(x) = \frac{\partial^2 f_j}{\partial x_l \, \partial x_i}(x)$$

erhalten wir

$$\frac{\partial \beta_l}{\partial x_i}(x) - \frac{\partial \beta_i}{\partial x_l}(x) = \sum_{k=1}^{n} \sum_{j=1}^{n} \left(\frac{\partial \alpha_k(y)}{\partial y_j} - \frac{\partial \alpha_j(y)}{\partial y_k} \right) \bigg|_{y=f(x)} \frac{\partial f_j(x)}{\partial x_i} \frac{\partial f_k(x)}{\partial x_l}. \qquad (72)$$

Mit (70) folgt hieraus die Behaupung:

$$\frac{\partial \beta_l}{\partial x_i} = \frac{\partial \beta_i}{\partial x_l}.$$

□

Bemerkung: Man beachte, dass zur Herleitung der Gleichung (72) nur die Gleichung $f^*\omega = f^*(\sum_j \alpha_j \, dy_j) = \sum_i \beta_i \, dx_i$ verwendet wurde, nicht jedoch die Geschlossenheit von ω. Die Gleichung (72) beschreibt also allgemein das Transformationsverhalten von Differenzen $D_k\alpha_j - D_j\alpha_k$ beim Rückzug einer 1-Form $\sum_j \alpha_j \, dy_j$.

Beispiel 2.115 Die 1-Form

$$\omega : \mathbb{R}^2 \setminus \{0\} \to (\mathbb{R}^2)', \quad \omega = \frac{x\,dy - y\,dx}{x^2 + y^2}$$

ist geschlossen, denn

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

Sie ist jedoch nicht exakt, wie wir unten sehen werden. Ihr Rückzug unter der Polarkoordinatenabbildung

$$f : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}^2 \setminus \{0\}, \quad f(r, \phi) = (r \cos \phi, r \sin \phi)$$

lautet

$$\begin{aligned} f^*\omega &= \frac{r \cos \phi \, d(r \sin \phi) - r \sin \phi \, d(r \cos \phi)}{(r \cos \phi)^2 + (r \sin \phi)^2} \\ &= \frac{1}{r^2} [r \cos \phi (\sin \phi \, dr + r \cos \phi \, d\phi) - r \sin \phi \, (\cos \phi \, dr - r \sin \phi \, d\phi) \\ &= (\cos^2 \phi + \sin^2 \phi) \, d\phi \\ &= d\phi. \end{aligned}$$

Der Rückzug $f^*\phi$ ist also exakt und daher auch geschlossen.
Allerdings ist die Einschränkung von ω auf $\mathbb{R}^+ \times \mathbb{R}$ dennoch exakt, wie Sie schon in Übung 2.29 gesehen haben: Es gilt nämlich

$$\omega|_{\mathbb{R}^+ \times \mathbb{R}} = dg$$

mit

$$g : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}, \quad g(x,y) = \arctan(y/x).$$

Insbesondere gilt für $r > 0$ und $-\frac{\pi}{2} < \phi < \frac{\pi}{2}$:

$$f^*g(r,\phi) = \arctan \frac{r \sin\phi}{r\cos\phi} = \arctan\tan\phi = \phi$$

also

$$d(f^*g) = d\phi = f^*\omega = f^*dg \text{ auf } \mathbb{R}^+ \times \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[.$$

Die fehlende Exaktheit von ω auf $\mathbb{R}^2 \setminus \{0\}$ wird also anschaulich gesprochen von der Mehrdeutigkeit des Winkels ϕ der Polarkoordinaten modulo 2π verursacht.

Kurvenintegrale. Man kann sich Differentialformen physikalisch als Kraftfelder vorstellen: Durchläuft ein Teilchen eine glatte Kurve $k : \mathbb{R} \to U$, so bedeutet $\omega_{k(t)}(k'(t))$ die Leistung, die das Kraftfeld zur Zeit t auf das Teilchen am Ort $k(t)$ mit der Geschwindigkeit $k'(t)$ überträgt.
Mit der vom euklidischen Skalarprodukt erzeugten Isomorphie $\mathbb{R}^n \to (\mathbb{R}^n)'$, $v \mapsto \langle v, \cdot\rangle$ kann man sich eine 1-Form ω auch mit Hilfe des Vektorfelds $V : U \to \mathbb{R}^n$ mit $\omega_x = \langle V(x), \cdot\rangle$ veranschaulichen. Bespielsweise haben wir uns früher Ableitungen $\omega = df$ mit Hilfe des Gradienten $V = \nabla f$ vorgestellt. So stellt man sich anschaulich oft Kräfte eher als Vektoren statt als Linearformen vor.
Das Zeitintegral über die Leistung ist die vom Kraftfeld erbrachte Arbeit. Dies gibt Anlaß zu der folgenden Definition:

Definition 2.116 (Kurvenintegral) *Es sei $\omega \in C(U, (\mathbb{R}^n)')$ eine stetige 1-Form auf einer offenen Menge $U \subseteq \mathbb{R}^n$, $[a,b]$ ein Intervall in $\mathbb{R}$ und $k \in C^1([a,b], U)$ eine stetig differenzierbare Kurve in U. Wir definieren das Kurvenintegral*

$$\int_k \omega := \int_a^b \omega_{k(t)}(k'(t))\,dt.$$

Anders gesagt:

$$\int_k \omega := \int_a^b k^*\omega,$$

wobei das Integral $\int_a^b g(t)\,dt$ der stetigen 1-Form $g(t)\,dt := (k^*\omega)_t$ auf dem Intervall $[a,b]$ im Sinne der Analysis 1 gelesen werden soll.
Analog definiert man das Integral eines stetigen Vektorfelds $V : U \to \mathbb{R}^n$ über die Kurve k durch

$$\int_a^b \langle V(k(t)), k'(t)\rangle\, dt.$$

Ist die Kurve k zwar stetig, aber nur stückweise stetig differenzierbar, so wird das Kurvenintegral $\int_k \omega$ als Summe der Integrale über die differenzierbaren Teilstücke definiert.

Beim Kurvenintegral über ein Vektorfeld V verwendet man also die Übersetzung in die Linearform $\langle V, \cdot \rangle$.
Das Kurvenintegral hängt nicht von der Geschwindigkeit ab, mit der die Kurve durchlaufen wird, sondern nur von der Durchlaufrichtung. Genauer gilt: Ist $[c, d]$ ein weiteres Intervall und $\pi : [c, d] \to [a, b]$ eine stetig differenzierbare Bijektion mit positiver Ableitung und bezeichnet $k \circ \pi$ die umparametrisierte Kurve (andere Durchlaufgeschwindigkeit bei gleicher Durchlaufrichtung), so gilt

$$\int_{k\circ\pi} \omega = \int_k \omega.$$

In der Tat gilt nach der Kettenregel und der Substitutionsregel:

$$\begin{aligned}
\int_{k\circ\pi} \omega &= \int_c^d \omega_{k(\pi(s))}((k \circ \pi)'(s))\, ds = \int_c^d \omega_{k(\pi(s))}(k'(\pi(s)))\pi'(s)\, ds \\
&= \int_{\pi(c)}^{\pi(d)} \omega_{k(t)}(k'(t))\, dt = \int_a^b \omega_{k(t)}(k'(t))\, dt = \int_k \omega.
\end{aligned}$$

Übung 2.117 *Überlegen Sie sich, dass sich nur das Vorzeichen des Kurvenintegrals ändert, wenn man die Durchlaufrichtung der Kurve umdreht.*

Die Parametrisierungsunabhängigkeit des Kurvenintegrals erlaubt uns manchmal eine etwas vergröberte Notation: Es sei $\mathcal{C} \subseteq \mathbb{R}^n$ eine unparametrisierte Kurve in $\mathbb{R}^n$. Parametrisieren wir sie durch eine Kurve

$$k : [a, b] \to \mathcal{C}.$$

Wenn keine Missverständnisse zu befürchten sind, schreibt man manchmal einfach vergröbert

$$\int_{\mathcal{C}} \omega := \int_k \omega.$$

Natürlich ist diese Schreibweise nur sinnvoll, wenn zusätzlich die Durchlaufrichtung der Kurve $\mathcal{C}$ spezifiziert wird. Zum Beispiel bedeutet für $x, y \in U$ mit $[x, y] \subseteq U$

$$\int_{[x,y]} \omega := \int_0^1 \omega_{(1-t)x+ty}(y - x)\, dt$$

das Kurvenintegral über die Strecke $[x, y]$, mit konstanter Geschwindigkeit durchlaufen von x nach y mit der Parametrisierung $k : [0, 1] \to [x, y]$, $k(t) = (1 - t)x + ty$.
In manchen Fällen hängt das Kurvenintegral nicht vom genauen Verlauf der Kurve ab:

Lemma 2.118 (Kurvenintegral von exakten 1-Formen) *Ist $\omega = df \in C(U, (R^n)')$ die Ableitung einer Funktion $f \in C^1(U, \mathbb{R})$ und ist $k : [a, b] \to U$ eine stetige, stückweise stetig differenzierbare Kurve, so gilt:*

$$\int_k df = f(k(b)) - f(k(a)).$$

Beweis: Aus der Definition des Kurvenintegrals, der Kettenregel und dem Hauptsatz der Differential- und Integralrechnung folgt

$$\int_k df = \int_a^b k^* df = \int_a^b d(f \circ k) = \int_a^b (f \circ k)'(t)\, dt = f(k(b)) - f(k(a)).$$

Das Lemma gibt uns eine Methode, eine "Stammfunktion" f einer exakten Form $\omega = df$ zu finden: Fixieren wir einen Referenzpunkt $x_0 \in U$ und nehmen wir an, dass jeder Punkt x in U durch eine stetige, stückweise stetig differenzierbare Kurve $k_x : [a, b] \to U$ mit x verbunden werden kann: $k_x(a) = x_0$, $k_x(b) = x$. Dann ist $f : U \to \mathbb{R}$, definiert durch

$$f(x) = \int_{k_x} \omega$$

eine Stammfunktion von ω.

Übung 2.119 Finden Sie eine Stammfunktion f der exakten 1-Form

$$\omega = 2xy\, dx + x^2\, dy$$

auf $\mathbb{R}^2$ auf drei verschiedene Weisen, indem Sie vom Referenzpunkt $(0, 0)$ zum Punkt (x, y) über folgende Kurven integrieren:

a) auf der direkten Verbindungsstrecke k_1 von $(0, 0)$ nach (x, y);

b) auf dem Polygonzug k_2, der durch Zusammenfügen der Verbindungsstrecke k_{2a} von $(0, 0)$ nach $(x, 0)$ und der Verbindungsstrecke k_{2b} von $(x, 0)$ nach (x, y) entsteht;

c) auf dem Polygonzug k_3, der durch Zusammenfügen der Verbindungsstrecke k_{3a} von $(0, 0)$ nach $(0, y)$ und der Verbindungsstrecke k_{3b} von $(0, y)$ nach (x, y) entsteht.

Skizzieren Sie die Bilder der Kurven k_1, k_2 und k_3 in der Ebene.

Stellen wir uns eine 1-Form ω physikalisch als "Kraftfeld" vor, so beschreiben die exakten 1-Formen die Kraftfelder, für die ein "*Energieerhaltungssatz*" in folgendem Sinn gilt: Die Arbeit, die das Kraftfeld leistet, wenn ein Teilchen vom Punkt x_0 zum Punkt x bewegt wird, hängt nur vom Anfangspunkt x_0 und vom Endpunkt x ab, nicht jedoch von der Wahl des Wegs dazwischen. Solche Kraftfelder nennt man in der Physik "*konservativ*".

2.14 Das Lemma von Poincaré und die erste de-Rham-Kohomologie

Wir beschäftigen uns jetzt mit dem Unterschied zwischen geschlossenen und exakten 1-Formen. Auf manchen Gebieten ist nicht jede geschlossene 1-Form ist exakt.

Beispiel 2.120 *Die 1-Form*

$$\omega = \frac{x\, dy - y\, dx}{x^2 + y^2} \in C^\infty(\mathbb{R}^2 \setminus \{0\}, (\mathbb{R}^2)')$$

aus Beispiel 2.115 ist zwar geschlossen, aber nicht exakt. Ihr Integral über den im positiven Sinn durchlaufenen Einheitskreis S^1 (Parametrisierung $k : [0, 2\pi] \to S^1,\ k(t) = (\cos t, \sin t)$) verschwindet nämlich nicht:

$$\begin{aligned}\int_{S^1} \omega = \int_k \frac{x\,dy - y\,dx}{x^2 + y^2} &= \int_0^{2\pi} \frac{\cos t\, d\sin t - \sin t\, d\cos t}{\cos^2 t + \sin^2 t} \\ &= \int_0^{2\pi} (\cos^2 t\, dt + \sin^2 t\, dt) = \int_0^{2\pi} dt = 2\pi.\end{aligned}$$

Anschaulich gesprochen bedeutet das, dass der Winkel in Polarkoordinaten sich um 2π ändert, wenn man einmal den Nullpunkt umläuft.

□

Integration einer gegebenen geschlossenen 1-Form ω über Kurven liefert uns dennoch in machen Fällen geeignete Kandidaten für eine Stammfunktion: Hierzu parametrisieren wir für $z, x \in \mathbb{R}^n$ die Verbindungsstrecke $[z, x]$ durch die Kurve $[z \to x] : [0, 1] \to [z, x]$, $t \mapsto (1 - t)z + tx$, die von z nach x mit konstanter Geschwindigkeit läuft.

Lemma 2.121 (Lemma von Poincaré für 1-Formen und sternförmige Gebiete) *Es sei $U \subseteq \mathbb{R}^n$ ein sternförmiges, offenes Gebiet mit Zentrum $z \in U$, d.h. $[z, x] \subseteq U$ für alle $x \in U$. Weiter sei $\omega \in Z^1(U)$ eine geschlossene Form. Dann ist*

$$f : U \to \mathbb{R}, \quad f(x) = \int_{[z \to x]} \omega = \int_0^1 \omega_{(1-t)z+tx}(x - z)\, dt$$

eine Stammfunktion von ω, also $\omega = df \in B^1(U)$. Insbesondere gilt:

$$\boxed{Z^1(U) = B^1(U) \text{ für sternförmige offene } U \subseteq \mathbb{R}^n}$$

Beweis: Man beachte, dass mit ω auch f glatt ist. Zur Vereinfachung der Notation nehmen wir o.B.d.A. an: $z = 0$; sonst verschieben wir das Gebiet U und die Form ω um $-z$ in den Nullpunkt. Es sei

$$\omega = \sum_{j=1}^n \alpha_j\, dx_j$$

die Komponentendarstellung von ω. Es folgt für $x \in U$:

$$f(x) = \int_0^1 \sum_{j=1}^n \alpha_j(tx) x_j\, dt$$

und daher für $k = 1, \ldots, n$:

$$\begin{aligned}
\frac{\partial f}{\partial x_k}(x) &= \int_0^1 \sum_{j=1}^n \frac{\partial}{\partial x_k}[\alpha_j(tx)x_j]\, dt \\
&= \int_0^1 \sum_{j=1}^n \left[\alpha_j(tx)\frac{\partial x_j}{\partial x_k} + tD_k\alpha_j(tx)x_j\right]\, dt \\
&= \int_0^1 \left[\alpha_k(tx) + \sum_{j=1}^n tD_k\alpha_j(tx)x_j\right] dt \\
&= \int_0^1 \left[\alpha_k(tx) + \sum_{j=1}^n tD_j\alpha_k(tx)x_j\right] dt \qquad \text{wegen } \omega \in Z^1(U) \\
&= \int_0^1 \left[\alpha_k(tx) + t\frac{\partial}{\partial t}\alpha_k(tx)\right] dt \\
&= \int_0^1 \frac{\partial}{\partial t}\,[t\alpha_k(tx)]\, dt \\
&= [t\alpha_k(tx)]_{t=0}^{t=1} \qquad \text{nach dem Hauptsatz} \\
&= \alpha_k(x).
\end{aligned}$$

Das bedeutet

$$df = \sum_{k=1}^n \frac{\partial f}{\partial x_k}\, dx_k = \sum_{j=1}^n \alpha_k\, dx_k = \omega.$$

□

Beispiel 2.122 Die Einschränkungen $\omega|_{U_1}$ und $\omega|_{U_2}$ der geschlossenen 1-Form

$$\omega = \frac{x\, dy - y\, dx}{x^2 + y^2} \in Z^1(\mathbb{R}^2 \setminus \{0\})$$

aus Beispiel 2.115 auf die "geschlitzten Ebenen"

$$\begin{aligned}
U_1 &:= \mathbb{R}^2 \setminus (]-\infty, 0] \times \{0\}) \\
U_2 &:= \mathbb{R}^2 \setminus ([0, +\infty[\times \{0\})
\end{aligned}$$

sind exakt, denn U_1 und U_2 sind sternförmig mit Zentren $(1, 0)$ bzw. $(-1, 0)$.

Übung 2.123 Zeigen Sie durch direkte Rechnung, dass

$$\begin{aligned}
\Phi_1 : U_1 \to\,]-\pi, \pi[, &\quad \Phi_1(x, y) = 2 \arctan \frac{y}{x + \sqrt{x^2 + y^2}}, \\
\Phi_2 : U_2 \to\,]-\pi, \pi[, &\quad \Phi_2(x, y) = 2 \arctan \frac{y}{x - \sqrt{x^2 + y^2}} \qquad (73)
\end{aligned}$$

Stammfunktionen von $\omega|_{U_1}$ bzw. von $\omega|_{U_2}$ sind:

$$d\Phi_1 = \omega|_{U_1}, \quad d\Phi_2 = \omega|_{U_2}.$$

Überlegen Sie sich mit elementargeometrischen Methoden, dass $\Phi_1(x,y)$ der Winkel zwischen der positiven x-Achse und dem Strahl von $(0,0)$ nach (x,y) ist, während $\Phi_2(x,y)$ der Winkel zwischen der negativen x-Achse und dem Strahl von $(0,0)$ nach (x,y) ist. Folgern Sie, dass $\Phi_1 - \Phi_2 : U_1 \cap U_2 \to \mathbb{R}$ lokal konstant ist, genauer

$$\Phi_1(x,y) - \Phi_2(x,y) = \begin{cases} \pi & \text{für } x \in \mathbb{R},\ y > 0, \\ -\pi & \text{für } x \in \mathbb{R},\ y < 0. \end{cases} \tag{74}$$

Der Unterschied zwischen dem Raum $Z^1(U)$ der geschlossenen Formen und dem Raum $B^1(U)$ der exakten Formen auf einer offenen Menge $U \subseteq \mathbb{R}^n$ wird durch den folgenden Quotientenraum beschrieben:

Definition 2.124 (1. de Rham Kohomologie) Es sei $U \subseteq \mathbb{R}^n$ offen. Der Quotientenraum

$$H^1(U) := Z^1(U)/B^1(U)$$

wird *1. de Rhamscher Kohomologieraum* von U genannt. Seine Elemente sind also Mengen der Gestalt

$$[\omega] := \omega + B^1(U) = \{\omega + df |\ f \in C^\infty(U, \mathbb{R})\}.$$

mit geschlossenen Formen $\omega \in Z^1(U)$. Die Klasse $[\omega]$ von $\omega \in Z^1(U)$ wird die *Kohomologieklasse* von ω genannt. Wir nennen zwei geschlossene Formen $\omega, \chi \in Z^1(U)$ *kohomolog*, wenn sie die gleiche Kohomologieklasse $[\omega] = [\chi]$ besitzen, d.h. wenn $\omega - \chi \in B^1(U)$ exakt ist.

Aus dem Lemma von Poincaré erhalten wir:

$$\boxed{H^1(U) = \{0\} \text{ für sternförmige offene } U \subseteq \mathbb{R}^n}$$

Anschaulich gesprochen misst $H^1(U)$ manche Typen von “Löchern” in U: Stellen wir uns U als einen “Schweizer Käse” im Anschauungsraum vor, so ist $H^1(U)$ sensitiv auf Löcher, “durch die man hindurchsehen kann”, jedoch nicht sensitiv auf Löcher, “die man von außen nicht sieht”.

Satz 2.125 (1. Kohomologie von $\mathbb{R}^2 \setminus \{0\}$) *Jede geschlossene Form $\chi \in Z^1(\mathbb{R}^2 \setminus \{0\})$ lässt sich in der Form*

$$\chi = \alpha \frac{x\,dy - y\,dx}{x^2 + y^2} + df \tag{75}$$

mit einem eindeutig bestimmten $\alpha \in \mathbb{R}$ und $f \in C^\infty(\mathbb{R}^2 \setminus \{0\}, \mathbb{R})$ schreiben. Das Gleiche anders gesagt: Der Kohomologieraum $H^1(\mathbb{R}^2 \setminus \{0\})$ ist ein eindimensionaler $\mathbb{R}$-Vektorraum mit der Basis

$$\left[\frac{x\,dy - y\,dx}{x^2 + y^2}\right].$$

Beweis: Wir verwenden die Bezeichnungen aus Übung 2.123. Insbesondere kürzen wir ab:

$$\omega = \frac{x\,dy - y\,dx}{x^2 + y^2} \in Z^1(\mathbb{R}^2 \setminus \{0\}).$$

Zur Eindeutigkeit der Darstellung (75). Wären $\alpha_1\omega + df_1 = \omega = \alpha_2\omega + df_2$ mit reellen Zahlen $\alpha_1 \neq \alpha_2$ zwei verschiedene Darstellungen wie in (75), so wäre

$$\omega = (\alpha_1 - \alpha_2)^{-1}(df_2 - df_1) = d\frac{f_2 - f_1}{\alpha_1 - \alpha_2}$$

exakt, ein Widerspruch.
Zur Existenz der Darstellung (75). Es sei $\chi \in Z^1(\mathbb{R}^2 \setminus \{0\})$ gegeben. Da die "geschlitzten Ebenen" U_1 und U_2 aus Übung 2.123 sternförmig sind, ist $\chi|_{U_j}$ für $j = 1, 2$ nach dem Lemma von Poincaré exakt; sagen wir $\chi|_{U_j} = dg_j$ mit $g_j \in C^\infty(U_j, \mathbb{R})$ für $j = 1, 2$. Die Differenz $g_1 - g_2$ ist auf $U_1 \cap U_2 = \mathbb{R} \times (\mathbb{R} \setminus \{0\})$ definiert, und es gilt dort $d(g_1 - g_2) = (\chi - \chi)|_{U_1 \cap U_2} = 0$. Also ist $g_1 - g_2$ auf $U_1 \cap U_2 = \mathbb{R} \times (\mathbb{R} \setminus \{0\})$ lokal konstant, also konstant auf $V_+ := \mathbb{R} \times \mathbb{R}^+$, sagen wir mit einem Wert $c_+ \in \mathbb{R}$, und konstant auf $V_- := \mathbb{R} \times \mathbb{R}^-$, sagen wir mit einem Wert $c_- \in \mathbb{R}$, da $\mathbb{R} \times \mathbb{R}^\pm$ zusammenhängend ist. Mit den Funktionen $\Phi_j : U_j \to \mathbb{R}$, $j = 1, 2$ aus Übung 2.123 setzen wir

$$\alpha := \frac{c_+ - c_-}{2\pi}$$

und

$$\begin{aligned} f_1 &:= g_1 - \frac{c_+ + c_-}{2} - \alpha\Phi_1, \\ f_2 &:= g_2 - \alpha\Phi_2. \end{aligned}$$

Insbesondere gilt für $j = 1, 2$:

$$df_j := dg_j - \alpha d\Phi_1 = \chi|_{U_j} - \alpha\omega|_{U_j}$$

also

$$\chi|_{U_j} = \alpha\omega|_{U_j} + df_j.$$

Dann folgt eingeschränkt auf $U_1 \cap U_2 = V_+ \cup V_-$ wegen (74):

$$\begin{aligned} f_1|_{U_1\cap U_2} - f_2|_{U_1\cap U_2} &= g_1|_{U_1\cap U_2} - g_2|_{U_1\cap U_2} - \frac{c_+ + c_-}{2} - \alpha(\Phi_1|_{U_1\cap U_2} - \Phi_2|_{U_1\cap U_2}) \\ &= c_+ 1_{V_+} + c_- 1_{V_-} - \frac{c_+ + c_-}{2} - \alpha\pi(1_{V_+} - 1_{V_-}) = 0. \end{aligned}$$

Hierbei bezeichnet $1_{V_\pm} : U_1 \cap U_2 \to \{0, 1\}$ die Indikatorfunktion von $V_\pm$, die den Wert 1 auf $V_\pm$ und den Wert 0 sonst annimmt. $f_1 : U_1 \to \mathbb{R}$ und $f_2 : U_2 \to \mathbb{R}$ stimmen also auf dem Durchschnitt $U_1 \cap U_2$ ihres Definitionsbereichs überein und besitzen daher eine gemeinsame Fortsetzung $f \in C^\infty(U_1 \cup U_2, \mathbb{R}) = C^\infty(\mathbb{R} \setminus \{0\}, \mathbb{R})$, $f|_{U_j} = f_j$ für $j = 1, 2$. Wir erhalten zusammen:

$$\chi = \alpha\omega + df,$$

wie gewünscht.

□

Übung 2.126 (1. Kohomologie von $\mathbb{R}^2 \setminus \{(\pm 1, 0)\}$) Zeigen Sie, dass $H^1(\mathbb{R}^2 \setminus \{(1,0),(-1,0)\})$ ein zweidimensionaler $\mathbb{R}$-Vektorraum mit der Basis

$$\left[\frac{y\,dx-(x-1)\,dy}{(x-1)^2+y^2}\right], \left[\frac{y\,dx-(x+1)\,dy}{(x+1)^2+y^2}\right]$$

ist. *Hinweis:* Betrachten Sie zu gegebenem $\chi \in Z^1(\mathbb{R}^2 \setminus \{(1,0),(-1,0)\})$ Stammfunktionen der Einschränkungen auf die sternförmigen Mengen $U_1 = \mathbb{R}^2 \setminus (]-\infty,-1] \times \{0\})$, $U_2 = \mathbb{R}^2 \setminus ([1,+\infty[\times \{0\})$ und $U_3 = \mathbb{R}^2 \setminus ([-1,1] \times \{0\})$. Gehen Sie damit analog zum Beweis von Satz 2.125 vor.

Übung 2.127 (1. Kohomologie von $\mathbb{R}^3 \setminus \{0\}$) Zeigen Sie, dass jede geschlossene 1-Form auf $\mathbb{R}^3 \setminus \{0\}$ exakt ist. Anders gesagt: $H^1(\mathbb{R}^3 \setminus \{0\}) = 0$. Betrachten Sie dazu die Überdeckung von $\mathbb{R}^3 \setminus \{0\}$ mit den beiden sternförmigen Mengen $U_1 = \mathbb{R}^3 \setminus \{(t,0,0) | \, t \le 0\}$ und $U_2 = \mathbb{R}^3 \setminus \{(t,0,0) | \, t \ge 0\}$. Beachten Sie, dass jede lokal konstante Funktion auf $U_1 \cap U_2$ konstant ist.[31]

Lemma 2.128 (Rückzug in der Kohomologie) *Es seien $V \subseteq \mathbb{R}^m$ und $U \subseteq \mathbb{R}^n$ offen und $f : V \to U$ glatt. Dann ist die Abbildung*

$$H^1(f) : H^1(U) \to H^1(V), \quad H^1(f)(\omega + B^1(U)) := f^*\omega + B^1(V) \text{ für } \omega \in Z^1(U)$$

wohldefiniert.

Diese Abbildung wird oft ebenfalls mit dem Symbol

$$f^* = H^1(f)$$

bezeichnet. Kurz gesagt: $f^*[\omega] := [f^*\omega]$.

Beweis des Lemmas: Ist $\omega \in Z^1(U)$ geschlossen, so ist $f^*\omega \in Z^1(V)$ nach Lemma 2.114 wieder geschlossen, und daher $[f^*\omega] \in H^1(V)$ definiert. Sind nun $\omega, \chi \in Z^1(U)$ kohomologe geschlossene Formen, also $\omega - \chi = dg$ für ein $g \in C^\infty(U,\mathbb{R})$, so folgt $f^*\omega - f^*\chi = f^*dg = df^*g \in B^1(V)$, also $[f^*\omega] = [f^*\chi] \in H^1(V)$. Daher hängt die Kohomologieklasse $[f^*\omega]$ nicht von der Wahl des Repräsentanten ω der Kohomologieklasse $[\omega]$ ab.

□

Wir betrachten nun "Deformationen von Abbildungen":

[31] Die erste de Rham Kohomologie $H^1(\mathbb{R}^3 \setminus \{0\})$ "sieht" also ein punktförmiges "Loch" in $\mathbb{R}^3$ nicht. Solch ein "Loch" wird erst von der zweiten de Rham Kohomologie $H^2(\mathbb{R}^3 \setminus \{0\})$ gesehen, die wir erst in der Analysis 3 definieren werden.

Definition 2.129 (Homotopie, Zusammenziehbarkeit) *Es seien V, U (genauer: $(U, \mathcal{T}_U)$, $(V, \mathcal{T}_V)$) zwei topologische Räume und $f, g : V \to U$ zwei stetige Abbildungen. Eine Homotopie von f nach g ist eine stetige Abbildung $h : [0,1] \times V \to U$, $(t, x) \mapsto h_t(x) := h(t, x)$, mit $h_0 = f$ und $h_1 = g$. Hierbei wird $[0,1] \times V$ mit der Produkttopologie versehen.*

Für uns von Interesse ist vorwiegend der Fall, dass $V \subseteq \mathbb{R}^m$, $U \subseteq \mathbb{R}^n$ offene Mengen sind, manchmal statt dessen auch mit Rändern, z.B. im Fall eines Intervalls $V = [a, b]$, und dass $f, g : V \to U$ und die Homotopie $h : [0,1] \times V \to U$ glatt sind.
Wir nennen eine offene Menge $U \subseteq \mathbb{R}^n$ auf einen Punkt $z \in U$ *(glatt) zusammenziehbar*, wenn es eine (glatte) Homotopie von der konstanten Abbildung $\mathrm{const}_z : U \to U$ mit dem Wert z zur Identität $\mathrm{id}_U : U \to U$ gibt.

Beispiel 2.130 Ist $U \subseteq \mathbb{R}^n$ sternförmig mit dem Zentrum $z \in U$, so ist

$$h : [0,1] \times U \to U, \quad h(t, x) = (1-t)z + tx$$

eine solche glatte Homotopie von const_z nach id_U. Mit dieser glatten Homotopie haben wir im Beweis des Poincaré-Lemmas schon gearbeitet. Jedes sternförmige Gebiet ist also glatt zusammenziehbar.

Nun sei $V \subseteq \mathbb{R}^m$ offen, $W := [0,1] \times V$. Die Koordinaten eines Punkts $(t, x) \in W$ bezeichnen wir mit $t, x_1, \ldots, x_m$. Gegeben $x \in V$, definieren wir die Kurve

$$k_x : [0,1] \to W, \quad k_x(t) := (t, x).$$

Analog setzen wir für $t \in [0,1]$:

$$\iota_t : V \to W, \quad \iota_t(x) := (t, x).$$

Wir definieren nun eine Abbildung $I_V : Z^1(W) \to C^\infty(V, \mathbb{R})$ wie folgt. Für eine geschlossene 1-Form

$$\chi = \alpha\, dt + \sum_{j=1}^{m} \beta_j\, dx_j \in Z^1(W)$$

setzen wir

$$I_V\chi : V \to \mathbb{R}, \quad I_V\chi(x) := \int_{k_x} \chi = \int_0^1 \alpha(t, x)\, dt.$$

Dann ist $I_V\chi$ glatt, und es gilt für $j = 1, \ldots, m$ und $x \in V$:

$$\begin{aligned}
\frac{\partial}{\partial x_j}(I_V\chi)(x) &= \int_0^1 \frac{\partial}{\partial x_j}\alpha(t, x)\, dt \\
&= \int_0^1 \frac{\partial}{\partial t}\beta_j(t, x)\, dt \qquad \text{da } \chi \in Z^1(W) \\
&= \beta_j(1, x) - \beta_j(0, x),
\end{aligned}$$

also

$$d(I_V\chi)_x = \sum_{j=1}^m \frac{\partial}{\partial x_j}(I_V\chi)(x)\,dx_j = \sum_{j=1}^m \beta_j(1,x)\,dx_j - \sum_{j=1}^m \beta_j(0,x)\,dx_j = (\iota_1^*\chi)_x - (\iota_0^*\chi)_x.$$

Es gilt also

$$\iota_1^*\chi - \iota_0^*\chi = d(I_V\chi) \in B^1(V). \tag{76}$$

Damit erhalten wir folgende Verallgemeinerung des Poincaré-Lemmas:

Satz 2.131 (Lemma von Poincaré für 1-Formen und Homotopien) *Es seien $U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^m$ offen*[32] *und $h : [0,1] \times V \to U$, $h(t,x) = h_t(x)$ eine glatte Homotopie. Dann gilt für alle $\omega \in Z^1(U)$:*

$$\boxed{h_1^*\omega - h_0^*\omega = dI_V h^*\omega \in B^1(V)}$$

Insbesondere sind $h_1^*\omega$ und $h_0^*\omega$ kohomolog. Die glatt homotopen Abbildungen $h_1 : V \to U$ und $h_0 : V \to U$ induzieren also die gleichen Abbildungen in der 1. Kohomologie:

$$\boxed{H^1(h_1) = H^1(h_0) : H^1(U) \to H^1(V)}$$

Beweis: Setzen wir $\chi := h^*\omega$, so ist $\chi \in Z^1(V)$ geschlossen nach Lemma 2.114, da ω geschlossen ist. Mit Formel (76) folgt wegen $h_t = h \circ \iota_t$, $t \in [0,1]$:

$$\begin{aligned} h_1^*\omega - h_0^*\omega &= (h \circ \iota_1)^*\omega - (h \circ \iota_0)^*\omega = \iota_1^* h^*\omega - \iota_0^* h^*\omega \\ &= \iota_1^*\chi - \iota_0^*\chi = d(I_V\chi) = dI_V h^*\omega \in B^1(V). \end{aligned}$$

Insbesondere folgt

$$H^1(h_1)([\omega]) = h_1^*\omega + B^1(V) = h_0^*\omega + B^1(V) = H^1(h_0)([\omega]).$$

□

Als eine wichtige Anwendung besprechen wir:

Korollar 2.132 (Fundamentalsatz der Algebra) *Jedes Polynom $f \in \mathbb{C}[X]$ vom Grad $n \geq 1$ besitzt mindestens eine Nullstelle in $\mathbb{C}$. Insbesondere zerfällt in $\mathbb{C}[X]$ jedes Polynom $f = a_0 + a_1 X + \ldots + a_n X^n \in \mathbb{C}[X]$, $a_n \neq 0$, in Linearfaktoren:*

$$f(X) = a_n \prod_{j=1}^n (X - \lambda_j), \qquad \textit{alle } \lambda_j \in \mathbb{C}.$$

[32] Auf die Voraussetzung der Offenheit kann man verzichten, wenn die benötigten Ableitungen noch an Randpunkten existieren und stetig sind.

Beweis: Für diesen Beweis ist es zweckmäßig, die Ebene $\mathbb{R}^2$ mit $\mathbb{C}$ mittels $(x, y) \equiv x + iy =: z$ zu identifizieren. Insbesondere steht $dx = \mathrm{Re} : \mathbb{C} \to \mathbb{R}$ für die Realteilbildung, $dy = \mathrm{Im} : \mathbb{C} \to \mathbb{R}$ für die Imaginärteilbildung, $dz = dx + i\,dy : \mathbb{C} \to \mathbb{C}$ für die Identität auf $\mathbb{C}$ und Multiplikation mit i für die 90°-Drehung in $\mathbb{R}^2$, also für eine spezielle $\mathbb{R}$-lineare Abbildung. Die geschlossene 1-Form

$$\omega = \frac{x\,dy - y\,dx}{x^2 + y^2}$$

auf $\mathbb{R}^2 \setminus \{(0, 0)\} \equiv \mathbb{C} \setminus \{0\} =: \mathbb{C}^*$ aus Beispiel 2.120 können wir damit in der Form

$$\omega = \mathrm{Im}\,\frac{x\,dx + ix\,dy - iy\,dx + y\,dy}{x^2 + y^2} = \mathrm{Im}\,\frac{(x - iy)(dx + idy)}{x^2 + y^2} = \mathrm{Im}\,\frac{\overline{z}\,dz}{\overline{z}z} = \mathrm{Im}\,\frac{dz}{z}$$

schreiben. Aus der (verallgemeinerten) Produktregel folgt: Die Potenzierung mit n, also $p : \mathbb{C}^* \to \mathbb{C}^*$, $p(z) = z^n$, aufgefasst als differenzierbare Abbildung vom Typ $\mathbb{R}^2 \setminus \{0\} \to \mathbb{R}^2 \setminus \{0\}$, besitzt die Ableitung $dp_z = nz^{n-1}\,dz$. Insbesondere folgt

$$(p^*\omega)_z = \mathrm{Im}\,\frac{dp_z}{p(z)} = \mathrm{Im}\,\frac{nz^{n-1}\,dz}{z^n} = n\,\mathrm{Im}\,\frac{dz}{z} = n\omega_z,$$

kurz geschrieben: $p^*\omega = n\omega$.

Nun sei ein Polynom $f(X) = a_0 + a_1X + \ldots + a_nX^n \in \mathbb{C}[X]$, vom Grad $n \in \mathbb{N}$ gegeben. O.B.d.A. nehmen wir $a_n = 1$ an, sonst dividieren wir f durch $a_n \neq 0$. *Indirekter Beweis:* Angenommen, f besitzt keine Nullstelle. Dann liefert f eine Abbildung $f : \mathbb{C} \to \mathbb{C}^*$. Insbesondere ist $f^*\omega$ nach Lemma 2.114 geschlossen, da ω geschlossen ist. Da $\mathbb{C}$ sternförmig ist und $f^*\omega \in Z^1(\mathbb{C})$, folgt aus dem Poincaré-Lemma 2.121: $f^*\omega$ ist exakt, sagen wir $f^*\omega = dg$ für ein $g \in C^\infty(\mathbb{C}, \mathbb{R})$. Insbesondere ist auch die Einschränkung von $f^*\omega$ auf $\mathbb{C}^*$ exakt: $(f|_{\mathbb{C}^*})^*\omega = (f^*\omega)|_{\mathbb{C}^*} = d(g|_{\mathbb{C}^*})$. Wir definieren nun eine glatte Homotopie $h : [0, 1] \times \mathbb{C}^* \to \mathbb{C}^*$ zwischen $p : \mathbb{C}^* \to \mathbb{C}^*$ und $f|_{\mathbb{C}^*} : \mathbb{C}^* \to \mathbb{C}^*$ so:

$$h(t, z) = \sum_{j=0}^{n} a_j t^{n-j} z^j.$$

Man beachte, dass für $0 < t \leq 1$ und $z \in \mathbb{C}^*$ gilt:

$$h(t, z) = t^n \sum_{j=0}^{n} a_j \left(\frac{z}{t}\right)^j = t^n f(z/t)$$

und $h(0, z) = z^n = p(z)$ sowie $h(1, z) = f(z)$, so dass in der Tat h nicht den Wert 0 annimmt, da f nach Annahme nicht den Wert 0 annimmt. Aus dem Poincaré Lemma 2.131 folgt: $(f|_{\mathbb{C}^*})^*\omega - p^*\omega$ ist exakt. Da $(f|_{\mathbb{C}^*})^*\omega$ exakt ist, folgt, dass auch $n\omega = p^*\omega$ exakt ist. Das ist ein Widerspruch, denn ω ist nicht exakt und $n \neq 0$. Also besitzt jedes nichtkonstante Polynom in $\mathbb{C}[X]$ eine Nullstelle, also mindestens einen Linearfaktor.

Die Linearfaktorzerlegung erhält man hieraus rekursiv durch sukzessives Abspalten von Linearfaktoren. Wir führen das hier nicht genauer aus, da dieser Beweisteil rein algebraisch ist und keine Analysis mehr braucht.

□

Korollar 2.133 (Lemma von Poincaré für 1-Formen und glatt zusammenziehbare Gebiete) *Ist $U \subseteq \mathbb{R}^n$ eine glatt zusammenziehbare offene Menge, so ist jede geschlossene 1-Form auf U exakt: $H^1(U) = \{0\}$.*

Beweis: Es sei $h : [0,1] \times U \to U$ eine Homotopie von einer konstanten Abbildung $h_0 = \mathrm{const}_z : U \to U$ nach der Identität $h_1 = \mathrm{id}_U$. Dann ist $d\,\mathrm{const}_z = 0$. Es folgt für alle $\omega \in Z^1(U)$ nach der Homotopie-Version des Poincaré-Lemmas:

$$\omega = \mathrm{id}_U^* \,\omega = \mathrm{id}_U^* \,\omega - \mathrm{const}_z^* \,\omega = h_1^*\omega - h_0^*\omega = dI_U h^*\omega \in B^1(U).$$

Insbesondere ist die 1-Form ω exakt. Wir erhalten $Z^1(U) = B^1(U)$, also $H^1(U) = \{0\}$.

□

Übung 2.134 Es sei $U \subseteq \mathbb{R}^n$ ein sternförmiges Gebiet mit Zentrum 0 und $h : [0,1] \times U \to U$, $h(t,x) = tx$. Berechnen Sie für $\omega = \sum_{j=1}^n \alpha_j \, dx_j \in Z^1(U)$ die Stammfunktion $f := I_U h^*\omega$ von ω in diesem Fall, und vollziehen Sie den Beweis von $df = \omega$ nach, herunterspezialisiert auf den Fall dieser Homotopie h. Sie rekonstruieren damit den Beweis des Poincaré-Lemmas für sternförmige Gebiete als Spezialfall des Beweises des Poincaré-Lemmas für glatt zusammenziehbare Gebiete.

Korollar 2.135 (Kurvenintegral über geschlossene Formen) *Es sei $U \subseteq \mathbb{R}^n$ offen, $h : [0,1] \times [a,b] \to U$, $h(t,s) =: h_t(s) =: k_s(t)$ eine glatte Homotopie von Pfaden. Es gelte mindestens eine der beiden folgenden Voraussetzungen:*

1. *Alle Pfade $h_t : [a,b] \to U$, $0 \le t \le 1$, besitzen den gleichen Anfangspunkt. $x = h_t(a)$ und den gleichen Endpunkt $y = h_t(b)$. Anders gesagt: $k_a = \mathrm{const}_x$ und $k_b = \mathrm{const}_y$.*
2. *Alle Pfade $h_t : [a,b] \to U$, $0 \le t \le 1$, sind geschlossen, d.h. $h_t(a) = h_t(b)$. Anders gesagt: $k_a = k_b$.*

Dann gilt für alle $\omega \in Z^1(U)$:

$$\int_{h_1} \omega = \int_{h_0} \omega,$$

d.h. das Kurvenintegral ändert sich nicht unter der Deformation mit h.

Beweis: Nach dem Poincaré-Lemma gilt

$$\begin{aligned}\int_{h_1} \omega - \int_{h_0} \omega &= \int_a^b (h_1^*\omega - h_0^*\omega) = \int_a^b dI_{[a,b]} h^*\omega \\ &= (I_{[a,b]} h^*\omega)(b) - (I_{[a,b]} h^*\omega)(a) = \int_0^1 (k_b^*\omega - k_a^*\omega) = 0\end{aligned}$$

wobei wir im letzten Schritt verwendet haben, dass $k_b^*\omega = 0 = k_a^*\omega$ für Voraussetzung 1. und $k_b^*\omega = k_a^*\omega$ für Voraussetzung 2. gilt.

□

Ausblick: Homologie von Kurven. Sozusagen "dual" zur Kohomologietheorie für 1-Formen gibt es auch eine *Homologietheorie* für geschlossene Kurven und "1-Zykeln", das sind aus mehreren geschlossenen Kurven mit "Vielfachheiten" zusammengesetzte Objekte. Die verbindende "Klammer" zwischen diesen beiden Theorien ist das (Kurven-)Integral. In dieser Vorlesung über Differentialrechnung liegt es nahe, die Homologierelation zwischen geschlossenen Kurven mittels "Dualität" mit geschlossenen 1-Formen, also mit Methoden der Differentialrechnung, zu definieren. Alternativ dazu gibt es auch einen rein topologischen Zugang zur Homologie, den wir hier nicht besprechen und den Sie in der Algebraischen Topologie lernen können.

Definition 2.136 (zusammenziehbare und nullhomologe Kurven) *Es sei $k : [a, b] \to U$ eine stetige, stückweise stetig differenzierbare geschlossene Kurve, d.h. $k(b) = k(a)$.*

1. *Die geschlossene Kurve k heißt (glatt) zusammenziehbar in U, wenn es eine (glatte) Homotopie $h : [0, 1] \times [a, b] \to U$ von geschlossenen Kurven h_t gibt, so dass $h_1 = k$ gilt und h_0 konstant ist.*
2. *Die geschlossene Kurve k heißt nullhomolog, in U, wenn für alle $\omega \in Z^1(U)$ gilt:*

$$\int_k \omega = 0.$$

Zwei stetige, stückweise stetig differenzierbare geschlossene Kurven k_1, k_2 mit Werten in U heißen homolog zueinander in U, wenn für alle $\omega \in Z^1(U)$ gilt:

$$\int_{k_1} \omega = \int_{k_2} \omega.$$

Nach dem vorhergehenden Korollar sind zwei glatte geschlossene Kurven homolog zueinander, wenn es eine glatte Homotopie geschlossener Kurven zwischen ihnen gibt. Insbesondere sind alle glatt zusammenziehbaren geschlossenen Kurven nullhomolog. Es gibt jedoch in manchen Gebieten nullhomologe Kurven, die nicht zusammenziehbar sind:

Beispiel 2.137 *Es sei $U = \mathbb{R}^2 \setminus \{(1, 0), (-1, 0)\}$. Weiter sei für $\sigma, \tau \in \{\pm 1\}$*

$$k_{\sigma,\tau} : [0, 2\pi] \to U, \quad k_{\sigma,\tau}(t) = (\sigma(1 - \cos t), -\sigma\tau \sin t)$$

die geschlossene Kurve, die den Kreis, mit Radius 1 um $(\sigma, 0)$ mit Umlaufrichtung τ und Start in $(0, 0)$ durchläuft. Durchläuft man erst $k_{1,1}$, dann $k_{-1,1}$, dann $k_{1,-1}$ und zuletzt $k_{-1,-1}$, so erhält man eine zusammengesetzte Kurve in U, die zwar nullhomolog, aber nicht zusammenziehbar ist.
Wir beweisen das hier nicht.

Die Umlaufzahl von Kurven.

Definition 2.138 (Umlaufzahl) *Es sei $k : [a, b] \to \mathbb{R}^2 \setminus \{(0, 0)\}$ eine stetige, stückweise stetig differenzierbare geschlossene Kurve. Die Umlaufzahl (synonym: Windungszahl) um 0 dieser Kurve wird durch*

$$N_0(k) := \frac{1}{2\pi} \int_k \frac{x\,dy - y\,dx}{x^2 + y^2}$$

definiert. Die Umlaufzahl von Kurven um andere Punkte $w \in \mathbb{R}^2$ statt 0 wird darauf mit einer Translation um w zurückgeführt:

$$N_w(k + w) := N_0(k).$$

Anschaulich gesprochen gibt die Umlaufzahl um 0 an, wie oft der Nullpunkt umlaufen wird, wobei Windungen im Uhrzeigersinn negativ und im Gegenuhrzeigersinn positiv gezählt werden.
Aus Korollar 2.135 folgt, dass zueinander homologe geschlossene Kurven die gleiche Umlaufzahl besitzen. Insbesondere besitzen nullhomologe und zusammenziehbare geschlossene Kurven k in $\mathbb{C}^*$ die Umlaufzahl $N_0(k) = 0$.

Lemma 2.139 (Ganzzahligkeit der Umlaufzahl) *Die Umlaufzahl $N_0(k)$ ist für jede geschlossene, stetige, stückweise stetig differenzierbare Kurve in $\mathbb{C}^*$ stets eine ganze Zahl.*

Beweis: Für $t \in [a, b]$ definieren wir:

$$f(t) := \frac{|k(t)|}{k(t)} \exp\left(i \int_a^t \operatorname{Im} \frac{dk(s)}{k(s)}\right).$$

Es bezeichne $\langle z, w\rangle = \operatorname{Re}(z\overline{w}) = \operatorname{Re} z \operatorname{Re} w + \operatorname{Im} z \operatorname{Im} w$ das reelle euklidische Skalarprodukt in $\mathbb{C}$. Insbesondere gilt

$$\frac{d}{dt}|k(t)| = \frac{\langle k'(t), k(t)\rangle}{|k(t)|}.$$

Nach der Kettenregel und dem Hauptsatz der Differential- und Integralrechnung erhalten wir:

$$\begin{aligned}\frac{d}{dt} f(t) &= -\frac{k'(t)}{k(t)} f(t) + \frac{\langle k'(t), k(t)\rangle}{|k(t)|^2} f(t) + f(t) i \operatorname{Im} \frac{k'(t)}{k(t)} \\ &= \frac{-k'(t)\overline{k(t)} + \operatorname{Re}(k'(t)\overline{k(t)}) + i \operatorname{Im}(k'(t)\overline{k(t)})}{|k(t)|^2} f(t) = 0,\end{aligned}$$

also ist f konstant. Es folgt $f(a) = f(b)$, also wegen $k(a) = k(b)$ auch

$$1 = \exp\left(i \int_a^b \operatorname{Im} \frac{dk(s)}{k(s)}\right) = \exp\left(i \int_k \operatorname{Im} \frac{dz}{z}\right) = \exp(2\pi i N_0(k)).$$

Da $\exp : \mathbb{C} \to \mathbb{C}^*$ genau auf der Menge $2\pi i\mathbb{Z}$ den Wert 1 annimmt, erhalten wir hieraus $N_0(k) \in \mathbb{Z}$.

□

Umlaufzahlbeweis des Fundamentalsatzes der Algebra. Als Alternative zum "kohomologischen" Beweis des Fundamentalsatzes der Algebra, den wir oben betrachtet haben, kann man auch sozusagen "dual" dazu einen "Umlaufzahlbeweis" angeben: Hierzu betrachtet man die geschlossene Kurve k_t, $t > 0$, die den Kreis um 0 mit Radius t mit Geschwindigkeit 1 im positiven Sinn umläuft. Wäre nun $f \in \mathbb{C}[X]$ ein nichtkonstantes Polynom ohne Nullstellen in $\mathbb{C}$, so liefert $f \circ k_t$, $t > 0$ eine Homotopie von geschlossenen Kurven, die Umlaufzahl 0 des Nullpunkts für kleine t und Umlaufzahl n des Nullpunkts für große t besitzt, ein Widerspruch. Wir führen die Details dieses Beweises hier nicht aus.

Übung 2.140 *Es sei* $f : \mathbb{R}^2 \to \mathbb{R}^2$ *eine glatte Abbildung mit der Eigenschaft*

$$\exists r > 0 \; \forall x \in \mathbb{R}^2 \setminus U_r(0) : \; f(x+z) = x.$$

Zeigen Sie, dass f *surjektiv ist. Sie können dabei mit der Umlaufzahl oder auch mit dem Poincaré-Lemma argumentieren. Lassen Sie sich vom Beweis des Fundamentalsatzes der Algebra inspirieren.*

Übung 2.141 *Es sei* $f : \mathbb{R}^2 \to \mathbb{R}^2$ *eine glatte Abbildung mit der Eigenschaft*

$$\forall x \in \mathbb{R}^2 \; \forall z \in \mathbb{Z}^2 : \; f(x+z) = f(x) + z.$$

Zeigen Sie, dass f *surjektiv ist. Sie können auch hier mit der Umlaufzahl oder mit dem Poincaré-Lemma argumentieren.*

Ausblick: 1-Formen auf Untermannigfaltigkeiten. 1-Formen ω kann man auch auf Untermannigfaltigkeiten statt nur auf offenen Teilmengen von $\mathbb{R}^n$ definieren. Statt dem Dualraum $\omega_x \in (\mathbb{R}^n)'$ verwendet man dann $\omega_x \in (T_xM)'$ (Dualraum des Tangentialraums). Auch geschlossene und exakte 1-Formen und damit den ersten de-Rham Kohomologieraum kann man ganz analog für 1-Formen auf Untermannigfaltigkeiten einführen. Wir führen das hier nicht aus.

Index