

Höhere Mathematik II für die Fachrichtung Physik Sommersemester 2024

Gerd Herzog, Peer Kunstmann, Christoph Schmoeger
Karlsruher Institut für Technologie
Institut für Analysis

Version: July 25, 2024

16 Skalarprodukt und Orthogonalität

16.1. Skalarprodukte: Sei V ein $\mathbb{K}$ -Vektorraum. Eine Abbildung $(\cdot|\cdot) : V \times V \rightarrow \mathbb{K}$ mit den Eigenschaften

- (S1) $\forall x, y \in V: (x|y) = \overline{(y|x)},$
- (S2) $\forall x, y, z \in V \forall \alpha \in \mathbb{K}: (\alpha x + y|z) = \alpha(x|z) + (y|z),$
- (S3) $\forall x \in V \setminus \{0\}: (x|x) > 0.$

heißt ein *Skalarprodukt auf V* .

Eigenschaften eines Skalarproduktes auf einem $\mathbb{K}$ -Vektorraum V sind

$$\begin{aligned} \forall x, y \in V \forall \alpha \in \mathbb{K}: (x|\alpha y + z) &= \overline{\alpha}(x|y) + (x|z), \\ \forall x \in V: (x|0) &= (0|x) = 0, \\ \forall x, y \in V: |(x|y)| &\leq \sqrt{(x|x)}\sqrt{(y|y)} \text{ (Cauchy-Schwarz-Ungleichung, (CSU)).} \end{aligned}$$

Zum Beweis der Cauchy-Schwarz-Ungleichung sei o.B.d.A. $y \neq 0$. Für $\alpha := (x|y)/(y|y)$ gilt

$$0 \leq (x - \alpha y|x - \alpha y) = (x|x) - \alpha(y|x) - \overline{\alpha}(x|y) + |\alpha|^2(y|y) = (x|x) - \frac{|(x|y)|^2}{(y|y)},$$

also $|(x|y)|^2 \leq (x|x)(y|y)$. Wegen (S3) gilt Gleichheit in der Cauchy-Schwarz-Ungleichung nur, wenn $x = \alpha y$, d.h. also nur dann, wenn x, y linear abhängig sind.

Bemerkung: Ist $\mathbb{K} = \mathbb{R}$ und also V ein $\mathbb{R}$ -Vektorraum, so kann auf die komplexe Konjugation verzichtet werden, da $\bar{r} = r$ für alle $r \in \mathbb{R}$.

Beispiele: (1) Das *Standardskalarprodukt* auf dem $\mathbb{K}^n$ (für $\mathbb{K} = \mathbb{R}$ heißt es auch *euklidisch*) ist gegeben durch

$$(\vec{x}|\vec{y}) := \sum_{j=1}^n x_j \overline{y_j} \quad \text{für } \vec{x} = (x_1, x_2, \dots, x_n), \vec{y} = (y_1, y_2, \dots, y_n) \in \mathbb{K}^n.$$

Für dieses Skalarprodukt von $\vec{x}, \vec{y} \in \mathbb{R}^n$ schreibt man auch $\vec{x} \cdot \vec{y}$ (oder $\vec{x}\vec{y}$).

(2) Sind $a_1, a_2, \dots, a_n > 0$, so definiert auch

$$(\vec{x}|\vec{y}) := \sum_{j=1}^n a_j x_j \overline{y_j} \quad \text{für } \vec{x} = (x_1, x_2, \dots, x_n), \vec{y} = (y_1, y_2, \dots, y_n) \in \mathbb{K}^n$$

ein Skalarprodukt auf $\mathbb{K}^n$: (S1) und (S2) sind leicht, für (S3) beachte man, dass

$$(\vec{x}|\vec{x}) = \sum_{j=1}^n a_j |x_j|^2 > 0 \quad (\vec{x} \neq 0).$$

(3) Sei $A \in \mathbb{K}^{n \times n}$ invertierbar und $(\cdot|\cdot)$ das Standardskalarprodukt. Dann definiert

$$\langle \vec{x} | \vec{y} \rangle := (A\vec{x} | A\vec{y})$$

ein Skalarprodukt auf $\mathbb{K}^n$. Auch hier sind (S1) und (S2) leicht. Für (S3) benötigt man die Invertierbarkeit von A .

(4) Durch Grenzübergang von (1) erhält man: Sei $l^2(\mathbb{N})$ der Raum der quadratsummierbaren Folgen in $\mathbb{K}$, also

$$l^2(\mathbb{N}) = \{x = (x_j)_{j \in \mathbb{N}} \in \mathbb{K}^{\mathbb{N}} : \sum_{j=1}^{\infty} |x_j|^2 < \infty\}.$$

Dann ist $l^2(\mathbb{N})$ ein Untervektorraum von $\mathbb{K}^{\mathbb{N}}$ und durch

$$(x|y) := \sum_{j=1}^{\infty} x_j \overline{y_j} \quad (x, y \in l^2(\mathbb{N})),$$

wird ein Skalarprodukt auf $l^2(\mathbb{N})$ definiert: Zunächst ist die Reihe in obiger Definition absolut konvergent. Für jedes $N \in \mathbb{N}$ folgt aus der CSU:

$$\begin{aligned} \sum_{j=1}^N |x_j \overline{y_j}| &= \sum_{j=1}^N |x_j| |y_j| \leq \left(\sum_{j=1}^N |x_j|^2 \right)^{1/2} \left(\sum_{j=1}^N |y_j|^2 \right)^{1/2} \\ &\leq \left(\sum_{j=1}^{\infty} |x_j|^2 \right)^{1/2} \left(\sum_{j=1}^{\infty} |y_j|^2 \right)^{1/2} =: M < \infty, \end{aligned}$$

und $N \rightarrow \infty$ impliziert absolute Konvergenz. Weiter haben wir für $x, y \in l^2(\mathbb{N})$:

$$\sum_{j=1}^N |x_j + y_j|^2 = \sum_{j=1}^N |x_j|^2 + \sum_{j=1}^N |y_j|^2 + \underbrace{\sum_{j=1}^N (x_j \overline{y_j} + \overline{x_j} y_j)}_{|\cdot| \leq 2M},$$

was für $N \rightarrow \infty$ impliziert, dass $x + y \in l^2(\mathbb{N})$ gilt. Die Eigenschaft

$$\alpha x \in l^2(\mathbb{N}) \quad (x \in l^2(\mathbb{N}), \alpha \in \mathbb{K})$$

ist offensichtlich. Die Eigenschaften (S1)-(S3) für $(x|y)$ sind nun leicht nachzurechnen.

Entsprechendes gilt, wenn man statt $\mathbb{N}$ als Indexmenge $\mathbb{Z}$ nimmt und

$$l^2(\mathbb{Z}) := \{x = (x_j)_{j \in \mathbb{Z}} \in \mathbb{K}^{\mathbb{Z}} : \sum_{j=-\infty}^{\infty} |x_j|^2 < \infty\}$$

betrachtet.

Beispiel: Wir betrachten die Folgen

$$x = \left(\frac{1}{j} \right)_{j \in \mathbb{N}}, \quad y = \left(\frac{1}{j+1} \right)_{j \in \mathbb{N}}.$$

Dann gilt $x, y \in l^2(\mathbb{N})$ und x, y sind linear unabhängig. Die CSU liefert nun

$$1 = \sum_{j=1}^{\infty} \frac{1}{j(j+1)} = |(x|y)| < \sqrt{(x|x)}\sqrt{(y|y)} = \left(\sum_{j=1}^{\infty} \frac{1}{j^2} \right)^{1/2} \left(\sum_{j=1}^{\infty} \frac{1}{(j+1)^2} \right)^{1/2}.$$

Mit $s := \sum_{j=1}^{\infty} 1/j^2$ erhalten wir $1 < s(s-1) = s^2 - s$, also

$$0 < s^2 - s - 1 = \left(s - \frac{1 + \sqrt{5}}{2} \right) \left(s - \frac{1 - \sqrt{5}}{2} \right),$$

und wegen $s > 0$ folgt

$$s = \sum_{j=1}^{\infty} \frac{1}{j^2} > \frac{1 + \sqrt{5}}{2}.$$

In 24.5 zeigen wir $s = \pi^2/6$. Obige Ungleichung lautet also

$$1,6449 \approx \frac{\pi^2}{6} > \frac{1 + \sqrt{5}}{2} \approx 1,6180.$$

(5) Seien $a, b \in \mathbb{R}$ mit $a < b$ und $V := C([a, b])$. Dann ist V ein $\mathbb{R}$ -Vektorraum. Durch

$$(f|g) := \int_a^b f(x)g(x) dx \quad (f, g \in C([a, b]))$$

wird auf $C([a, b])$ ein Skalarprodukt definiert. Beachte, dass mit f und g auch fg auf $[a, b]$ stetig und damit integrierbar ist. (S1) und (S2) sind klar. Für $f \in C([a, b])$ ist $(f|f) \geq 0$ klar. Ist $(f|f) = 0$, so folgt $\int_a^x |f(t)|^2 dt = 0$ ($x \in [a, b]$). Da $x \mapsto |f(x)|^2$ auf $[a, b]$ stetig ist, folgt durch Ableiten nach dem Hauptsatz $|f(x)|^2 = 0$ ($x \in [a, b]$), also $f = 0$ auf $[a, b]$. Damit ist (S3) gezeigt.

16.2. Normen: Ist V ein $\mathbb{K}$ -Vektorraum und $(\cdot|\cdot) : V \times V \rightarrow \mathbb{K}$ ein Skalarprodukt, so hat die Abbildung $\|\cdot\| : V \rightarrow [0, \infty)$, $v \mapsto \sqrt{(v|v)}$ folgende Eigenschaften:

(N1) $\forall v \in V: \|v\| = 0 \Rightarrow v = 0$,

(N2) $\forall v \in V \forall \alpha \in \mathbb{K}: \|\alpha v\| = |\alpha| \|v\|$,

(N3) $\forall u, v \in V: \|u + v\| \leq \|u\| + \|v\|$ (Dreiecksungleichung).

Beweis der Dreiecksungleichung. Es gilt:

$$\begin{aligned}\|u + v\|^2 &= (u + v|u + v) = \|u\|^2 + 2\operatorname{Re}(u|v) + \|v\|^2 \\ &\leq \|u\|^2 + 2|(u|v)| + \|v\|^2 \leq_{CSU} \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2.\end{aligned}$$

□

Außerdem gilt die *Parallelogrammgleichung*:

$$\forall u, v \in V : \|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2).$$

Definition: Sei V ein $\mathbb{K}$ -Vektorraum. Eine Abbildung $\|\cdot\| : V \rightarrow [0, \infty)$ mit den Eigenschaften (N1)–(N3) heißt eine *Norm auf V* .

Ist $\|\cdot\|$ eine Norm auf V , so wird für $u, v \in V$ die Zahl $\|u - v\| \geq 0$ als **Abstand** von u und v interpretiert. Es gilt $\|0\| = 0$, sowie

$$|||u\| - \|v\|| \leq \|u - v\| \quad (\text{umgekehrte Dreiecksungleichung}).$$

Beispiele: (1) Ist M eine Menge mit mindestens zwei Punkten, so wird auf dem $\mathbb{R}$ -Vektorraum $B(M)$ aller beschränkten Funktionen $f : M \rightarrow \mathbb{R}$ durch

$$\|f\|_\infty := \sup\{|f(m)| : m \in M\}$$

eine Norm definiert. Zu dieser Norm gibt es kein Skalarprodukt.

(2) Durch $\|(x_1, x_2)\|_1 := |x_1| + |x_2|$ wird auf dem $\mathbb{R}^2$ eine Norm definiert, zu der es kein Skalarprodukt gibt.

In beiden Beispielen ist die Parallelogrammgleichung nicht erfüllt. In (2) ist z.B.

$$\|\vec{e}_1 + \vec{e}_1\|_1^2 + \|\vec{e}_1 - \vec{e}_1\|_1^2 = 8 \neq 4 = 2(\|\vec{e}_1\|_1^2 + \|\vec{e}_2\|_1^2)$$

Bemerkung: Durch eine Norm hat man also einen *Abstandsbegriff* auf einem Vektorraum. Durch ein Skalarprodukt hat man aber außerdem noch die Möglichkeit *Winkel* zu betrachten:

Sind z.B. $\vec{x}, \vec{y} \in \mathbb{R}^n \setminus \{\vec{0}\}$, so ist nach der CSU

$$\frac{(\vec{x}|\vec{y})}{\|\vec{x}\|\|\vec{y}\|} \in [-1, 1], \text{ also } (\vec{x}|\vec{y}) = \|\vec{x}\|\|\vec{y}\| \cos \varphi,$$

für genau ein $\varphi \in [0, \pi]$, das man als den von $\vec{x}, \vec{y}$ eingeschlossenen **Winkel** interpretiert. In $[0, 2\pi]$ gibt es i.a. zwei φ die dieser Gleichung genügen (Winkel zwischen " $\vec{x}$ und $\vec{y}$ " bzw. zwischen " $\vec{y}$ und $\vec{x}$ "). Da das Skalarprodukt in jeder Komponente linear ist, muss man das nur für $\|\vec{x}\| = \|\vec{y}\| = 1$ einsehen. Am besten geht das für $n = 2$ und etwa $\vec{x} = \vec{e}_1$, $\vec{y} = (y_1, y_2)$. Dann ist

$$(\vec{x}|\vec{y}) = (\vec{e}_1|\vec{y}) = y_1.$$

16.3. Orthogonalität: Es sei V ein $\mathbb{K}$ -Vektorraum mit Skalarprodukt $(\cdot|\cdot)$. Vektoren $v_1, v_2, \dots, v_m \in V$ heißen *orthogonal*, falls

$$\forall j, k \in \{1, 2, \dots, m\} : j \neq k \Rightarrow (v_j|v_k) = 0.$$

Statt $(v|w) = 0$ schreibt man auch $v \perp w$.

Die Vektoren $v_1, v_2, \dots, v_m$ heißen *orthonormal* oder ein *Orthonormalsystem* (*ONS*), falls

$$\forall j, k \in \{1, 2, \dots, m\} : (v_j|v_k) = \delta_{jk},$$

d.h. also falls die Vektoren orthogonal sind und zusätzlich alle Norm 1 haben.

Ist V endlich-dimensional, so ist eine *Orthonormalbasis* (*ONB*) von V eine Basis von V , die ein Orthonormalsystem ist.

Beispiele: Die Standardbasis $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ ist eine Orthonormalbasis von $\mathbb{K}^n$, denn es gilt

$$(\vec{e}_j|\vec{e}_k) = \delta_{jk} \quad (j, k = 1, \dots, n).$$

Satz: Sei V ein $\mathbb{K}$ -Vektorraum mit Skalarprodukt und seien $v_1, v_2, \dots, v_m \in V \setminus \{0\}$ orthogonal. Dann sind $v_1, v_2, \dots, v_m$ linear unabhängig.

Beweis. Seien $\alpha_1, \alpha_2, \dots, \alpha_m \in \mathbb{K}$ mit

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_m v_m = 0.$$

Sei $k \in \{1, 2, \dots, m\}$. Wir nehmen das Skalarprodukt der Gleichung mit v_k und erhalten

$$0 = (0|v_k) = (\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_m v_m|v_k) = \sum_{j=1}^m \alpha_j (v_j|v_k) = \alpha_k (v_k|v_k).$$

Wegen $v_k \neq 0$ ist auch $(v_k|v_k) \neq 0$, also $\alpha_k = 0$. Da k beliebig war, sind $v_1, v_2, \dots, v_m$ linear unabhängig. $\square$

Bemerkung: Ist $v_1, v_2, \dots, v_m$ ein Orthonormalsystem in V und $v \in \text{lin}\{v_1, v_2, \dots, v_m\}$, so lassen sich die Koordinaten von v bzgl. $v_1, v_2, \dots, v_m$ leicht bestimmen. Es gilt nämlich

$$v = \sum_{j=1}^m (v|v_j) v_j.$$

Zum Beweis schreibt man $v = \sum_{j=1}^m \alpha_j v_j$ und bildet das Skalarprodukt mit v_k :

$$\forall k \in \{1, \dots, m\} : (v|v_k) = \sum_{j=1}^m \alpha_j \underbrace{(v_j|v_k)}_{=\delta_{jk}} = \alpha_k.$$

Satz (Pythagoras): Sind $v_1, \dots, v_m \in V$ orthogonal, so gilt

$$\|v_1 + \dots + v_m\|^2 = \|v_1\|^2 + \dots + \|v_m\|^2.$$

Z.B. $m = 2$:

$$\|v_1 + v_2\|^2 = (v_1 + v_2 | v_1 + v_2) = (v_1 | v_1) + (v_1 | v_2) + (v_2 | v_1) + (v_2 | v_2) = \|v_1\|^2 + \|v_2\|^2.$$

16.4. Das Gram-Schmidt-Verfahren: Gegeben sei ein $\mathbb{K}$ -Vektorraum V mit Skalarprodukt und $v_1, v_2, \dots, v_m \in V$ seien linear unabhängig. Wir werden ein Orthonormalsystem $b_1, b_2, \dots, b_m$ konstruieren mit

$$\text{lin}\{v_1, v_2, \dots, v_m\} = \text{lin}\{b_1, b_2, \dots, b_m\}.$$

Falls $v_1, v_2, \dots, v_m$ eine Basis von V ist, so ist damit $b_1, b_2, \dots, b_m$ eine Orthonormalbasis von V .

Die Vektoren $b_1, \dots, b_m$ werden sukzessive so konstruiert, dass gilt:

$$\text{lin}\{v_1, v_2, \dots, v_k\} = \text{lin}\{b_1, b_2, \dots, b_k\} \quad (k = 1, \dots, m).$$

Wir setzen $b_1 := v_1 / \|v_1\|$ und für $k = 2, \dots, m$:

$$c_k := v_k - \sum_{j=1}^{k-1} (v_k | b_j) b_j, \quad b_k := \frac{c_k}{\|c_k\|},$$

oder gleichbedeutend

$$c_1 := v_1, \quad c_k := v_k - \sum_{j=1}^{k-1} \frac{(v_k | c_j)}{(c_j | c_j)} c_j \quad (k = 2, \dots, m), \quad b_k := \frac{c_k}{\|c_k\|} \quad (k = 1, \dots, m).$$

Dann ist $b_1, b_2, \dots, b_m$ ein Orthonormalsystem (zum Beweis siehe Bemerkung in 16.7).

Beispiele: (1) Wir betrachten $n = 3$, $V = \mathbb{C}^3$ und

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

Dann ist

$$\begin{aligned}
\vec{b}_1 &= \frac{\vec{v}_1}{\|\vec{v}_1\|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \\
\vec{c}_2 &= \vec{v}_2 - (\vec{v}_2|\vec{b}_1)\vec{b}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{2} \cdot 2 \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\
\vec{b}_2 &= \frac{\vec{c}_2}{\|\vec{c}_2\|} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\
\vec{c}_3 &= \vec{v}_3 - (\vec{v}_3|\vec{b}_1)\vec{b}_1 - (\vec{v}_3|\vec{b}_2)\vec{b}_2 \\
&= \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} - \frac{1}{2} \cdot (-1) \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} - \frac{1}{3} \cdot 0 \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1/2 \\ -1 \\ 1/2 \end{pmatrix} \\
\vec{b}_3 &= \frac{\vec{c}_3}{\|\vec{c}_3\|} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.
\end{aligned}$$

(2) $\vec{v}_1, \vec{v}_2$ wie oben, aber $\vec{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$. Dann sind $\vec{b}_1, \vec{b}_2$ wie oben, aber

$$\begin{aligned}
\vec{c}_3 &= \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{2} \cdot (-1) \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} - \frac{1}{3} \cdot 2 \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1/6 \\ 1/3 \\ -1/6 \end{pmatrix} \\
\vec{b}_3 &= \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}.
\end{aligned}$$

Man beachte, dass sich dieses $\vec{b}_3$ von dem in Beispiel (1) nur durch das Vorzeichen unterscheidet. Dies ist nicht erstaunlich, da es genau zwei Möglichkeiten gibt, $\vec{b}_1, \vec{b}_2$ zu einer Orthonormalbasis zu ergänzen.

Folgerung: Jeder endlichdimensionale Vektorraum V mit Skalarprodukt besitzt eine Orthonormalbasis.

16.5. Transponierte und adjungierte Matrizen: Für eine Matrix $A = (a_{jk}) \in \mathbb{K}^{m \times n}$ heißt die Matrix in $\mathbb{K}^{n \times m}$, die durch Vertauschen von Zeilen und Spalten entsteht, die *transponierte Matrix* zu A und wird mit A^T bezeichnet. Für alle $k \in \{1, \dots, n\}$, $j \in \{1, \dots, m\}$ steht an der Stelle (k, j) in der Matrix A^T also der Eintrag a_{jk} , der in der

Matrix A an der Stelle (j, k) steht. Setzen wir $B := A^T$ mit $B = (b_{kj})_{k=1}^n_{j=1}^m \in \mathbb{K}^{n \times m}$, so gilt also

$$b_{kj} = a_{jk} \quad (k \in \{1, \dots, n\}, j \in \{1, \dots, m\}).$$

Im Falle $\mathbb{K} = \mathbb{C}$ heißt die Matrix aus $\mathbb{C}^{n \times m}$, für die an jeder Stelle (k, j) der Eintrag $\overline{a_{jk}}$ steht, die *adjungierte Matrix* zu A und wird mit A^* bezeichnet.

Bemerkungen: 1. Setzt man $\bar{A} := (\overline{a_{jk}}) \in \mathbb{C}^{m \times n}$ (konjugiert komplexe Matrix zu $A = (a_{jk})$), so gilt also

$$A^* = \overline{A^T} = (\bar{A})^T.$$

2. Ist $A \in \mathbb{R}^{m \times n}$, so gilt

$$A^* = A^T.$$

3. Ist $A \in \mathbb{K}^{m \times n}$, so gilt

$$(A^T)^T = A, \quad (A^*)^* = A.$$

Schreibweisen des Standardskalarprodukts:

Für $\vec{x}, \vec{y} \in \mathbb{R}^n = \mathbb{R}^{n \times 1}$ gilt $(\vec{x}|\vec{y}) = \vec{y}^T \vec{x} = \vec{x}^T \vec{y}$.

Für $\vec{x}, \vec{y} \in \mathbb{C}^n = \mathbb{C}^{n \times 1}$ gilt $(\vec{x}|\vec{y}) = \vec{y}^* \vec{x} = \overline{\vec{x}^* \vec{y}}$.

Rechenregeln: Für Matrizen A, B , deren Produkt definiert ist, gilt:

$$(AB)^T = B^T A^T \quad \text{und} \quad (AB)^* = B^* A^*.$$

Für eine invertierbare Matrix A sind auch A^T und A^* invertierbar, und es gilt

$$(A^T)^{-1} = (A^{-1})^T, \quad (A^*)^{-1} = (A^{-1})^*.$$

Wende dazu die Rechenregeln auf $B = A^{-1}$ an und beachte $I^T = I = I^*$, wobei I die jeweilige Einheitsmatrix sei.

Folgerung: Sei $A \in \mathbb{K}^{m \times n}$. Dann gilt:

(a) Im Fall $\mathbb{K} = \mathbb{R}$ ist $(A\vec{x}|\vec{y}) = (\vec{x}|A^T \vec{y})$ für alle $\vec{x} \in \mathbb{R}^n, \vec{y} \in \mathbb{R}^m$.

(b) Im Fall $\mathbb{K} = \mathbb{C}$ ist $(A\vec{x}|\vec{y}) = (\vec{x}|A^* \vec{y})$ für alle $\vec{x} \in \mathbb{C}^n, \vec{y} \in \mathbb{C}^m$.

Beweis. Man schreibe das Skalarprodukt wie oben angegeben und benutze die Rechenregeln. Z.B. (b):

$$(\vec{x}|A^* \vec{y}) = (A^* \vec{y})^* \vec{x} = (\vec{y}^* A) \vec{x} = \vec{y}^* (A \vec{x}) = (A \vec{x}|\vec{y}).$$

□

16.6. Orthogonale und unitäre Matrizen: Eine wichtige Rolle spielen Matrizen $A \in \mathbb{K}^{n \times n}$, deren zugehörige lineare Abbildung $\mathbb{K}^n \rightarrow \mathbb{K}^n$, $\vec{x} \mapsto A\vec{x}$, das Skalarprodukt invariant lässt, d.h. für die gilt:

$$\forall \vec{x}, \vec{y} \in \mathbb{K}^n : (A\vec{x} | A\vec{y}) = (\vec{x} | \vec{y}).$$

Damit verändert A auch *Winkel* und *Abstände* nicht. Eine solche Matrix A hat $\text{Kern } A = \{\vec{0}\}$, ist also invertierbar. Aus den Rechenregeln folgt

$$A^T A = I_n \quad (\text{für } \mathbb{K} = \mathbb{R}), \quad A^* A = I_n \quad (\text{für } \mathbb{K} = \mathbb{C}).$$

Eine Matrix $A \in \mathbb{K}^{n \times n}$ mit dieser Eigenschaft heißt *orthogonal* (für $\mathbb{K} = \mathbb{R}$) bzw. *unitär* (für $\mathbb{K} = \mathbb{C}$), also

$$A \in \mathbb{R}^{n \times n} \text{ ist orthogonal} : \iff A^T = A^{-1},$$

$$A \in \mathbb{C}^{n \times n} \text{ ist unitär} : \iff A^* = A^{-1}.$$

Bemerkung: Sei $A \in \mathbb{C}^{n \times n}$.

- 1) A ist genau dann unitär, wenn die Spalten von A eine Orthonormalbasis von $\mathbb{C}^n$ bilden.
- 2) A ist genau dann unitär, wenn für jedes Orthonormalsystem $v_1, \dots, v_m$ in $\mathbb{C}^n$ auch $Av_1, \dots, Av_m$ ein Orthonormalsystem von $\mathbb{C}^n$ ist.
- 3) Produkte, Inverse, Transponierte und Adjungierte von unitären Matrizen sind unitär.

Beispiele: 1) Spiegelungen in $\mathbb{C}^n$, etwa $A\vec{e}_1 = -\vec{e}_1$, $A\vec{e}_j = \vec{e}_j$ für $j = 2, \dots, n$.

2) Rotation in $\mathbb{R}^2$ um den Winkel $\theta \in \mathbb{R}$ (im mathematisch positiven Sinn):

$$A = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

3) Im $\mathbb{R}^3$ Rotation um die z -Achse bei Spiegelung an der (x, y) -Ebene:

$$A = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

16.7. Orthogonalprojektionen: Es sei V ein $\mathbb{K}$ -Vektorraum mit Skalarprodukt $(\cdot | \cdot)$, $b_1, b_2, \dots, b_m$ ein Orthonormalsystem in V und $U := \text{lin}\{b_1, b_2, \dots, b_m\}$. Die lineare Abbildung

$$P : V \rightarrow U, \quad v \mapsto Pv = \sum_{j=1}^m (v | b_j) b_j$$

hat folgende Eigenschaften:

$$P \circ P = P, \quad \text{Bild } P = U, \quad \text{Kern } P = U^\perp := \{v \in V : v \perp u \text{ für alle } u \in U\}.$$

Es gilt $(v - Pv|u) = 0$ ($u \in U, v \in V$), und

$$\|v - Pv\| = \min\{\|v - u\| : u \in U\} \quad (v \in V),$$

d.h. Pv ist die (eindeutig bestimmte) *Bestapproximation* von v in U . Die Abbildung P heit *Orthogonalprojektion* von V auf U .

Beweis. Klar ist $\text{Bild } P \subseteq U$. Nach der Bemerkung in 16.3 gilt $Pu = u$ ($u \in U$), also ist $P \circ P = P$ und $\text{Bild } P = U$. Fr jedes $v \in V$ gilt:

$$Pv = 0 \iff (v|b_j) = 0 \quad (j = 1, \dots, m) \iff (v|u) = 0 \quad (u \in U).$$

Auerdem ist fr $k = 1, \dots, m$:

$$(v - Pv|b_k) = (v|b_k) - (Pv|b_k) = (v|b_k) - \sum_{j=1}^m (v|b_j) \underbrace{(b_j|b_k)}_{=\delta_{jk}} = (v|b_k) - (v|b_k) = 0,$$

und folglich auch $(v - Pv|u) = 0$ ($u \in U$). Insbesondere ist $(v - Pv|Pv - u) = 0$ ($u \in U$). Damit folgt mit de Satz von Pythagoras fr $u \in U$:

$$\|v - u\|^2 = \|v - Pv + Pv - u\|^2 = \|v - Pv\|^2 + \|Pv - u\|^2.$$

Hieran sieht man, dass Pv die eindeutige Bestapproximation von v in U ist. □

Bemerkung: Im Gram-Schmidt-Verfahren in 16.4 hat man im k -ten Schritt (fr $k = 2, \dots, m$) und fr $U = \text{lin}\{b_1, b_2, \dots, b_{k-1}\}$, dass $c_k = v_k - Pv_k$ und somit $c_k \perp u$ ($u \in U$) wie gewnscht.

16.8. Gruppen und Matrixgruppen:

Definition: Eine *Gruppe* ist eine nicht-leere Menge G , versehen mit einer Verknpfung $\circ : G \times G \rightarrow G$, geschrieben $a \circ b$ (bzw. meist nur ab), mit folgenden Eigenschaften:

- (G1) $\forall a, b, c \in G : (ab)c = a(bc)$ (Assoziativitt),
- (G2) $\exists e \in G \forall a \in G : ae = ea = a$ (Existenz des neutralen Elements),
- (G3) $\forall a \in G \exists b \in G : ab = ba = e$ (Existenz inverser Elemente).

Das neutrale Element e und inverse Elemente sind eindeutig bestimmt: Schreibweise a^{-1} , also $aa^{-1} = a^{-1}a = e$.

Eine Gruppe $(G, \circ)$ mit kommutativer Verknpfung $\circ$, d.h. mit $ab = ba$ ($a, b \in G$), heit *abelsch*.

Bemerkung: In einer Gruppe G lassen sich die Gleichungen $ax = b$ und $xa = b$ (für beliebig gegebene $a, b \in G$) eindeutig lösen:

$$\begin{aligned} ax = b &\iff x = ex = a^{-1}ax = a^{-1}b \\ xa = b &\iff x = xe = xaa^{-1} = ba^{-1}. \end{aligned}$$

Beispiele: 1.) $(\mathbb{Z}, +)$ ist eine abelsche Gruppe, neutrales Element ist 0.

2) $(\mathbb{R}, +)$ und $(\mathbb{C}, +)$ sind abelsche Gruppen. Ist V ein Vektorraum, so ist $(V, +)$ eine abelsche Gruppe. Neutrales Element ist die jeweilige 0.

3) Die Menge

$$S := \{x \in \mathbb{R} : \lim_{n \rightarrow \infty} \sin(n! \pi x) = 0\}$$

ist eine abelsche Gruppe bzgl. „+“ (also eine *Untergruppe* von $(\mathbb{R}, +)$). Es ist nicht leicht zu sehen welche Zahlen zu dieser Gruppe gehören. Z.B. gilt $\mathbb{Q} \cup \{e\} \subseteq S \neq \mathbb{R}$. Achtung: Das neutrale Element ist 0, und e ist hier die Eulersche Zahl.

4) $(\mathbb{K} \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe, neutrales Element ist 1.

5) Bzgl. der Matrixmultiplikation ist

$$GL(n, \mathbb{K}) := \{A \in \mathbb{K}^{n \times n} : A \text{ ist invertierbar}\}$$

eine Gruppe, die *allgemeine (generelle) lineare Gruppe*, die für $n \geq 2$ nicht abelsch ist. Neutrales Element ist die Einheitsmatrix I_n .

In $GL(n, \mathbb{K})$ gibt es viele kleinere Gruppen (also Untergruppen). Neutrales Element ist dabei stets die Einheitsmatrix I_n .

6) $O(n) := \{A \in \mathbb{R}^{n \times n} : A \text{ ist orthogonal}\}$ ist eine Gruppe, die *orthogonale Gruppe*.

7) $U(n) := \{A \in \mathbb{C}^{n \times n} : A \text{ ist unitär}\}$ ist eine Gruppe, die *unitäre Gruppe*.

8) Die Teilmenge

$$\left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} : a, b \in \mathbb{R}, |a| + |b| > 0 \right\} \subseteq \mathbb{R}^{2 \times 2}$$

ist eine abelsche Gruppe, die $(\mathbb{C} \setminus \{0\}, \cdot)$ entspricht, wenn man

$$a + ib \quad \text{mit} \quad \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

identifiziert.

9) Die Menge

$$\{A \in O(2) : A\vec{e}_1 = \vec{e}_1\}$$

ist eine abelsche Gruppe mit zwei Elementen.

10) Sei $M := \{(x_1, x_2) \in \mathbb{R}^2 : |x_1| + |x_2| \leq 1\}$. Die Menge

$$\{A \in O(2) : A(M) = M\}$$

ist eine Gruppe mit acht Elementen (Drehungen um $0, \pi/2, \pi, 3\pi/2$, sowie die Spiegelungen an den 4 Symmetrieachsen). Diese Gruppe ist nicht abelsch.

17 Determinanten und Kreuzprodukt

Idee der Determinante: Zu gegebenen Vektoren

$$\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n \in \mathbb{R}^n$$

gibt $\det(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n)$ das mit einem Vorzeichen versehene Volumen des von den Vektoren $\vec{a}_1, \dots, \vec{a}_n$ aufgespannten *Spates*

$$\{\lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 + \dots + \lambda_n \vec{a}_n : \lambda_j \in [0, 1] \ (j = 1, \dots, n)\}$$

an. “Mit Vorzeichen versehen” bedeutet dabei, dass $\det(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n)$ das Vorzeichen wechselt, wenn man einen Vektor $\vec{a}_j$ durch $-\vec{a}_j$ ersetzt. Bei $\det(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n)$ stellen wir uns die Vektoren $\vec{a}_1, \dots, \vec{a}_n$ als *Spalten* einer *Matrix* vor.

Beispiel: Der von den Einheitsvektoren $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ im $\mathbb{R}^n$ aufgespannte Spat ist der *Einheitswürfel*

$$Q := [0, 1]^n.$$

Dieser hat Volumen = 1 (Produkt der Seitenlängen).

17.1. Definierende Eigenschaften der Determinante: Die Determinante ist eine Abbildung $\det : \underbrace{\mathbb{K}^n \times \mathbb{K}^n \times \dots \times \mathbb{K}^n}_n \rightarrow \mathbb{K}$ mit den Eigenschaften

$$(D1) \det(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n) = 1,$$

$$(D2) \text{ für alle } j \in \{1, \dots, n\}, \vec{a}_1, \dots, \vec{a}_n, \vec{b}_j \in \mathbb{K}^n \text{ und } \alpha, \beta \in \mathbb{K} \text{ gilt}$$

$$\begin{aligned} & \det(\vec{a}_1, \dots, \alpha \vec{a}_j + \beta \vec{b}_j, \dots, \vec{a}_n) \\ &= \alpha \det(\vec{a}_1, \dots, \vec{a}_j, \dots, \vec{a}_n) + \beta \det(\vec{a}_1, \dots, \vec{b}_j, \dots, \vec{a}_n), \end{aligned}$$

$$(D3) \det(\vec{a}_1, \dots, \vec{a}_n) = 0, \text{ falls es } j \neq k \text{ gibt mit } \vec{a}_j = \vec{a}_k.$$

Bemerkung: Durch die Eigenschaften (D1)–(D3) ist die Determinante $\det$ eindeutig bestimmt (ohne Beweis). (D1) bedeutet eine Normierung (vgl. Beispiel oben). (D2) bedeutet, dass die Determinante in jeder Spalte *linear* ist. Zusammen mit (D3) bedeutet (D2), dass $\det$ eine *alternierende Multilinearform* ist, d.h. Vertauschen zweier Spalten ändert das Vorzeichen.

Beachte außerdem, dass (D2) und (D3) die Idee eines mit Vorzeichen versehenen Volumens umsetzen: Bei (D3) ist der aufgespannte Spat “flach”, hat also kein Volumen; (D2) kann man sich mit Streckung und Scherung veranschaulichen.

Schreibweise: Ist $A \in \mathbb{K}^{n \times n}$ die Matrix mit den Spalten $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n \in \mathbb{K}^n$, so schreibt man auch

$$|A| := \det(A) := \det A := \det(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n).$$

Wir betrachten im folgenden $\det$ meist als Funktion auf $\mathbb{K}^{n \times n}$, d.h. auf dem Raum der quadratischen Matrizen.

17.2. Folgerungen: (a) Ist eine Spalte $= \vec{0}$, so ist auch die Determinante $= 0$.

(b) Man kann zu einer Spalte ein Vielfaches einer anderen Spalte dazuaddieren, ohne den Wert der Determinante zu ändern.

(c) Vertauscht man zwei Spalten miteinander, so ändert sich das Vorzeichen der Determinante.

(d) Sind die Spalten von $A \in \mathbb{K}^{n \times n}$ linear abhängig (d.h. gilt $\text{Rang } A < n$), so ist

$$\det(A) = 0.$$

(e) Es gilt: $\det(A) \neq 0 \Leftrightarrow A$ ist regulär.

Erinnerung: Eine Matrix $\in \mathbb{K}^{n \times n}$ ist regulär, genau dann wenn sie invertierbar ist, bzw. falls die zugehörige lineare Abbildung $\mathbb{K}^n \rightarrow \mathbb{K}^n, \vec{x} \mapsto A\vec{x}$ bijektiv (oder injektiv oder surjektiv) ist (vgl. 15.16 in HM I).

Der Rang einer Matrix ist die Maximalzahl linear unabhängiger Spalten (oder die Maximalzahl linear unabhängiger Zeilen).

Beweis. (a) folgt sofort aus (D2). (b) folgt leicht aus (D2) und (D3).

Zu (c): Wegen (b) und (D2) ist

$$\begin{aligned} & \det(\vec{a}_1, \dots, \underbrace{\vec{a}_j}_k, \dots, \underbrace{\vec{a}_k}_j, \dots, \vec{a}_n) \\ &= \det(\vec{a}_1, \dots, \underbrace{\vec{a}_j + \vec{a}_k}_k, \dots, \underbrace{\vec{a}_k}_j, \dots, \vec{a}_n) \\ &= \det(\vec{a}_1, \dots, \underbrace{\vec{a}_j + \vec{a}_k}_k, \dots, \underbrace{-\vec{a}_j}_j, \dots, \vec{a}_n) \\ &= \det(\vec{a}_1, \dots, \underbrace{\vec{a}_k}_k, \dots, \underbrace{-\vec{a}_j}_j, \dots, \vec{a}_n) \\ &= -\det(\vec{a}_1, \dots, \vec{a}_n). \end{aligned}$$

Zu (d): Es sei etwa die letzte Spalte Linearkombination der anderen Spalten. Wegen (D2) und (D3) gilt dann:

$$\det(\vec{a}_1, \dots, \vec{a}_{n-1}, \sum_{j=1}^{n-1} \alpha_j \vec{a}_j) = \sum_{j=1}^{n-1} \alpha_j \det(\vec{a}_1, \dots, \vec{a}_{n-1}, \vec{a}_j) = 0.$$

Wegen (d) muss man bei (e) nur noch zeigen: $\text{Rang } A = n$ impliziert $\det(A) \neq 0$. Dazu bringen wir A durch elementare Spaltenumformungen (analog zu Zeilenumformungen, nur für Spalten statt für Zeilen) auf die Gestalt der Einheitsmatrix I_n . Dabei wird nach (D2) (für $\vec{b}_j = 0$) und (b) und (c) in jedem Umformungsschritt $\det(A)$ nur mit Zahlen $\neq 0$ multipliziert. Wegen (D1) ist schließlich für ein $c \neq 0$

$$\det(A) = c \det(I_n) = c \neq 0.$$

□

17.3. Der Fall $n = 1$: Es gilt $\det(a) = a$ ($a \in \mathbb{K}$). (D1), (D2) und (D3) sind klar.

Der Fall $n = 2$: Es gilt

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc \quad (\text{das ist das } \delta \text{ aus Beispiel 15.16, HM I}),$$

denn die Eigenschaften (D1) und (D3) sind klar, und (D2) ist leicht.

Beispiele: (1)

$$\begin{vmatrix} 2 & -3 \\ -4 & 6 \end{vmatrix} = 2 \cdot 6 - (-3) \cdot (-4) = 0,$$

die Matrix ist nicht regulär.

(2)

$$\begin{vmatrix} i & -4 \\ 0 & -1 \end{vmatrix} = i \cdot (-1) - (-4) \cdot 0 = -i,$$

die Matrix ist regulär.

17.4. Der Fall $n = 3$: Es gilt

$$\begin{vmatrix} a & b & c \\ u & v & w \\ x & y & z \end{vmatrix} = a \begin{vmatrix} v & w \\ y & z \end{vmatrix} - b \begin{vmatrix} u & w \\ x & z \end{vmatrix} + c \begin{vmatrix} u & v \\ x & y \end{vmatrix} = avz + bwx + cuy - awy - buz - cvx.$$

(D1) ist klar, (D2) ist leicht. (D3) kann man nachrechnen.

Die *Regel von Sarrus* gilt für $n = 3$, **aber nicht für $n \geq 4$!**

Schema:

$$\begin{array}{ccccccc} & & & a & b & c & \\ & & w & \left| \begin{array}{ccc} u & v & w \end{array} \right| & u & & \\ & y & z & \left| \begin{array}{ccc} x & y & z \end{array} \right| & x & y & \\ \swarrow & \swarrow & \swarrow & & \searrow & \searrow & \searrow \\ - & - & - & & + & + & + \end{array}$$

Beispiel:

$$\begin{vmatrix} 1 & 0 & 0 \\ 4 & 2 & 0 \\ 6 & 5 & 3 \end{vmatrix} = 1 \cdot 2 \cdot 3 + 0 \cdot 0 \cdot 6 + 0 \cdot 4 \cdot 5 - 0 \cdot 2 \cdot 6 - 0 \cdot 4 \cdot 3 - 1 \cdot 0 \cdot 5 = 6.$$

17.5. Der allgemeine Fall: Gegeben sei die Matrix $A = (a_{jk})_{jk} \in \mathbb{K}^{n \times n}$. Für jedes $k \in \{1, \dots, n\}$ bezeichne $A_{1k} \in \mathbb{K}^{(n-1) \times (n-1)}$ diejenige Matrix, die aus A durch Streichen der ersten Zeile und der k -ten Spalte entsteht.

Ohne Beweis: Man hat die folgende Formel, die das Berechnen von $\det(A)$ auf das Berechnen der Determinanten kleinerer Matrizen zurückführt:

$$\det(A) = \sum_{k=1}^n (-1)^{k+1} a_{1k} \det(A_{1k}).$$

Für $n = 3$ steht hier gerade die Formel aus 17.4, für $n = 2$ diejenige aus 17.3 (und man kann induktiv (D1), (D2), (D3) nachweisen).

Beispiel: Ist

$$A = \begin{pmatrix} d_1 & 0 & 0 & \cdots & 0 \\ * & d_2 & 0 & \cdots & 0 \\ * & * & \ddots & \ddots & 0 \\ * & * & \cdots & d_{n-1} & 0 \\ * & * & \cdots & * & d_n \end{pmatrix} \in \mathbb{K}^{n \times n},$$

also eine *untere Dreiecksmatrix*, so gilt $\det(A) = d_1 \cdot d_2 \cdot \dots \cdot d_n$.

17.6. Determinantenentwicklungssatz: Die Formel in 17.5 nennt man *Entwicklung von $\det(A)$ nach der ersten Zeile von $A = (a_{jk}) \in \mathbb{K}^{n \times n}$* . Ebenso kann man $\det(A)$ nach einer beliebigen Zeile oder Spalte entwickeln:

Für jedes $l \in \{1, \dots, n\}$ gilt

$$\det(A) = \sum_{k=1}^n (-1)^{k+l} a_{lk} \det(A_{lk}) = \sum_{j=1}^n (-1)^{j+l} a_{jl} \det(A_{jl}),$$

wobei $A_{jk} \in \mathbb{K}^{(n-1) \times (n-1)}$ die Matrix bezeichne, die aus A durch Streichen der j -ten Zeile und der k -ten Spalte entsteht.

Beispiel: Man entwickelt möglichst nach einer Zeile oder Spalte mit vielen Nullen, hier z.B. nach der zweiten Spalte:

$$\begin{vmatrix} 2 & 0 & 3 \\ 4 & 2 & 5 \\ 6 & 0 & 7 \end{vmatrix} = -0 \cdot \begin{vmatrix} 4 & 5 \\ 6 & 7 \end{vmatrix} + 2 \cdot \begin{vmatrix} 2 & 3 \\ 6 & 7 \end{vmatrix} - 0 \cdot \begin{vmatrix} 2 & 3 \\ 4 & 5 \end{vmatrix} = 2 \cdot (2 \cdot 7 - 3 \cdot 6) = -8.$$

17.7. Das Signum einer Permutation: Eine *Permutation* ist eine bijektive Abbildung $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$. Die Menge S_n aller Permutationen von $\{1, \dots, n\}$ hat genau $n!$ Elemente, und bzgl. Komposition “ $\circ$ ” ist $(S_n, \circ)$ eine Gruppe.

Eine Permutation $\sigma \in S_n$ schreibt man oft in der Form

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$

oder auch nur $(\sigma(1), \sigma(2), \dots, \sigma(n))$.

Für eine Permutation $\sigma \in S_n$ definiert man das *Signum* (oder *Vorzeichen*) $\text{sgn } \sigma$ von σ durch

$$\text{sgn } \sigma := \prod_{i < j} \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Das Produkt hat genau $\binom{n}{2} = \frac{n(n-1)}{2}$ Faktoren.

Bemerkung: Es gilt stets $\text{sgn } \sigma \in \{1, -1\}$, genauer ist $\text{sgn } \sigma = (-1)^m$, wobei m die Anzahl der Paare (i, j) mit $i, j \in \{1, \dots, n\}$ und $i < j$, aber $\sigma(i) > \sigma(j)$ ist.

Beispiel: Sei $n = 4$ und

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}.$$

Hier ist $\text{sgn } \sigma = -1$, da die Paare (i, j) mit $i < j$ und $\sigma(i) > \sigma(j)$ gerade $(1, 4), (2, 4), (3, 4)$ sind.

Bemerkung: Es gilt:

$$\forall \sigma, \tau \in S_n : \text{sgn } (\sigma \circ \tau) = (\text{sgn } \sigma) \cdot (\text{sgn } \tau).$$

Insbesondere ist damit $\text{sgn } (\sigma^{-1}) = \text{sgn } \sigma$ ($\sigma \in S_n$).

17.8. Die Leibnizformel für Determinanten: Es sei $A = (a_{jk}) \in \mathbb{K}^{n \times n}$. Dann gilt:

$$\det(A) = \sum_{\sigma \in S_n} \text{sgn } \sigma \cdot a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

(ohne Beweis).

Folgerung: Für $A \in \mathbb{K}^{n \times n}$ gilt $\det(A^T) = \det(A)$.

Beweis. Unter Beachtung von $\{\sigma^{-1} : \sigma \in S_n\} = S_n$ gilt

$$\begin{aligned} \det(A) &= \sum_{\sigma \in S_n} \text{sgn } \sigma \cdot a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} = \sum_{\sigma \in S_n} \text{sgn } \sigma \cdot a_{\sigma^{-1}(1)1} a_{\sigma^{-1}(2)2} \cdots a_{\sigma^{-1}(n)n} \\ &= \sum_{\sigma \in S_n} \text{sgn } (\sigma^{-1}) \cdot a_{\sigma^{-1}(1)1} a_{\sigma^{-1}(2)2} \cdots a_{\sigma^{-1}(n)n} = \sum_{\sigma \in S_n} \text{sgn } \sigma \cdot a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} = \det(A^T). \end{aligned}$$

□

Beispiel: Ist $A \in \mathbb{K}^{n \times n}$ von der Form

$$A = \begin{pmatrix} d_1 & * & * & \cdots & * \\ 0 & d_2 & * & \cdots & * \\ 0 & 0 & \ddots & \ddots & * \\ 0 & 0 & \cdots & d_{n-1} & * \\ 0 & 0 & \cdots & 0 & d_n \end{pmatrix} \in \mathbb{K}^{n \times n},$$

also eine *obere Dreiecksmatrix*, so gilt $\det(A) = d_1 \cdot d_2 \cdot \dots \cdot d_n$.

17.9. Determinantenmultiplikationssatz:

$$\forall A, B \in \mathbb{K}^{n \times n} : \det(AB) = \det(A) \det(B).$$

(ohne Beweis).

Folgerung: Ist A regulär, so gilt: $\det(A^{-1}) = 1/\det(A)$.

17.10. Die Cramersche Regel für lineare Gleichungssysteme: Gegeben sei ein lineares Gleichungssystem $A\vec{x} = \vec{b}$, wobei $\vec{b} \in \mathbb{K}^n$ sei und $A \in \mathbb{K}^{n \times n}$ die Spalten $\vec{a}_1, \dots, \vec{a}_n \in \mathbb{K}^n$ habe. Ist A regulär, so hat das Gleichungssystem eine eindeutige Lösung $\vec{x} = (x_1, x_2, \dots, x_n)^T \in \mathbb{K}^n$. Es gilt:

$$\forall j \in \{1, \dots, n\} : x_j = \det(\vec{a}_1, \dots, \underbrace{\vec{b}}_j, \dots, \vec{a}_n) / \det(A).$$

Zur Berechnung der j -ten Komponente x_j der Lösung muss man also die j -te Spalte von A durch den Vektor $\vec{b}$ ersetzen, die Determinante berechnen und durch die Determinante von A dividieren.

Beweis. Es gilt $\vec{b} = \sum_{k=1}^n x_k \vec{a}_k$, also ist wegen (D2) und (D3):

$$\det(\vec{a}_1, \dots, \underbrace{\vec{b}}_j, \dots, \vec{a}_n) = \sum_{k=1}^n x_k \det(\vec{a}_1, \dots, \underbrace{\vec{a}_k}_j, \dots, \vec{a}_n) = x_j \det(A).$$

□

17.11. Eine Formel für die inverse Matrix: Sei $A \in \mathbb{K}^{n \times n}$ regulär mit Spalten $\vec{a}_1, \dots, \vec{a}_n \in \mathbb{K}^n$. Geht man zur Berechnung von A^{-1} wie in 15.16 (HM I) vor und verwendet die Cramersche Regel 17.10, so erhält man

$$A^{-1} = \frac{1}{\det(A)} \left(\det(\vec{a}_1, \dots, \underbrace{\vec{e}_k}_j, \dots, \vec{a}_n) \right)_{j,k=1}^n.$$

Die Formel für $n = 2$ in Beispiel 15.16 (HM I) ist ein Spezialfall.

17.12. Orientierung: Ist V ein n -dimensionaler reeller Vektorraum und $\phi : V \rightarrow V$ eine lineare Abbildung, so kann man ϕ eine Determinante $\det \phi$ zuordnen, indem man eine Basis $b_1, b_2, \dots, b_n$ von V wählt, die Abbildung ϕ bzgl. dieser Basis durch eine Matrix A darstellt und $\det \phi := \det(A)$ setzt. Diese Definition hängt nicht von der Basis ab (ist also sinnvoll):

Ist $c_1, \dots, c_n$ eine weitere Basis von V , so erhalten wir die Darstellungsmatrix $\tilde{A}$ von ϕ bzgl. dieser Basis als $\tilde{A} = S^{-1}AS$, wobei S die Darstellungsmatrix der Identität $\text{Id}_V : V \rightarrow V$ ist, wenn man im Argumentraum die Basis $c_1, \dots, c_n$ und im Zielraum die Basis $b_1, \dots, b_n$ nimmt. Wegen 17.9 ist dann

$$\det(\tilde{A}) = \det(S^{-1})\det(A)\det(S) = \det(S)^{-1}\det(A)\det(S) = \det(A).$$

Definition: Eine bijektive lineare Abbildung $\phi : V \rightarrow V$ heißt *orientierungstreu*, falls $\det \phi > 0$ ist, und *orientierungsumkehrend*, falls $\det \phi < 0$ ist.

Eine geordnete Basis $\vec{b}_1, \vec{b}_2, \vec{b}_3$ von $\mathbb{R}^3$ heißt *Rechtssystem*, falls $\det(\vec{b}_1, \vec{b}_2, \vec{b}_3) > 0$ ist. Meist ist dabei $\vec{b}_1, \vec{b}_2, \vec{b}_3$ eine Orthonormalbasis.

Satz: Ist $\vec{b}_1, \vec{b}_2, \vec{b}_3$ ein Rechtssystem in $\mathbb{R}^3$ und $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ orientierungstreu, so ist auch $\phi(\vec{b}_1), \phi(\vec{b}_2), \phi(\vec{b}_3)$ ein Rechtssystem.

Beweis. Ist A die Darstellungsmatrix von ϕ bzgl. $\vec{b}_1, \vec{b}_2, \vec{b}_3$, so ist

$$\det(\phi(\vec{b}_1), \phi(\vec{b}_2), \phi(\vec{b}_3)) = \det\left((\vec{b}_1, \vec{b}_2, \vec{b}_3) \cdot A\right) = \det(\vec{b}_1, \vec{b}_2, \vec{b}_3) \cdot \det(A) > 0.$$

□

Beispiel: $\vec{e}_1, \vec{e}_2, \vec{e}_3$ ist ein Rechtssystem, $\vec{e}_1, \vec{e}_3, \vec{e}_2$ ist kein Rechtssystem.

17.13. Das Kreuzprodukt (Vektorprodukt) im $\mathbb{R}^3$: Für zwei Vektoren

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \in \mathbb{R}^3$$

ist das *Kreuzprodukt* $\vec{x} \times \vec{y} \in \mathbb{R}^3$ definiert durch

$$\vec{x} \times \vec{y} := \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}.$$

Der Vektor $\vec{x} \times \vec{y} \in \mathbb{R}^3$ ist senkrecht zu $\vec{x}$ und $\vec{y}$, seine Länge ist der Flächeninhalt des von $\vec{x}$ und $\vec{y}$ aufgespannten Parallelogramms, und es gilt $\det(\vec{x}, \vec{y}, \vec{x} \times \vec{y}) \geq 0$.

Es gelten folgende **Rechenregeln**:

$$(1) \vec{x} \times \vec{y} = -\vec{y} \times \vec{x}.$$

(2) $\vec{e}_1 \times \vec{e}_2 = \vec{e}_3$, allgemeiner $\vec{e}_{\sigma(1)} \times \vec{e}_{\sigma(2)} = \operatorname{sgn} \sigma \vec{e}_{\sigma(3)}$ für jede Permutation $\sigma \in S_3$ (denn $\det(\vec{e}_{\sigma(1)}, \vec{e}_{\sigma(2)}, \vec{e}_{\sigma(3)}) = \operatorname{sgn} \sigma$).

(3)

$$(\alpha \vec{x} + \beta \vec{w}) \times \vec{y} = \alpha(\vec{x} \times \vec{y}) + \beta(\vec{w} \times \vec{y})$$

und

$$\vec{x} \times (\alpha \vec{y} + \beta \vec{z}) = \alpha(\vec{x} \times \vec{y}) + \beta(\vec{x} \times \vec{z})$$

für alle $\alpha, \beta \in \mathbb{R}$ und $\vec{x}, \vec{w}, \vec{y}, \vec{z} \in \mathbb{R}^3$, d.h. das Kreuzprodukt ist linear in jeder Komponente.

(4) Für alle $\vec{x}, \vec{y} \in \mathbb{R}^3$ und $\alpha \in \mathbb{R}$ gilt

$$\vec{x} \times \vec{y} = \vec{x} \times (\vec{y} + \alpha \vec{x}) = (\vec{x} + \alpha \vec{x}) \times \vec{y},$$

d.h. man kann zu einer Variablen ein Vielfaches der anderen dazuaddieren.

(5) Es gilt $\vec{x} \times \vec{y} = \vec{0}$ genau dann, wenn $\vec{x}, \vec{y}$ linear abhängig sind.

Warnung: Das Kreuzprodukt ist nicht assoziativ! So ist z.B.

$$(\vec{e}_1 \times \vec{e}_2) \times \vec{e}_2 = \vec{e}_3 \times \vec{e}_2 = -\vec{e}_1, \quad \vec{e}_1 \times (\vec{e}_2 \times \vec{e}_2) = \vec{e}_1 \times \vec{0} = \vec{0}.$$

Anwendung Lorentzkraft: Eine in einem Magnetfeld der Flussdichte $\vec{B}$ mit der Geschwindigkeit $\vec{v}$ bewegte elektrische Ladung q erfährt die *Lorentzkraft*

$$\vec{F} = q(\vec{v} \times \vec{B})$$

und wird dadurch abgelenkt. Das durch $\vec{v}$ und $\vec{B}$ aufgespannte Parallelogramm hat den Flächeninhalt $\|\vec{v}\| \|\vec{B}\| \sin \varphi$, wobei φ der Winkel zwischen $\vec{v}$ und $\vec{B}$ ist. Es gilt also

$$\|\vec{F}\| = |q| \|\vec{v} \times \vec{B}\| = |q| \|\vec{v}\| \|\vec{B}\| \sin \varphi.$$

Die größte Kraft wirkt somit, wenn $\vec{v}$ und $\vec{B}$ zueinander senkrecht sind. Sind hingegen $\vec{v}$ und $\vec{B}$ parallel, so ist $\vec{F} = \vec{0}$ und es wirkt keine Kraft.

Berechnung: Man berechnet $\vec{x} \times \vec{y}$ formal über eine "Determinante":

$$\vec{x} \times \vec{y} = \begin{vmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} \vec{e}_1 - \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} \vec{e}_2 + \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \vec{e}_3 = \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}.$$

Beispiel:

$$\vec{x} = \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} -1 \\ 5 \\ 3 \end{pmatrix}, \quad \vec{x} \times \vec{y} = \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix} \times \begin{pmatrix} -1 \\ 5 \\ 3 \end{pmatrix} = \begin{pmatrix} (-1) \cdot 3 - 2 \cdot 5 \\ 2 \cdot (-1) - 2 \cdot 3 \\ 2 \cdot 5 - (-1) \cdot (-1) \end{pmatrix}.$$

17.14. Das Spatprodukt: Für $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^3$ heißt $(\vec{x} \times \vec{y}) \cdot \vec{z} = (\vec{x} \times \vec{y} | \vec{z})$ das *Spatprodukt* von $\vec{x}, \vec{y}, \vec{z}$.

Satz: Es gilt $(\vec{x} \times \vec{y}) \cdot \vec{z} = \det(\vec{x}, \vec{y}, \vec{z})$.

Beweis. Man weist für die linke Seite die Eigenschaften (D1)–(D3) der Determinante nach. Wegen der Eindeutigkeit der Determinante folgt die Behauptung (alternativ kann man die Formel auch direkt nachrechnen). $\square$

Beispiel: Es gilt

$$\|\vec{x} \times \vec{y}\|^2 = (\vec{x} \times \vec{y}) \cdot (\vec{x} \times \vec{y}) = \det(\vec{x}, \vec{y}, \vec{x} \times \vec{y})$$

(setze $\vec{z} = \vec{x} \times \vec{y}$). Dies zeigt: Sind $\vec{x}, \vec{y}$ linear unabhängig, so ist $\vec{x}, \vec{y}, \vec{x} \times \vec{y}$ eine Basis von $\mathbb{R}^3$ und zwar ein Rechtssystem.

18 Eigenwerte und Diagonalisierung von Matrizen

18.1. Eigenwerte und Eigenvektoren: Es sei $A \in \mathbb{C}^{n \times n}$. Ein $\lambda \in \mathbb{C}$ heißt *Eigenwert* von A , falls es ein $\vec{x} \in \mathbb{C}^n \setminus \{\vec{0}\}$ gibt mit

$$A\vec{x} = \lambda\vec{x}.$$

Jedes solche $\vec{x}$ heißt *Eigenvektor zum Eigenwert λ* (von A).

Der *Eigenraum von A zum Eigenwert λ* ist

$$E_A(\lambda) := \{\vec{x} \in \mathbb{C}^n : A\vec{x} = \lambda\vec{x}\}.$$

Er besteht aus $\vec{0}$ und allen Eigenvektoren von A zum Eigenwert λ . **Beachte:** $\vec{0}$ ist kein Eigenvektor!

Sei V ein $\mathbb{C}$ -Vektorraum und $\phi : V \rightarrow V$ linear. Ein $\lambda \in \mathbb{C}$ heißt Eigenwert von ϕ , falls es ein $v \in V \setminus \{0\}$ gibt mit

$$\phi(v) = \lambda v.$$

Jedes solche v heißt *Eigenvektor zum Eigenwert λ* (von ϕ). Den Eigenraum definiert man entsprechend:

$$E_\phi(\lambda) := \{v \in V : \phi(v) = \lambda v\}.$$

Beispiele: (1) 1 ist der einzige Eigenwert von I_n und $E_{I_n}(1) = \mathbb{C}^n$.

(2) Die Diagonalelemente $d_1, d_2, \dots, d_n$ einer Diagonalmatrix

$$D = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & d_n \end{pmatrix} =: \text{diag}(d_1, \dots, d_n)$$

sind Eigenwerte von D . Für $j = 1, \dots, n$ ist der j -te Einheitsvektor $\vec{e}_j$ ein Eigenvektor zum Eigenwert d_j .

(3) Für die reelle Matrix

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \text{ ist z.B. } A \begin{pmatrix} i \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ i \end{pmatrix} = i \begin{pmatrix} i \\ 1 \end{pmatrix}.$$

Also ist i ein Eigenwert von A und $\begin{pmatrix} i \\ 1 \end{pmatrix}$ ist ein zugehöriger Eigenvektor. Dieses Beispiel zeigt:

Reelle Matrizen können nichtreelle Eigenwerte haben!

Übung: Hat eine reelle Matrix $A \in \mathbb{R}^{n \times n}$ einen reellen Eigenwert $\lambda \in \mathbb{R}$, so existiert dazu auch ein reeller Eigenvektor $x \in \mathbb{R}^n \setminus \{0\}$.

Wir betrachten von nun an komplexe Matrizen.

18.2. Bemerkungen (geometrische Vielfachheit): Sei $A \in \mathbb{C}^{n \times n}$. Dann gilt:

(a) Ein $\lambda \in \mathbb{C}$ ist Eigenwert von A genau dann, wenn $\text{Kern}(A - \lambda I_n) \neq \{\vec{0}\}$ ist. In diesem Fall ist

$$E_A(\lambda) = \text{Kern}(A - \lambda I_n).$$

(b) Zu jedem Eigenwert λ von A ist $E_A(\lambda)$ ein Untervektorraum von $\mathbb{C}^n$ mit

$$m := \dim(E_A(\lambda)) \geq 1.$$

Die Zahl m heißt *geometrische Vielfachheit des Eigenwertes λ* .

(c) Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig.

(d) Sind $\lambda_1, \dots, \lambda_k$ verschiedenen Eigenwerte von A , so gilt $k \leq n$ und

$$\dim(E_A(\lambda_1)) + \dim(E_A(\lambda_2)) + \dots + \dim(E_A(\lambda_k)) \leq n,$$

d.h. **die Summe der geometrischen Vielfachheiten aller Eigenwerte von A ist höchstens n und es gibt höchstens n verschiedene Eigenwerte.**

Beweis. (a) und (b) folgen aus der Definition von $E_A(\lambda)$ und (d) folgt aus (c).

(c) zeigt man durch (endliche) Induktion nach der Anzahl k verschiedener Eigenwerte. Der Induktionsanfang $k = 1$ ist klar. Für den Induktionsschritt seien $\lambda_1, \dots, \lambda_{k+1}$ verschiedene Eigenwerte mit zugehörigen Eigenvektoren $\vec{x}_1, \dots, \vec{x}_{k+1}$. Zum Beweis von deren Unabhängigkeit seien $\alpha_1, \dots, \alpha_{k+1} \in \mathbb{C}$ mit $\sum_{j=1}^{k+1} \alpha_j \vec{x}_j = \vec{0}$. Durch Multiplikation mit A bzw. mit λ_{k+1} folgt

$$\vec{0} = \sum_{j=1}^{k+1} \alpha_j A \vec{x}_j = \sum_{j=1}^{k+1} \alpha_j \lambda_j \vec{x}_j \quad \text{und} \quad \vec{0} = \sum_{j=1}^{k+1} \alpha_j \lambda_{k+1} \vec{x}_j.$$

Durch Differenzbildung erhalten wir

$$\vec{0} = \sum_{j=1}^{k+1} \alpha_j (\lambda_j - \lambda_{k+1}) \vec{x}_j = \sum_{j=1}^k \alpha_j (\lambda_j - \lambda_{k+1}) \vec{x}_j.$$

Nach Induktionsvoraussetzung ist dann $\alpha_j (\lambda_j - \lambda_{k+1}) = 0$ für $j = 1, \dots, k$, also $\alpha_j = 0$ für $j = 1, \dots, k$ wegen $\lambda_j \neq \lambda_{k+1}$. Schließlich folgt auch $\alpha_{k+1} = 0$ wegen $\vec{x}_{k+1} \neq 0$. $\square$

Beispiel: Die Matrix A aus Beispiel 18.1(3) hat genau die Eigenwerte $i, -i$ und es gilt

$$E_A(i) = \text{lin}\left\{\begin{pmatrix} i \\ 1 \end{pmatrix}\right\}, \quad E_A(-i) = \text{lin}\left\{\begin{pmatrix} i \\ -1 \end{pmatrix}\right\}.$$

Die geometrische Vielfachheit von i und $-i$ ist also jeweils 1.

18.3. Charakteristisches Polynom und algebraische Vielfachheit: Sei $A \in \mathbb{C}^{n \times n}$ eine Matrix. Das *charakteristische Polynom* p_A von A ist definiert durch

$$p_A(\lambda) := \det(A - \lambda I_n) \quad (\lambda \in \mathbb{C}).$$

Mit der Leibnizformel (vgl. 17.8) sieht man: Das charakteristische Polynom ist ein Polynom vom Grad n mit führendem Koeffizient $(-1)^n$, also $p_A(\lambda) = (-1)^n \lambda^n + \dots$.

Mit 17.2(e) und 18.2(a) folgt für $\lambda \in \mathbb{C}$:

$$\lambda \text{ ist Eigenwert von } A \Leftrightarrow p_A(\lambda) = 0.$$

Definition: Die *algebraische Vielfachheit* eines Eigenwertes λ von A ist die Vielfachheit von λ als Nullstelle des charakteristischen Polynoms p_A .

Bemerkung: Die geometrische Vielfachheit eines Eigenwertes ist immer $\leq$ seiner algebraischen Vielfachheit.

Sind $\lambda_1, \dots, \lambda_k$ die verschiedenen Eigenwerte und $m_1, m_2, \dots, m_k$ die jeweiligen algebraischen Vielfachheiten, so gilt $m_1 + m_2 + \dots + m_k = n$ (vgl. 5.5, HM I).

Beispiele: (1) Der Eigenwert 1 von I_n hat algebraische und geometrische Vielfachheit n : Es ist

$$E_{I_n}(1) = \mathbb{C}^n \text{ und } p_{I_n}(\lambda) = \det(I_n - \lambda I_n) = (1 - \lambda)^n.$$

(2) In Beispiel 18.2 haben i und $-i$ jeweils algebraische und geometrische Vielfachheit 1. Es ist

$$p_A(\lambda) = \det \begin{pmatrix} -\lambda & -1 \\ 1 & -\lambda \end{pmatrix} = \lambda^2 + 1 = (\lambda - i)(\lambda + i).$$

(3) Der Eigenwert 1 der Matrix $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ hat algebraische Vielfachheit 2 und geometrische Vielfachheit 1: Es gilt

$$p_A(\lambda) = (1 - \lambda)^2 = (\lambda - 1)^2 \text{ und } E_A(1) = \text{Kern} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \text{lin} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}.$$

18.4. Ähnliche Matrizen: Für Matrizen $A, B \in \mathbb{C}^{n \times n}$ definieren wir die Relation

$$A \sim B :\Leftrightarrow \text{es gibt eine reguläre Matrix } S \in \mathbb{C}^{n \times n} \text{ mit } B = S^{-1}AS.$$

Ist $A \sim B$, so nennen wir A und B *ähnlich*.

Bemerkungen:

(a) Die Relation $\sim$ ist eine Äquivalenzrelation:

(i) $A \sim A$ gilt mit $S = I_n$.

(ii) Ist $B = S^{-1}AS$ so folgt $A = SBS^{-1}$, denn

$$SBS^{-1} = S(S^{-1}AS)S^{-1} = \underbrace{(SS^{-1})}_{=I} A \underbrace{(SS^{-1})}_{=I} = A.$$

(iii) Gilt $A \sim B$ und $B \sim C$, so gilt auch $A \sim C$, denn $B = S^{-1}AS$ und $C = R^{-1}BR$ implizieren

$$C = R^{-1}S^{-1}ASR = (SR)^{-1}A(SR)$$

und mit S und R ist auch SR regulär.

(b) Ähnliche Matrizen haben dieselbe Determinante und dieselben Eigenwerte mit denselben algebraischen und geometrischen Vielfachheiten: Ist $B = S^{-1}AS$, so gilt

$$\det(B) = (\det(S))^{-1} \det(A) \det(S) = \det(A)$$

und

$$p_B(\lambda) = \det(S^{-1}AS - \lambda I) = \det(S^{-1}(A - \lambda I)S) = \det(A - \lambda I) = p_A(\lambda).$$

Also sind Eigenwerte und algebraische Vielfachheiten gleich. Weiter gilt für jeden Eigenwert λ von A :

$$\begin{aligned} E_A(\lambda) &= \{\vec{x} \in \mathbb{C}^n : A\vec{x} = \lambda\vec{x}\} = \{\vec{x} \in \mathbb{C}^n : S^{-1}A\vec{x} = \lambda S^{-1}\vec{x}\} \\ &= S(\{\vec{y} \in \mathbb{C}^n : \underbrace{S^{-1}AS}_B \vec{y} = \lambda \vec{y}\}) = S(E_B(\lambda)), \end{aligned}$$

wobei wir $\vec{x} = S\vec{y}$ geschrieben haben. Umgekehrt gilt $E_B(\lambda) = S^{-1}(E_A(\lambda))$. Da S regulär ist, haben $E_A(\lambda)$ und $E_B(\lambda)$ dieselbe Dimension.

Skizze:

$$\begin{array}{ccccc} \mathbb{C}^n & \xrightarrow{A} & \mathbb{C}^n & & \\ S \uparrow & & \downarrow & S^{-1} & \\ \mathbb{C}^n & \xrightarrow{B} & \mathbb{C}^n & & \end{array}$$

Man kann die Skizze so interpretieren, dass die zur Matrix A gehörige lineare Abbildung $\vec{x} \mapsto A\vec{x}$ in der durch die Spalten der regulären Matrix S gegebenen Basis durch die Matrix B dargestellt wird.

Beispiel: Sei $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ und $B = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$. Die Matrix $S = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ ist regulär.

Es gilt $S^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ und $S^{-1}AS = B$. Also sind A, B ähnlich.

Ist ϕ die lineare Abbildung $\vec{x} \mapsto A\vec{x}$, so wird ϕ bzgl. der Basis $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ durch die Matrix B dargestellt, d.h. ϕ streckt in Richtung von $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ um den Faktor 4 und in Richtung $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ um den Faktor 2.

18.5. Die Spur einer Matrix: Für $A = (a_{jk})_{jk} \in \mathbb{C}^{n \times n}$ definiert man die *Spur* von A durch

$$\text{Spur}(A) = \sum_{j=1}^n a_{jj},$$

d.h. als Summe der Diagonaleinträge.

Satz: (a) Für $A, C \in \mathbb{C}^{n \times n}$ gilt $\text{Spur}(AC) = \text{Spur}(CA)$.

(b) Ähnliche Matrizen haben dieselbe Spur.

(c) Ist $p_A(\lambda) = (-1)^n \lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_1 \lambda + a_0$ das charakteristische Polynom der Matrix $A \in \mathbb{C}^{n \times n}$, so gilt $a_{n-1} = (-1)^{n-1} \text{Spur}(A)$ und $a_0 = \det(A)$.

Beweis. (a) Gilt $C = (c_{kl})_{kl}$, so ist

$$AC = \left(\sum_{k=1}^n a_{jk} c_{kl} \right)_{jl} \quad \text{und} \quad CA = \left(\sum_{j=1}^n c_{lj} a_{jk} \right)_{lk},$$

also

$$\text{Spur}(AC) = \sum_{j=1}^n \left(\sum_{k=1}^n a_{jk} c_{kj} \right) = \sum_{k=1}^n \left(\sum_{j=1}^n c_{kj} a_{jk} \right) = \text{Spur}(CA).$$

(b) Unter Verwendung von (a) gilt für eine reguläre Matrix S :

$$\text{Spur}(\underbrace{S^{-1}AS}_{=C}) = \text{Spur}(SS^{-1}A) = \text{Spur}(A).$$

(c) $a_0 = p_A(0) = \det(A - 0 \cdot I) = \det(A)$. Aussage über a_{n-1} ohne Beweis ($n = 2$ siehe unten, man kann einen Induktionsbeweis führen und z.B. $\det(A - \lambda I)$ nach der ersten Zeile entwickeln). $\square$

Beispiel: Für $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ gilt

$$p_A(\lambda) = \begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} = \underbrace{ad - bc}_{=\det(A)} - \underbrace{(a + d)}_{=\text{Spur}(A)} \lambda + \lambda^2.$$

Hier sind die Eigenwerte von A also durch $\det A$ und $\text{Spur } A$ eindeutig bestimmt.

18.6. Diagonalisierung von Matrizen: Eine Matrix $A \in \mathbb{C}^{n \times n}$ heißt *diagonalisierbar*, falls sie ähnlich zu einer Diagonalmatrix D ist, d.h. falls es eine reguläre Matrix $S \in \mathbb{C}^{n \times n}$ so gibt, dass $S^{-1}AS$ eine Diagonalmatrix ist.

Beispiel: Nach dem Beispiel in 18.4 ist $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ diagonalisierbar.

Bemerkung: Auf der Diagonalen von D müssen dann die Eigenwerte von A stehen, gemäß ihrer algebraischen Vielfachheit wiederholt. Ist nämlich $D = \text{diag}(d_1, \dots, d_n)$ so ist jedes $d \in \{d_j : j = 1, \dots, n\}$ ein Eigenwert von D mit algebraischer Vielfachheit = Anzahl der $j \in \{1, \dots, n\}$ mit $d_j = d$. Der Eigenraum zu d ist

$$E_D(d) = \text{Kern}(D - dI_n) = \text{lin}\{\vec{e}_j : j \in \{1, \dots, n\}, d_j = d\},$$

also stimmt die geometrische Vielfachheit von d mit der algebraischen Vielfachheit überein.

Satz: Eine Matrix $A \in \mathbb{C}^{n \times n}$ ist genau dann diagonalisierbar, wenn für jeden Eigenwert von A geometrische und algebraische Vielfachheit übereinstimmen.

Eine entsprechende Matrix S erhält man folgendermaßen: Man wähle in jedem Eigenraum eine Basis und schreibe die Vektoren als Spalten $\vec{s}_1, \vec{s}_2, \dots, \vec{s}_n$ in eine Matrix S . Ist λ_j der Eigenwert zum Eigenvektor $\vec{s}_j$, so erhält man $AS = SD$, mit $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ (die Matrix SD hat die Spalten $\lambda_1 \vec{s}_1, \lambda_2 \vec{s}_2, \dots, \lambda_n \vec{s}_n$). Die Matrix S ist regulär und somit ist $S^{-1}AS = D$.

Beispiel: Wir betrachten $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$. Es gilt

$$p_A(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 2 - \lambda & 1 & 1 \\ 1 & 2 - \lambda & 1 \\ 1 & 1 & 2 - \lambda \end{vmatrix} = -(\lambda - 4)(\lambda - 1)^2,$$

also ist 4 Eigenwert mit algebraischer Vielfachheit 1, und 1 ist Eigenwert mit algebraischer Vielfachheit 2. Für die Eigenräume gilt

$$E_A(4) = \text{Kern} \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} = \text{lin}\left\{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right\}$$

und

$$E_A(1) = \text{Kern} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \text{lin}\left\{\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}\right\},$$

d.h. für jeden Eigenwert von A sind algebraische und geometrische Vielfachheit gleich und A ist diagonalisierbar. Wir setzen

$$S := \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix}$$

und erhalten

$$S^{-1} = \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & -1 \\ -1 & 2 & -1 \end{pmatrix} \quad \text{und} \quad S^{-1}AS = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Bemerkung: Folgende Eigenschaften sind ebenfalls äquivalent zur Diagonalisierbarkeit von $A \in \mathbb{C}^{n \times n}$:

- (a) $\mathbb{C}^n$ hat eine Basis aus Eigenvektoren von A .
- (b) Die Summe der geometrischen Vielfachheiten der Eigenwerte von A ist n .

Folgerung: Sind die Eigenwerte von $A \in \mathbb{C}^{n \times n}$ paarweise verschieden, so ist A diagonalisierbar (dann ist für jeden Eigenwert algebraische und geometrische Vielfachheit = 1).

18.7. Symmetrische und hermitesche Matrizen: Eine Matrix $A \in \mathbb{R}^{n \times n}$ mit $A = A^T$ heißt *symmetrisch*. Eine Matrix $A \in \mathbb{C}^{n \times n}$ mit $A = A^*$ heißt *hermitesch* oder *selbstadjungiert*.

Bemerkung: Jede symmetrische Matrix ist also hermitesch.

Beispiele: Die Matrix $A \in \mathbb{R}^{2 \times 2}$ aus dem Beispiel in 18.4 ist symmetrisch. Die Matrix

$$\begin{pmatrix} 0 & -2i \\ 2i & 0 \end{pmatrix}$$

ist hermitesch.

Satz: Es sei $A \in \mathbb{C}^{n \times n}$ eine hermitesche Matrix. Dann gilt:

- (a) A hat nur reelle Eigenwerte. Die Eigenvektoren zu verschiedenen Eigenwerten sind paarweise orthogonal.
- (b) A ist diagonalisierbar, wobei die Matrix S unitär, d.h. mit $S^* = S^{-1}$, gewählt werden kann. Für reelle symmetrische Matrizen kann $S \in \mathbb{R}^{n \times n}$ orthogonal, d.h. mit $S^T = S^{-1}$ gewählt werden (ohne Beweis).
- (c) Für A existiert die folgende *Spektralzerlegung*: Es seien $\lambda_1, \dots, \lambda_m$ die verschiedenen Eigenwerte von A , und $P_k : \mathbb{C}^n \rightarrow E_A(\lambda_k)$ die Orthogonalprojektion mit Bild $P_k = E_A(\lambda_k)$ und Kern $P_k = (E_A(\lambda_k))^\perp$ ($k = 1, \dots, m$) (vgl. 16.7). Dann gilt

$$\forall \vec{x} \in \mathbb{C}^n : \quad \vec{x} = \sum_{k=1}^m P_k \vec{x}, \quad A\vec{x} = \sum_{k=1}^m \lambda_k P_k \vec{x}$$

und $P_j P_k = \delta_{jk} P_j$ ($j, k = 1, \dots, m$) (ohne Beweis).

Beweis. zu (a): Für jedes $\vec{x} \in \mathbb{C}^n$ gilt

$$(A\vec{x}|\vec{x}) = (\vec{x}|A^*\vec{x}) = (\vec{x}|A\vec{x}) = \overline{(A\vec{x}|\vec{x})}, \text{ d.h. } (A\vec{x}|\vec{x}) \in \mathbb{R}.$$

Sei λ ein Eigenwert von A und $\vec{x}$ ein zugehöriger Eigenvektor. Dann gilt

$$\lambda \|\vec{x}\|^2 = (\lambda\vec{x}|\vec{x}) = (A\vec{x}|\vec{x}) \in \mathbb{R}, \text{ also } \lambda = \frac{(A\vec{x}|\vec{x})}{\|\vec{x}\|^2} \in \mathbb{R}.$$

Ist $\mu \neq \lambda$ ein weiterer Eigenwert und $\vec{y}$ ein zugehöriger Eigenvektor, so gilt

$$\lambda(\vec{x}|\vec{y}) = (\lambda\vec{x}|\vec{y}) = (A\vec{x}|\vec{y}) = (\vec{x}|A\vec{y}) = (\vec{x}|\mu\vec{y}) = \mu(\vec{x}|\vec{y}),$$

also $(\vec{x}|\vec{y}) = 0$ wegen $\lambda \neq \mu$. □

Beispiele: Die symmetrische Matrix $\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ lässt sich durch die orthogonale Matrix $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ diagonalisieren (siehe 18.4 und 18.6).

Die Matrix $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ aus dem Beispiel in 18.6 ist symmetrisch, die dort angegebene Matrix S ist aber nicht orthogonal. Hier reicht es nicht, die gefundenen Eigenvektoren auf Länge 1 zu bringen, da die beiden Eigenvektoren zum Eigenwert 1 nicht orthogonal sind. Wir können aber das Gram-Schmidt-Verfahren verwenden und setzen

$$\begin{aligned} \vec{s}_1 &:= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, & \vec{s}_2 &:= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \\ \vec{c} &:= \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} - \left(\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \middle| \vec{s}_2 \right) \vec{s}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}, \\ \vec{s}_3 &:= \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}. \end{aligned}$$

Die Matrix mit den Spalten $\vec{s}_1, \vec{s}_2, \vec{s}_3$ ist orthogonal und diagonalisiert A : Mit

$$S = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{2} & \sqrt{3} & -1 \\ \sqrt{2} & 0 & 2 \\ \sqrt{2} & -\sqrt{3} & -1 \end{pmatrix}, \quad S^{-1} = S^T = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{2} & \sqrt{2} & \sqrt{2} \\ \sqrt{3} & 0 & -\sqrt{3} \\ -1 & 2 & -1 \end{pmatrix}$$

ist

$$S^T A S = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Definition: Eine Matrix $A \in \mathbb{C}^{n \times n}$ heißt *normal* wenn $AA^* = A^*A$ gilt.

Satz (ohne Beweis): Eine Matrix $A \in \mathbb{C}^{n \times n}$ läßt sich genau dann unitär diagonalisieren, wenn sie normal ist.

Es sei noch der folgende tiefliegendere Satz angegeben:

18.8. Satz (Jordan-Normalform): Ohne Beweis: Zu jeder Matrix $A \in \mathbb{C}^{n \times n}$ gibt es eine reguläre Matrix S so, dass $S^{-1}AS$ die folgende Blockmatrix-Struktur hat

$$S^{-1}AS = J = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_l \end{pmatrix},$$

wobei jedes J_j (*Jordanblock*) die Gestalt

$$J_j = \begin{pmatrix} \lambda_j & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda_j & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_j \end{pmatrix}$$

hat, d.h. auf der Hauptdiagonalen steht bei J_j ein Eigenwert, auf der oberen Nebendiagonalen stehen Einsen.

Insgesamt stehen in einer *Jordannormalform* J auf der Diagonalen die Eigenwerte von A , gemäß ihrer algebraischen Vielfachheiten wiederholt und auf der Nebendiagonalen stehen nur Einsen und Nullen. Dabei ist die Anzahl der Einsen gerade n minus die Summe der geometrischen Vielfachheiten aller Eigenwerte.

Bemerkung: Die Spalten der Matrix S erhält man hier durch eine geeignete Wahl von Basen in den *Haupträumen* $\text{Kern}((A - \lambda I_n)^n)$, wobei λ die Eigenwerte von A durchläuft.

Folgerung: Sei $A \in \mathbb{C}^{n \times n}$ und seien $\lambda_1, \lambda_2, \dots, \lambda_n$ die Eigenwerte von A , wobei jeder Eigenwert gemäß seiner algebraischen Vielfachheit wiederholt sei. Dann gilt

$$\det(A) = \det(J) = \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_n \quad \text{und} \quad \text{Spur}(A) = \text{Spur}(J) = \lambda_1 + \lambda_2 + \dots + \lambda_n,$$

d.h. $\det(A)$ ist das Produkt der Eigenwerte und $\text{Spur}(A)$ ist die Summe der Eigenwerte.

Beispiel (zur Jordan-Normalform): Sei $A = \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix}$. Dann gilt $p_A(\lambda) = (\lambda - 2)^2$, also ist 2 Eigenwert mit algebraischer Vielfachheit 2. Weiter ist

$$E_A(2) = \text{Kern}(A - 2I) = \text{Kern} \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} = \text{lin}\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\},$$

also hat der Eigenwert 2 die geometrische Vielfachheit 1 und $\vec{s}_1 := \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ist ein Eigenvektor. Nun sucht man einen Vektor $\vec{s}_2$ mit $(A - 2I)\vec{s}_2 = \vec{s}_1$ und findet $\vec{s}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Sei nun S die Matrix mit den Spalten $\vec{s}_1, \vec{s}_2$. Dann haben wir

$$S = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad S^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}, \quad J = S^{-1}AS = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix},$$

wobei die Jordan-Normalform hier nur einen Jordan-Block hat.

Bemerkungen: (1) Ist $J = D + N$ eine Jordannormalform von $A \in \mathbb{C}^{n \times n}$ wobei D die Diagonalmatrix mit den Eigenwerten von A ist und $N := J - D$ die Matrix mit den zugehörigen Einsen und Nullen auf der oberen Nebendiagonalen, so gilt

$$DN = ND \quad \text{und} \quad N^n = 0$$

(nachrechnen). Matrizen für die eine Potenz die Nullmatrix ergibt heißen *nilpotent*. Die Jordannormalform zeigt also:

Jedes $A \in \mathbb{C}^{n \times n}$ ist Summe einer diagonalisierbaren Matrix und einer nilpotenten Matrix, welche kommutieren:

$$A = SJS^{-1} = SDS^{-1} + SNS^{-1} =: B + C$$

mit B ist diagonalisierbar, $C^n = 0$ und $BC = CB$.

(2) Ohne Beweis: Zwei Matrizen $A, B \in \mathbb{C}^{n \times n}$ sind genau dann ähnlich, wenn sie die gleiche Jordannormalform haben, wobei zu beachten ist, daß die Jordannormalform nur bis auf eine Permutation der Jordanblöcke eindeutig bestimmt ist.

18.9. Definitheit reeller symmetrischer Matrizen: Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch und $q : \mathbb{R}^n \rightarrow \mathbb{R}$, $q(\vec{x}) = \vec{x}^T A \vec{x}$, die zugehörige *quadratische Form*.

A heißt *positiv definit*, falls

$$\forall \vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\} : \vec{x}^T A \vec{x} > 0.$$

A heißt *negativ definit*, falls

$$\forall \vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\} : \vec{x}^T A \vec{x} < 0.$$

A heißt *positiv semidefinit*,

$$\forall \vec{x} \in \mathbb{R}^n : \vec{x}^T A \vec{x} \geq 0.$$

A heißt *negativ semidefinit*, falls

$$\forall \vec{x} \in \mathbb{R}^n : \vec{x}^T A \vec{x} \leq 0.$$

A heißt *indefinit*, falls

$$\exists \vec{x}, \vec{y} \in \mathbb{R}^n : \vec{x}^T A \vec{x} > 0 \wedge \vec{y}^T A \vec{y} < 0.$$

Bemerkungen: (1) A ist positive [semi]definit genau dann wenn $-A$ negativ [semi]definit ist. Ist A indefinit, so ist auch $-A$ indefinit.

(2) Ist $A = (a_{jk}) \in \mathbb{R}^{n \times n}$, so gilt

$$q(\vec{x}) = \vec{x}^T A \vec{x} = \sum_{j=1}^n \sum_{k=1}^n a_{jk} x_j x_k \quad (\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n).$$

Beachte, dass jedes $A \in \mathbb{R}^{n \times n}$ eine quadratische Form erzeugt. Wir betrachten aber nur symmetrische Matrizen, da für beliebiges $A \in \mathbb{R}^{n \times n}$ und $\vec{x} \in \mathbb{R}^n$ stets

$$q(\vec{x}) = \vec{x}^T A \vec{x} = \vec{x}^T \left(\frac{A + A^T}{2} \right) \vec{x} + \vec{x}^T \left(\frac{A - A^T}{2} \right) \vec{x} = \vec{x}^T \left(\frac{A + A^T}{2} \right) \vec{x}.$$

Beispiel: Für $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ und $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$ gilt

$$\begin{aligned} \vec{x}^T A \vec{x} &= (x_1 \ x_2) A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 3x_1^2 + x_1x_2 + x_2x_1 + 3x_2^2 \\ &= 3x_1^2 + 2x_1x_2 + 3x_2^2 = 2x_1^2 + 2x_2^2 + (x_1 + x_2)^2. \end{aligned}$$

Also ist A positiv definit.

Definition: Für $A \in \mathbb{C}^{n \times n}$ heißt die Menge

$$\sigma(A) := \{ \lambda \in \mathbb{C} : \lambda \text{ ist Eigenwert von } A \}$$

das *Spektrum* von A .

Satz: Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch. Dann gilt:

A ist positiv definit $\Leftrightarrow \forall \lambda \in \sigma(A) : \lambda > 0$.

A ist negativ definit $\Leftrightarrow \forall \lambda \in \sigma(A) : \lambda < 0$.

A ist positiv semidefinit $\Leftrightarrow \forall \lambda \in \sigma(A) : \lambda \geq 0$.

A ist negativ semidefinit $\Leftrightarrow \forall \lambda \in \sigma(A) : \lambda \leq 0$.

A ist indefinit $\Leftrightarrow \exists \lambda, \mu \in \sigma(A) : \lambda > 0 \text{ und } \mu < 0$.

Beweis. Seien $\lambda_1, \dots, \lambda_n$ die Eigenwerte von A (gemäß algebraischer Vielfachheit wiederholt). Diagonalisiere A mit einer orthogonalen Matrix S : $S^T A S = D = \text{diag}(\lambda_1, \dots, \lambda_n)$. Schreibe

$$\vec{x} = S\vec{y}, \quad \text{wobei} \quad \vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Dann ist $\vec{x} = \vec{0} \Leftrightarrow \vec{y} = \vec{0}$ und

$$\vec{x}^T A \vec{x} = \vec{y}^T S^T A S \vec{y} = \vec{y}^T D \vec{y} = \sum_{j=1}^n \lambda_j y_j^2.$$

Hieraus lassen sich alle Fälle ablesen. □

Folgerung: Sei $A \in \mathbb{R}^{2 \times 2}$ symmetrisch. Dann gilt:

A ist indefinit genau dann, wenn $\det(A) < 0$ ist.

A ist positiv definit genau dann, wenn $\det(A) > 0$ und $\text{Spur}(A) > 0$ ist.

A ist negativ definit genau dann, wenn $\det(A) > 0$ und $\text{Spur}(A) < 0$ ist.

Beispiele: Für

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$$

gilt: A ist indefinit, B ist positiv definit und C ist negativ definit.

Für größere Matrizen kann man das folgende Kriterium verwenden:

Kriterium von Hurwitz: Eine symmetrische Matrix $A = (a_{jk}) \in \mathbb{R}^{n \times n}$ ist positiv definit genau dann, wenn alle *Hauptunterdeterminanten* positiv sind, d.h. wenn

$$\begin{vmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mm} \end{vmatrix} > 0 \quad (m = 1, 2, \dots, n).$$

Ohne Beweis.

Warnung: Die Aussage “ A ist positive semidefinit genau dann wenn alle Hauptunterdeterminanten ≥ 0 sind” ist **falsch!** Betrachte z.B.

$$A = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}.$$

19 Mehrdimensionale Differentialrechnung

Wie in den vorigen Kapiteln sind Vektoren $\vec{x} \in \mathbb{R}^n$ mit den Komponenten $x_1, x_2, \dots, x_n$ Spaltenvektoren:

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \text{also } \mathbb{R}^n = \mathbb{R}^{n \times 1}.$$

Um die Notation übersichtlich zu halten schreiben wir dennoch oft $\vec{x} = (x_1, \dots, x_n)$ (wenn es keine Rolle spielt). Gemeint sind aber immer Spaltenvektoren! Spielt es eine Rolle (z.B. beim Matrix-Vektor-Produkt) so schreibt man auch Spaltenvektoren.

Im Fall $n = 2, 3$ schreiben wir häufig x, y, z statt x_1, x_2, x_3 .

Weiter sei im folgenden $\|\cdot\|$ die *Euklidnorm*: $\|\vec{x}\| = \sqrt{x_1^2 + \dots + x_n^2}$.

19.1. Konvergenz von Folgen in $\mathbb{R}^n$: Eine Folge $(\vec{x}_k)_{k \in \mathbb{N}}$ von Vektoren in $\mathbb{R}^n$ heißt *konvergent*, falls es einen Vektor $\vec{x}_0 \in \mathbb{R}^n$ gibt mit

$$\|\vec{x}_k - \vec{x}_0\| \rightarrow 0 \quad (k \rightarrow \infty).$$

Wir schreiben dann $\lim_{k \rightarrow \infty} \vec{x}_k = \vec{x}_0$ oder $\vec{x}_k \rightarrow \vec{x}_0$ ($k \rightarrow \infty$).

Konvergenz in $\mathbb{R}^n$ bedeutet koordinatenweise Konvergenz: Wegen

$$|y_l| \leq \|\vec{y}\| = \sqrt{\sum_{j=1}^n |y_j|^2} \quad (l = 1, \dots, n)$$

gilt für eine Folge $(\vec{x}_k)_{k \in \mathbb{N}}$ in $\mathbb{R}^n$ mit $\vec{x}_k = (\xi_1^{(k)}, \dots, \xi_n^{(k)})$ und $\vec{x}_0 = (\eta_1, \dots, \eta_n) \in \mathbb{R}^n$:

$$\vec{x}_k \rightarrow \vec{x}_0 \quad (k \rightarrow \infty) \iff \forall j \in \{1, \dots, n\} : \xi_j^{(k)} \rightarrow \eta_j \quad (k \rightarrow \infty).$$

Beispiel: Sei $\vec{x}_k := (e^{-k}, 1 - 1/k)$ ($k \in \mathbb{N}$). Dann gilt $\vec{x}_k \rightarrow (0, 1)$ ($k \rightarrow \infty$).

19.2. Offene, abgeschlossene und kompakte Mengen in $\mathbb{R}^n$: Ist $\vec{x}_0 \in \mathbb{R}^n$ und $r > 0$, so heißt

$$K(\vec{x}_0, r) := \{\vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{x}_0\| < r\}$$

die *offene Kugel um $\vec{x}_0$ mit Radius r* . Eine Teilmenge $Q \subseteq \mathbb{R}^n$ heißt *offen*

$$:\Leftrightarrow \forall \vec{x}_0 \in Q \exists r > 0 : K(\vec{x}_0, r) \subseteq Q.$$

Beispiele: Folgende Teilmengen des $\mathbb{R}^n$ sind offen:

$$\mathbb{R}^n, \quad \emptyset, \quad K(\vec{y}_0, s), \quad (0, 1)^n.$$

Beweis. Wir zeigen, dass $K(\vec{y}_0, s)$ offen ist. Dazu sei $\vec{x}_0 \in K(\vec{y}_0, s)$, d.h. $\|\vec{x}_0 - \vec{y}_0\| < s$. Wir setzen $r := s - \|\vec{x}_0 - \vec{y}_0\|$ und zeigen $K(\vec{x}_0, r) \subseteq K(\vec{y}_0, s)$. Dazu sei $\vec{x} \in K(\vec{x}_0, r)$, d.h. $\|\vec{x} - \vec{x}_0\| < r$. Dann gilt

$$\|\vec{x} - \vec{y}_0\| = \|\vec{x} - \vec{x}_0 + \vec{x}_0 - \vec{y}_0\| \leq \|\vec{x} - \vec{x}_0\| + \|\vec{x}_0 - \vec{y}_0\| < r + \|\vec{x}_0 - \vec{y}_0\| = s,$$

also $\vec{x} \in K(\vec{y}_0, s)$. □

Definition: Eine Teilmenge $A \subseteq \mathbb{R}^n$ heißt *abgeschlossen*, falls $\mathbb{R}^n \setminus A$ offen ist.

Satz: Eine Teilmenge $A \subseteq \mathbb{R}^n$ ist abgeschlossen genau dann, wenn für jede Folge $(\vec{x}_k)_{k \in \mathbb{N}}$ in A , die in $\mathbb{R}^n$ konvergiert, der Grenzwert $\lim_{k \rightarrow \infty} \vec{x}_k$ wieder zu A gehört.

Beweis. Sei A abgeschlossen, also $\mathbb{R}^n \setminus A$ offen. Sei $(\vec{x}_k)$ eine Folge in A mit $\vec{x}_k \rightarrow \vec{x}_0 \in \mathbb{R}^n$. Ist $\vec{x}_0 \notin A$, so existiert ein $\varepsilon > 0$ mit $K(\vec{x}_0, \varepsilon) \subseteq \mathbb{R}^n \setminus A$. Wegen $x_k \in A$ ($k \in \mathbb{N}$) und $\vec{x}_k \rightarrow \vec{x}_0$ folgt $x_k \in K(\vec{x}_0, \varepsilon) \cap A$ ffa $k \in \mathbb{N}$, im Widerspruch zu $K(\vec{x}_0, \varepsilon) \cap A = \emptyset$.

Sei umgekehrt A nicht abgeschlossen, d.h. $\mathbb{R}^n \setminus A$ nicht offen. Dann gibt es ein $\vec{x}_0 \notin A$ so, dass für jedes $r > 0$ gilt $A \cap K(\vec{x}_0, r) \neq \emptyset$. Wir finden dann zu jedem $k \in \mathbb{N}$ ein $\vec{x}_k \in A \cap K(\vec{x}_0, 1/k)$. Es ist dann $(\vec{x}_k)_{k \in \mathbb{N}}$ eine Folge in A mit $\vec{x}_k \rightarrow \vec{x}_0 \notin A$. □

Beispiele: $\mathbb{R}^n$ und $\emptyset$ sind abgeschlossen. Endliche Mengen sind abgeschlossen.

Abgeschlossene Kugeln

$$\{\vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{x}_0\| \leq r\}$$

sind abgeschlossen.

Definition: Eine Teilmenge $A \subseteq \mathbb{R}^n$ heißt *beschränkt*, falls gilt:

$$\exists c \geq 0 \forall \vec{x} \in A : \|\vec{x}\| \leq c.$$

Eine Folge in $\mathbb{R}^n$ heißt *beschränkt*, falls die Menge ihrer Folgenglieder beschränkt ist.

Beispiele: Offene und abgeschlossene Kugeln sind beschränkt. Die Menge

$$\{(x, y) \in \mathbb{R}^2 : x^2 - y^2 = 0\}$$

ist unbeschränkt.

Definition: Eine Teilmenge $A \subseteq \mathbb{R}^n$ heißt *kompakt* wenn jede Folge $(\vec{x}_k)$ in A eine konvergente Teilfolge $(\vec{x}_{j(k)})$ hat mit

$$\lim_{k \rightarrow \infty} \vec{x}_{j(k)} \in A.$$

Wie in HM I gilt auch im mehrdimensionalen Fall der

Satz von Bolzano Weierstraß: Jede beschränkte Folge in $\mathbb{R}^n$ hat eine konvergente Teilfolge.

Beweis. Wir betrachten den Fall $n = 2$. Es sei $((x_k, y_k))_{k=1}^\infty$ eine beschränkte Folge in $\mathbb{R}^2$. Wegen

$$\max\{|x_k|, |y_k|\} \leq \|(x_k, y_k)\| \quad (k \in \mathbb{N})$$

sind (x_k) und (y_k) beschränkte Folgen in $\mathbb{R}$. Also hat (x_k) eine konvergente Teilfolge $(x_{j(k)})$. Weiter hat die beschränkte Folge $(y_{j(k)})$ eine konvergente Teilfolge $(y_{j(l(k))})$. Nun ist $((x_{j(l(k))}, y_{j(l(k))}))$ eine konvergente Teilfolge von $((x_k, y_k))$. Im allgemeinen Fall $n \in \mathbb{N}$ wird sukzessive n -mal eine konvergente Teilfolge gewählt. $\square$

Ebenso wie in HM I ergibt sich daraus:

Satz: $A \subseteq \mathbb{R}^n$ ist kompakt $\Leftrightarrow A$ ist beschränkt und abgeschlossen.

Beispiele: Abgeschlossene Kugeln sind kompakt. Offene Kugeln sind nicht kompakt (da nicht abgeschlossen). Die Menge

$$\{(x, y) \in \mathbb{R}^2 : x^2 - y^2 = 0\}$$

ist abgeschlossen aber nicht kompakt (da unbeschränkt).

Definition: Sei $A \subseteq \mathbb{R}^n$. Wir definieren

(1) das *Innere* von A :

$$A^\circ := \{\vec{x} \in A : \exists r > 0 : K(\vec{x}, r) \subseteq A\},$$

(2) den *Rand* von A :

$$\partial A := \{\vec{x} \in \mathbb{R}^n : \forall r > 0 : K(\vec{x}, r) \cap A \neq \emptyset \wedge K(\vec{x}, r) \cap (\mathbb{R}^n \setminus A) \neq \emptyset\}.$$

(3) den *Abschluß* von A :

$$\overline{A} := A \cup \partial A.$$

Bemerkung: Es gilt stets: A° ist offen und $\partial A, \overline{A}$ sind abgeschlossen. Punkte aus ∂A bzw. A° heißen *Randpunkte* bzw. *innere Punkte* von A .

Beispiele: (1) Es gilt

$$\overline{K(\vec{x}_0, r)} = \{\vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{x}_0\| \leq r\}$$

und

$$\partial K(\vec{x}_0, r) = \{\vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{x}_0\| = r\}.$$

(2) $\partial \mathbb{R}^n = \emptyset$, $K(\vec{x}_0, r)^\circ = K(\vec{x}_0, r)$, $([0, 1] \times [0, 1])^\circ = (0, 1) \times (0, 1)$, $(\mathbb{Q} \times \mathbb{R})^\circ = \emptyset$.

19.3. Stetigkeit von Funktionen: Sei $D \subseteq \mathbb{R}^n$ nichtleer und $f : D \rightarrow \mathbb{R}^m$ eine Funktion. Dann gibt es Funktionen $f_1, \dots, f_m : D \rightarrow \mathbb{R}$, die *Komponentenfunktionen* von f mit

$$f(\vec{x}) = f(x_1, \dots, x_n) = \begin{pmatrix} f_1(\vec{x}) \\ \vdots \\ f_m(\vec{x}) \end{pmatrix} = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix} \quad (\vec{x} = (x_1, \dots, x_n) \in D).$$

Bemerkung: Der Graph einer solchen Funktion f ist Teilmenge des $\mathbb{R}^{n+m}$. Eine Skizze des Graphen solcher Funktionen ist daher möglich falls $n + m \leq 3$ ist, also den Fällen

1. $m = n = 1$ (bekannt),
2. $n = 2, m = 1$,
3. $n = 1, m = 2$.

Definition: Die Funktion f heißt *stetig in* $\vec{x}_0 \in D$, falls für alle Folgen $(\vec{x}_k)_{k \in \mathbb{N}}$ in D mit $\vec{x}_k \rightarrow \vec{x}_0$ gilt: $f(\vec{x}_k) \rightarrow f(\vec{x}_0)$.

f heißt *stetig in* D , falls f in **jedem** $\vec{x}_0 \in D$ stetig ist. Wir schreiben dann $f \in C(D, \mathbb{R}^m)$.

Definition: Ein $\vec{x}_0 \in \mathbb{R}^n$ heißt *Häufungspunkt von* D , falls es eine Folge $(\vec{x}_k)_{k \in \mathbb{N}}$ in $D \setminus \{\vec{x}_0\}$ gibt mit $\vec{x}_k \rightarrow \vec{x}_0$ ($k \rightarrow \infty$).

Sei $\vec{c} \in \mathbb{R}^n$ und $\vec{x}_0 \in \mathbb{R}^n$ ein Häufungspunkt von D . Wir schreiben

$$\lim_{\vec{x} \rightarrow \vec{x}_0} f(\vec{x}) = \vec{c},$$

falls $f(\vec{x}_k) \rightarrow \vec{c}$ für jede Folge $(\vec{x}_k)_{k \in \mathbb{N}}$ in $D \setminus \{\vec{x}_0\}$, die gegen $\vec{x}_0$ konvergiert. In diesem Fall ist $\vec{c}$ eindeutig bestimmt.

Für $n \geq 2$ definiert man außerdem (da keine Ordnung vorhanden ist):

$$\vec{x}_k \rightarrow \infty \ (k \rightarrow \infty) : \iff \|x_k\| \rightarrow \infty.$$

Damit ergeben sich die Definitionen von

$$\lim_{\vec{x} \rightarrow \vec{x}_0} f(\vec{x}) = \infty, \quad \lim_{\vec{x} \rightarrow \infty} f(\vec{x}) = \vec{c}, \quad \lim_{\vec{x} \rightarrow \infty} f(\vec{x}) = \infty$$

analog zum Fall $n = 1$. Beachte: Bei $\vec{x} \rightarrow \infty$ muß D unbeschränkt sein.

Wie in HM I gilt dann auch hier: $f : D \rightarrow \mathbb{R}^m$ ist genau dann in D stetig, wenn für alle $\vec{x}_0 \in D$, die Häufungspunkt von D sind, gilt:

$$\lim_{\vec{x} \rightarrow \vec{x}_0} f(\vec{x}) = f(\vec{x}_0).$$

Beispiel: Sei $D := \{(x, y) \in \mathbb{R}^2 : x > 0\}$ und

$$f : D \rightarrow \mathbb{R}^2, \quad f(x, y) = \begin{pmatrix} \sqrt{x^2 + y^2} \\ \arctan(y/x) \end{pmatrix}.$$

Dann ist $f \in C(D, \mathbb{R}^2)$:

Es sei $(x_0, y_0) \in D$ und $((x_k, y_k))_{k \in \mathbb{N}}$ eine Folge in D mit $(x_k, y_k) \rightarrow (x_0, y_0)$, also $x_k \rightarrow x_0$, $y_k \rightarrow y_0$. Es folgt $\sqrt{x_k^2 + y_k^2} \rightarrow \sqrt{x_0^2 + y_0^2}$ und (wegen $x_0 > 0$ und der Stetigkeit von $\arctan$) $\arctan(y_k/x_k) \rightarrow \arctan(y_0/x_0)$. Also $f(x_k, y_k) \rightarrow f(x_0, y_0)$.

Bemerkung: Wegen 19.1 ist f stetig in $\vec{x}_0$ (bzw. in D) genau dann, wenn alle Komponentenfunktionen $f_1, \dots, f_m$ stetig in $\vec{x}_0$ (bzw. in D) ist.

Beispiele: (1) Betrachte

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

f ist in jedem Punkt $(x, y) \neq (0, 0)$ stetig. f ist aber in $(0, 0)$ nicht stetig, denn für $(x_k, y_k) = (1/k, 1/k)$ gilt $(x_k, y_k) \rightarrow (0, 0)$, aber

$$f(x_k, y_k) = \frac{x_k^2}{2x_k^2} = \frac{1}{2} \not\rightarrow 0 = f(0, 0) \quad (k \rightarrow \infty).$$

(2) Sei $\beta > 1$ und

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = \begin{cases} \frac{x|y|^\beta}{x^2+y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}.$$

Die Funktion f ist in jedem Punkt $(x, y) \neq (0, 0)$ stetig. f ist auch $(0, 0)$ stetig: Wegen $|xy| \leq (x^2 + y^2)/2$ gilt

$$|f(x, y)| = \frac{|y|^{\beta-1}}{2} \frac{2|xy|}{x^2 + y^2} \leq \frac{|y|^{\beta-1}}{2} \rightarrow 0 \quad ((x, y) \rightarrow (0, 0)).$$

(3) Sei $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ linear. Dann ist ϕ stetig: Wegen $\|\phi(\vec{x}) - \phi(\vec{y})\| = \|\phi(\vec{x} - \vec{y})\|$ reicht es, Stetigkeit in $\vec{x}_0 = \vec{0}$ zu zeigen. Für $\vec{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ gilt

$$\|\phi(\vec{x})\| = \left\| \sum_{j=1}^n x_j \phi(\vec{e}_j) \right\| \leq \sum_{j=1}^n |x_j| \|\phi(\vec{e}_j)\| \leq \|\vec{x}\| \sqrt{\sum_{j=1}^n \|\phi(\vec{e}_j)\|^2} \rightarrow 0 \quad (\|\vec{x}\| \rightarrow 0).$$

Für $\vec{x} \rightarrow \vec{0}$ gilt also $\phi(\vec{x}) \rightarrow \vec{0} = \phi(\vec{0})$. Insbesondere gilt: Ist $A \in \mathbb{R}^{m \times n}$, so ist $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $f(\vec{x}) = A\vec{x}$ in $\mathbb{R}^m$ stetig, also $f \in C(\mathbb{R}^n, \mathbb{R}^m)$. Man identifiziert oft die Matrix A mit der von ihr erzeugten linearen Abbildung und schreibt $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$.

Bemerkung: Kompositionen von stetigen Funktionen sind stetig. Die Addition und die Multiplikation mit Skalaren, also

$$+ : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n \quad \text{und} \quad \cdot : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n,$$

sind stetig. Ebenfalls stetig sind: Das Skalarprodukt

$$(\cdot | \cdot) : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R},$$

das Matrix-Vektor-Produkt,

$$\mathbb{R}^{n \times m} \times \mathbb{R}^m \rightarrow \mathbb{R}^n, \quad (A, \vec{x}) \mapsto A\vec{x}$$

Multiplikation von Matrizen

$$\mathbb{R}^{n \times m} \times \mathbb{R}^{m \times q} \rightarrow \mathbb{R}^{n \times q}, \quad (A, B) \rightarrow AB,$$

die Determinante, die Spur und das Kreuzprodukt

$$\det : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}, \quad \text{Spur} : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}, \quad \times : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3.$$

Auch die auf den regulären Matrizen erklärte Matrixinversion

$$\cdot^{-1} : GL(n, \mathbb{R}) \rightarrow \mathbb{R}^{n \times n}, \quad A \mapsto A^{-1}$$

ist stetig (Cramersche Regel!). In all diesen Fällen wird dabei sowohl $\mathbb{R}^n \times \mathbb{R}^m$ als auch $\mathbb{R}^{n \times m}$ als $\mathbb{R}^{nm}$ aufgefasst und mit der zugehörigen Euklidnorm versehen. Z.B. für $A = (a_{jk}) \in \mathbb{R}^{n \times m}$ ist

$$\|A\| = \left(\sum_{j=1}^n \sum_{k=1}^m a_{jk}^2 \right)^{1/2}.$$

Beispiel: Sei $D \subseteq \mathbb{R}^n$, $f, g \in C(D, \mathbb{R}^m)$ und $\alpha, \beta \in \mathbb{R}$, so gilt

$$\alpha f + \beta g \in C(D, \mathbb{R}^m).$$

Somit ist $C(D, \mathbb{R}^m)$ ein reeller Vektorraum. Weiter ist

$$h : D \rightarrow \mathbb{R}, \quad h(\vec{x}) = (f(\vec{x})|g(\vec{x}))$$

in $C(D, \mathbb{R})$.

Ebenso wie in HM I kann man das folgende ε - δ -Kriterium beweisen:

Satz: $f : D \rightarrow \mathbb{R}^m$ ist stetig in $\vec{x}_0 \in D$ genau dann, wenn gilt:

$$\forall \varepsilon > 0 \exists \delta > 0 \forall \vec{x} \in D : \|\vec{x} - \vec{x}_0\| < \delta \Rightarrow \|f(\vec{x}) - f(\vec{x}_0)\| < \varepsilon.$$

Ebenso wie in HM I gilt:

Satz: Ist $\emptyset \neq D \subseteq \mathbb{R}^n$ kompakt und $f \in C(D, \mathbb{R}^m)$, so ist $f(D)$ kompakt. Im Fall $m = 1$, existieren dann $\vec{x}_1, \vec{x}_2 \in D$ mit

$$\forall \vec{x} \in D : f(\vec{x}_1) \leq f(x) \leq f(\vec{x}_2).$$

Auch der Begriff “gleichmäßig stetig” kann wie in HM I definiert werden und es gilt auch hier das zugehörige ε - δ -Kriterium:

Definition: Die Funktion f heißt *gleichmäßig stetig in D* , falls für alle Folgen $(\vec{x}_k)_{k \in \mathbb{N}}$ und $(\vec{y}_k)_{k \in \mathbb{N}}$ in D mit $\vec{x}_k - \vec{y}_k \rightarrow 0$ gilt: $f(\vec{x}_k) - f(\vec{y}_k) \rightarrow 0$. Gleichbedeutend ist

$$\forall \varepsilon > 0 \exists \delta > 0 \forall \vec{x}, \vec{y} \in D : \|\vec{x} - \vec{y}\| < \delta \Rightarrow \|f(\vec{x}) - f(\vec{y})\| < \varepsilon.$$

19.4. Differenzierbarkeit von Funktionen einer Veränderlichen: Es sei $I \subseteq \mathbb{R}$ ein Intervall und

$$f : I \rightarrow \mathbb{R}^n, \quad t \mapsto f(t) = \begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix}$$

eine Funktion.

Definition: Die Funktion f heißt *differenzierbar in* $t_0 \in I$, falls es einen Vektor $\vec{a} \in \mathbb{R}^n$ gibt mit

$$\lim_{t \rightarrow t_0} \frac{f(t) - f(t_0)}{t - t_0} = \vec{a}.$$

In diesem Fall heißt $\vec{a}$ *Ableitung von* f *in* t_0 , geschrieben $f'(t_0) := \vec{a}$.

f heißt *differenzierbar in* I , falls f in **jedem** $t_0 \in I$ differenzierbar ist, und f heißt *stetig differenzierbar in* I oder *eine* C^1 -*Funktion*, falls $f' : I \rightarrow \mathbb{R}^n$ zusätzlich stetig ist. Schreibweise: $f \in C^1(I, \mathbb{R}^n)$.

Bemerkung: Wegen 19.1 ist f differenzierbar in t_0 [bzw. in I] genau dann, wenn alle Komponentenfunktionen $f_1, \dots, f_n$ differenzierbar in t_0 [bzw. in I] sind. Es gilt dann

$$f'(t_0) = \begin{pmatrix} f'_1(t_0) \\ \vdots \\ f'_n(t_0) \end{pmatrix}, \quad [\text{bzw. } f' = \begin{pmatrix} f'_1 \\ \vdots \\ f'_n \end{pmatrix} \text{ auf } I].$$

Außerdem ist $f' : I \rightarrow \mathbb{R}^n$ stetig genau dann, wenn alle $f'_1, \dots, f'_n : I \rightarrow \mathbb{R}$ stetig sind.

Beispiele: Sei

$$f : I \rightarrow \mathbb{R}^2, \quad f(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

gegeben durch

(1) $x(t) = r \cos t$, $y(t) = r \sin t$ mit $r > 0$ und $I = [0, 2\pi]$ (hier ist $f(I) = \partial K((0, 0), r)$). Dann gilt $f \in C^1(I, \mathbb{R}^2)$ und

$$f'(t) = \begin{pmatrix} -r \sin t \\ r \cos t \end{pmatrix} \quad (t \in I).$$

(2) $x(t) = e^{-t} \cos t$, $y(t) = e^{-t} \sin t$ mit $I = [0, \infty)$ (f heißt auch *logarithmische Spirale*). Dann ist $f \in C^1(I, \mathbb{R}^2)$ und

$$f'(t) = \begin{pmatrix} -e^{-t} \cos t - e^{-t} \sin t \\ -e^{-t} \sin t + e^{-t} \cos t \end{pmatrix} \quad (t \in I).$$

19.5. Raumkurven: Eine *Raumkurve* (kurz *Kurve* oder auch C^1 -Weg) ist eine Abbildung $\gamma \in C^1(I, \mathbb{R}^n)$, wobei $I \subseteq \mathbb{R}$ ein Intervall ist. Meist ist I von der Form $[a, b]$. Das Bild $\gamma(I) \subseteq \mathbb{R}^n$ heißt hier auch *Spur von γ* .

Sind $\gamma_1 : I \rightarrow \mathbb{R}^n$ und $\gamma_2 : J \rightarrow \mathbb{R}^n$ Kurven, so heißt γ_2 eine *Umparametrisierung von γ_1* , falls es eine stetig differenzierbare und bijektive Abbildung $\phi : I \rightarrow J$ gibt mit

$$\phi'(t) \neq 0 \quad (t \in I) \quad \text{und} \quad \gamma_1(t) = \gamma_2(\phi(t)) \quad (t \in I).$$

Nach dem ZWS ist dann $\phi'(t)$ stets > 0 oder stets < 0 , also ϕ streng monoton wachsend oder streng monoton fallend, und $\phi^{-1} : J \rightarrow I$ ist ebenfalls stetig differenzierbar.

Weiter haben dann γ_1 und γ_2 dieselbe Spur. Die Bemerkung in 19.4 und die Kettenregel aus HM I zeigen, dass dann gilt:

$$\gamma_1'(t) = \gamma_2'(\phi(t))\phi'(t) \quad (t \in I).$$

Ist ϕ streng monoton wachsend [bzw. fallend], so heißt die Umparametrisierung *orientierungserhaltend* [bzw. *orientierungsumkehrend*].

Eine Kurve $\gamma : I \rightarrow \mathbb{R}^n$ heißt *regulär*, falls $\gamma'(t) \neq \vec{0}$ ($t \in I$) gilt. Dann ist auch jede Umparametrisierung regulär.

Eine Kurve $\gamma : [a, b] \rightarrow \mathbb{R}^n$ heißt *geschlossen*, falls $\gamma(a) = \gamma(b)$ gilt, und *doppelpunktfrei*, falls $\gamma : [a, b] \rightarrow \mathbb{R}^n$ injektiv ist.

Beispiele: (1) Seien $\gamma_1, \gamma_2 : [0, 1] \rightarrow \mathbb{R}^2$ gegeben durch

$$\gamma_1(t) = \begin{pmatrix} t \\ t \end{pmatrix}, \quad \gamma_2(t) = \begin{pmatrix} t^2 \\ t^2 \end{pmatrix}.$$

Dann ist γ_1 regulär. Wegen $\gamma_2'(0) = \vec{0}$ ist γ_2 nicht regulär. Insbesondere ist γ_2 keine Umparametrisierung von γ_1 , obwohl γ_1 und γ_2 dieselbe Spur haben.

(2) Sei

$$\gamma_1 : [0, 2\pi] \rightarrow \mathbb{R}^2, \quad \gamma_1(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \quad \text{und} \quad \gamma_2 : [2\pi, 8\pi] \rightarrow \mathbb{R}^2, \quad \gamma_2(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}.$$

Dann haben γ_1 und γ_2 dieselbe Spur, sind regulär und geschlossen. γ_1 ist doppelpunktfrei, aber γ_2 durchläuft das Bild von γ_1 dreimal, ist also nicht doppelpunktfrei.

19.6. Bogenlänge von Raumkurven: Ist $\gamma : [a, b] \rightarrow \mathbb{R}^n$ eine Kurve, so ist ihre *Länge* (auch *Bogenlänge* genannt) definiert durch

$$L(\gamma) := \int_a^b \|\gamma'(t)\| dt.$$

Bemerkungen: (a) $L(\gamma)$ ändert sich nicht unter Umparametrisierung.

(b) Eine reguläre Kurve $\gamma : [a, b] \rightarrow \mathbb{R}^n$ lässt sich nach der Bogenlänge s parametrisieren:
Für

$$\psi(t) := \int_a^t \|\gamma'(\tau)\| d\tau \quad (t \in [a, b]),$$

gilt $\psi \in C^1([a, b], \mathbb{R})$ und $\psi'(t) = \|\gamma'(t)\| > 0$ ($t \in [a, b]$). Sei $\phi := \psi^{-1} : [0, L(\gamma)] \rightarrow [a, b]$.
Für $\tilde{\gamma} := \gamma \circ \phi : [0, L(\gamma)] \rightarrow \mathbb{R}^n$ gilt

$$\|\tilde{\gamma}'(s)\| = \|\gamma'(\phi(s))\phi'(s)\| = \left\| \frac{\gamma'(\psi^{-1}(s))}{\psi'(\psi^{-1}(s))} \right\| = \frac{\|\gamma'(\psi^{-1}(s))\|}{\|\gamma'(\psi^{-1}(s))\|} = 1 \quad (s \in [0, L(\gamma)]).$$

Diese Parametrisierung heißt auch *natürliche Parametrisierung*, und es gilt

$$s = \int_0^s \|\tilde{\gamma}'(\tau)\| d\tau \quad (s \in [0, L(\gamma)]).$$

Beispiele: Die Abbildungen aus Beispiel 19.4 (1) und (2) sind Kurven.

(1) Sei $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2$, $\gamma(t) = \begin{pmatrix} r \cos t \\ r \sin t \end{pmatrix}$. Dann gilt (vgl. 19.4):

$$L(\gamma) = \int_0^{2\pi} \left\| \begin{pmatrix} -r \sin t \\ r \cos t \end{pmatrix} \right\| dt = \int_0^{2\pi} r dt = 2\pi r.$$

Das ist der Umfang eines Kreises mit Radius r . Die natürliche Parametrisierung ist wegen

$$s = \psi(t) = \int_0^t \|\gamma'(\tau)\| d\tau = \int_0^t r d\tau = rt$$

hier gegeben durch $\tilde{\gamma} : [0, 2\pi r] \rightarrow \mathbb{R}^2$, $\tilde{\gamma}(s) = \begin{pmatrix} r \cos(s/r) \\ r \sin(s/r) \end{pmatrix}$.

(2) Obige Definitionen können auf Kurven übertragen werden deren Definitionsintervall nicht kompakt ist. Wir zeigen dies anhand eines Beispiels:

Für die logarithmische Spirale

$$\gamma : [0, \infty) \rightarrow \mathbb{R}^2, \quad \gamma(t) = \begin{pmatrix} e^{-t} \cos t \\ e^{-t} \sin t \end{pmatrix}$$

berechnen wir unter der Verwendung der Formel für $\gamma'(t)$ aus Beispiel 19.4 (2):

$$L(\gamma) = \int_0^\infty e^{-t} \sqrt{(\cos t + \sin t)^2 + (\cos t - \sin t)^2} dt = \sqrt{2} \int_0^\infty e^{-t} dt = \sqrt{2}.$$

Hier gilt für $t \geq 0$:

$$s = \psi(t) = \int_0^t \|\gamma'(\tau)\| d\tau = \sqrt{2} \int_0^t e^{-\tau} d\tau = \sqrt{2}(1 - e^{-t}).$$

Auflösung nach t ergibt

$$t = \phi(s) = -\log \frac{\sqrt{2}-s}{\sqrt{2}} = \log \frac{\sqrt{2}}{\sqrt{2}-s}.$$

Die natürliche Parametrisierung ist hier also

$$\tilde{\gamma} : [0, \sqrt{2}) \rightarrow \mathbb{R}^2, \quad \tilde{\gamma}(s) = \begin{pmatrix} \frac{\sqrt{2}-s}{\sqrt{2}} \cos(\log \frac{\sqrt{2}}{\sqrt{2}-s}) \\ \frac{\sqrt{2}-s}{\sqrt{2}} \sin(\log \frac{\sqrt{2}}{\sqrt{2}-s}) \end{pmatrix}.$$

19.7. Richtungsableitungen: Es sei $D \subseteq \mathbb{R}^n$ offen und $f : D \rightarrow \mathbb{R}^m$ eine Funktion.

Definition: f heißt in $\vec{x}_0 \in D$ in Richtung $\vec{v} \in \mathbb{R}^n \setminus \{\vec{0}\}$ differenzierbar, falls der Limes

$$\frac{\partial f}{\partial \vec{v}}(\vec{x}_0) := \lim_{t \rightarrow 0} \frac{f(\vec{x}_0 + t\vec{v}) - f(\vec{x}_0)}{t}$$

in $\mathbb{R}^m$ existiert. $\frac{\partial f}{\partial \vec{v}}(\vec{x}_0)$ heißt *Richtungsableitung von f in $\vec{x}_0$ in Richtung $\vec{v}$* .

Bemerkung: Da D offen ist, gibt es $\delta > 0$ mit $\vec{x}_0 + t\vec{v} \in D$ für $t \in (-\delta, \delta)$. Setzt man $g(t) := f(\vec{x}_0 + t\vec{v})$ ($|t| < \delta$), so ist $\frac{\partial f}{\partial \vec{v}}(\vec{x}_0) = g'(0)$ (vgl. 19.4).

19.8. Partielle Ableitungen: Richtungsableitungen von f in Richtung der Einheitsvektoren $\vec{e}_1, \dots, \vec{e}_n$ heißen *partielle Ableitungen von f , d.h.*

$$f_{x_k}(\vec{x}_0) := \frac{\partial f}{\partial x_k}(\vec{x}_0) := \frac{\partial f}{\partial \vec{e}_k}(\vec{x}_0) \quad \text{partielle Ableitung von } f \text{ nach } x_k \text{ im Punkt } \vec{x}_0$$

für $k = 1, \dots, n$.

Beispiele: (1) $D = \{(x, y, z) \in \mathbb{R}^3 : z > 0\}$ ist offen. Betrachte

$$f : D \rightarrow \mathbb{R}, \quad f(x, y, z) = e^{-x} \cos y + \log z.$$

In jedem Punkt $(x, y, z) \in D$ existieren alle partiellen Ableitungen, und es gilt

$$\frac{\partial f}{\partial x}(x, y, z) = -e^{-x} \cos y, \quad \frac{\partial f}{\partial y}(x, y, z) = -e^{-x} \sin y, \quad \frac{\partial f}{\partial z}(x, y, z) = \frac{1}{z}.$$

(2) Betrachte

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

Wir untersuchen Richtungsableitungen im Punkt $(0, 0)$: Sei $\vec{v} = (\xi, \eta) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ eine Richtung. Es gilt für $t \neq 0$:

$$\frac{f((0, 0) + t(\xi, \eta)) - f(0, 0)}{t} = \frac{f(t\xi, t\eta)}{t} = \frac{1}{t} \frac{\xi\eta}{\xi^2 + \eta^2}.$$

Der Limes für $t \rightarrow 0$ existiert in $\mathbb{R}$ genau dann, wenn $\xi\eta = 0$ ist, d.h. genau dann, wenn $\xi = 0$ oder $\eta = 0$ ist. Also existiert $\frac{\partial f}{\partial \vec{v}}(0, 0)$ genau dann, wenn $\vec{v}$ ein Vielfaches von $\vec{e}_1$ oder ein Vielfaches von $\vec{e}_2$ ist. Insbesondere gilt $f_x(0, 0) = f_y(0, 0) = 0$.

(3) Betrachte

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = \begin{cases} \frac{x|x|^{1/2}|y|^{3/2}}{x^2 + y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}.$$

Wir untersuchen Richtungsableitungen im Punkt $(0, 0)$. Für $(\xi, \eta) \neq (0, 0)$ gilt hier

$$\frac{f((0, 0) + t(\xi, \eta)) - f(0, 0)}{t} = \frac{f(t\xi, t\eta)}{t} = \frac{t^3 \xi |\xi|^{1/2} |\eta|^{3/2}}{t^3 (\xi^2 + \eta^2)} = \frac{\xi |\xi|^{1/2} |\eta|^{3/2}}{\xi^2 + \eta^2}$$

und wir erhalten für $t \rightarrow 0$:

$$\frac{\partial f}{\partial (\xi, \eta)}(0, 0) = \frac{\xi |\xi|^{1/2} |\eta|^{3/2}}{\xi^2 + \eta^2}$$

für jede Richtung $(\xi, \eta) \neq (0, 0)$.

Übung: Existiert in $\vec{x}_0$ die Richtungsableitung von f in Richtung $\vec{v}$ und ist $\alpha \in \mathbb{R} \setminus \{0\}$, so gilt

$$\frac{\partial f}{\partial (\alpha \vec{v})}(\vec{x}_0) = \alpha \frac{\partial f}{\partial \vec{v}}(\vec{x}_0).$$

19.9. Differenzierbarkeit: Sei $D \subseteq \mathbb{R}^n$ offen und $f = (f_1, \dots, f_m) : D \rightarrow \mathbb{R}^m$ eine Funktion, sowie $\vec{x}_0 \in D$. Idee der Differenzierbarkeit in 19.4 (und in HM I) war im Fall $n = 1$:

$$f(\vec{x}) \approx f(\vec{x}_0) + f'(\vec{x}_0)(\vec{x} - \vec{x}_0) \quad \text{für } \vec{x} \text{ nahe } \vec{x}_0.$$

Dabei ist $f'(\vec{x}_0) \in \mathbb{R}^m = \mathbb{R}^{m \times 1}$, d.h. $h \mapsto f'(\vec{x}_0)h$ ist eine lineare Abbildung $\mathbb{R} \rightarrow \mathbb{R}^m$.

Definition: f heißt *differenzierbar* in $\vec{x}_0 \in D$ (manchmal auch *total differenzierbar* in $\vec{x}_0$), falls es eine lineare Abbildung $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ (also eine Matrix $A \in \mathbb{R}^{m \times n}$) gibt mit

$$\frac{\|f(\vec{x}) - f(\vec{x}_0) - A(\vec{x} - \vec{x}_0)\|}{\|\vec{x} - \vec{x}_0\|} \rightarrow 0 \quad (\vec{x} \rightarrow \vec{x}_0).$$

Äquivalent ist

$$\frac{\|f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) - A\vec{h}\|}{\|\vec{h}\|} \rightarrow 0 \quad (\vec{h} \rightarrow 0).$$

In diesem Fall ist die lineare Abbildung A eindeutig bestimmt und heißt *Ableitung von f in $\vec{x}_0$* , Bezeichnung: $f'(\vec{x}_0) := A$. Andere Bezeichnungen: $Df(\vec{x}_0)$, $J_f(\vec{x}_0)$, *Jacobimatrix*, *Funktionalmatrix*.

f heißt *differenzierbar in D* , falls f in **jedem** $\vec{x}_0 \in D$ differenzierbar ist. In diesem Fall ist f' eine Funktion von D nach $\mathbb{R}^{m \times n}$:

$$f' : D \rightarrow \mathbb{R}^{m \times n}, \quad \vec{x} \mapsto f'(\vec{x}).$$

Satz: (a) Ist f differenzierbar in $\vec{x}_0$, so ist f stetig in $\vec{x}_0$.

(b) Ist f differenzierbar in $\vec{x}_0$, so existieren in $\vec{x}_0$ alle Richtungsableitungen von f und es gilt

$$\frac{\partial f}{\partial \vec{v}}(\vec{x}_0) = f'(\vec{x}_0)\vec{v} \quad (\vec{v} \in \mathbb{R}^n \setminus \{\vec{0}\}).$$

Bemerkungen: (1) Insbesondere ist damit

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_k}(\vec{x}_0) \\ \vdots \\ \frac{\partial f_m}{\partial x_k}(\vec{x}_0) \end{pmatrix} = \frac{\partial f}{\partial x_k}(\vec{x}_0) = f'(\vec{x}_0)\vec{e}_k$$

die k -te Spalte von $f'(\vec{x}_0)$ ($k = 1, \dots, n$), also

$$f'(\vec{x}_0) = \left(\frac{\partial f_j}{\partial x_k}(\vec{x}_0) \right)_{j=1, k=1}^{m \quad n} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\vec{x}_0) & \frac{\partial f_1}{\partial x_2}(\vec{x}_0) & \dots & \frac{\partial f_1}{\partial x_n}(\vec{x}_0) \\ \frac{\partial f_2}{\partial x_1}(\vec{x}_0) & \frac{\partial f_2}{\partial x_2}(\vec{x}_0) & \dots & \frac{\partial f_2}{\partial x_n}(\vec{x}_0) \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\vec{x}_0) & \frac{\partial f_m}{\partial x_2}(\vec{x}_0) & \dots & \frac{\partial f_m}{\partial x_n}(\vec{x}_0) \end{pmatrix}.$$

(2) Im Fall $m = 1$ ist also

$$f'(\vec{x}_0) = \left(\frac{\partial f}{\partial x_1}(\vec{x}_0) \quad \frac{\partial f}{\partial x_2}(\vec{x}_0) \quad \dots \quad \frac{\partial f}{\partial x_n}(\vec{x}_0) \right) \in \mathbb{R}^{1 \times n}.$$

In diesem Fall schreibt man $f'(\vec{x}_0)$ manchmal auch als Zeilenvektor

$$f'(\vec{x}_0) = \left(\frac{\partial f}{\partial x_1}(\vec{x}_0), \dots, \frac{\partial f}{\partial x_n}(\vec{x}_0) \right).$$

Beweis. (a) Es gilt für $\vec{x} \in D \setminus \{\vec{x}_0\}$:

$$f(\vec{x}) - f(\vec{x}_0) = \frac{f(\vec{x}) - f(\vec{x}_0) - f'(\vec{x}_0)(\vec{x} - \vec{x}_0)}{\|\vec{x} - \vec{x}_0\|} \|\vec{x} - \vec{x}_0\| + f'(\vec{x}_0)(\vec{x} - \vec{x}_0) \rightarrow \vec{0}$$

für $\vec{x} \rightarrow \vec{x}_0$, da die lineare Abbildung $f'(\vec{x}_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ stetig ist.

(b) Für kleine $|t| \neq 0$ haben wir

$$\left\| \frac{f(\vec{x}_0 + t\vec{v}) - f(\vec{x}_0)}{t} - f'(\vec{x}_0)\vec{v} \right\| = \frac{\|f(\vec{x}_0 + t\vec{v}) - f(\vec{x}_0) - f'(\vec{x}_0)(t\vec{v})\|}{\|t\vec{v}\|} \|\vec{v}\| \rightarrow 0$$

für $t \rightarrow 0$. □

Beispiele: (1) Die Funktion f aus Beispiel 19.8 (2) ist **nicht** differenzierbar in $(0, 0)$, da in $(0, 0)$ nicht alle Richtungsableitungen existieren.

(2) Die Funktion f aus Beispiel 19.8 (3) ist **nicht** differenzierbar in $(0, 0)$ (obwohl in $(0, 0)$ alle Richtungsableitungen existieren): Es gilt $f_x(0, 0) = 0 = f_y(0, 0)$. Wäre f differenzierbar in $(0, 0)$, so wäre $f'(0, 0) = (0, 0)$ und somit

$$\frac{\partial f}{\partial(1, 1)}(0, 0) = f'(0, 0) \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0.$$

In Beispiel 19.8 (3) ist aber

$$\frac{\partial f}{\partial(1, 1)}(0, 0) = \frac{1}{2} \neq 0.$$

(3) Die Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = \begin{cases} \frac{x|x|^{1/2}y^2}{x^2+y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$$

ist in $(0, 0)$ differenzierbar: Zunächst berechnen wir $f_x(0, 0) = 0$ und $f_y(0, 0) = 0$. Unser Kandidat für $A \in \mathbb{R}^{1 \times 2}$ ist also $A := (0, 0)$. Es gilt: Für $(x, y) \neq (0, 0)$ ist (bea. $|xy| \leq 2|xy| \leq x^2 + y^2$)

$$\begin{aligned} \frac{\|f(x, y) - f(0, 0) - A \begin{pmatrix} x \\ y \end{pmatrix}\|}{\|(x, y) - (0, 0)\|} &= \frac{|f(x, y)|}{\sqrt{x^2 + y^2}} = \frac{|x|^{3/2}y^2}{(x^2 + y^2)^{3/2}} = \frac{|xy|^{3/2}}{(x^2 + y^2)^{3/2}} |y|^{1/2} \\ &\leq \frac{(x^2 + y^2)^{3/2}}{(x^2 + y^2)^{3/2}} |y|^{1/2} = |y|^{1/2} \rightarrow 0 \quad ((x, y) \rightarrow (0, 0)). \end{aligned}$$

Somit ist f in $(0, 0)$ differenzierbar, und $f'(0, 0) = (0, 0)$.

(4) Es sei $B \in \mathbb{R}^{m \times n}$ und $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ definiert durch $f(\vec{x}) = B\vec{x}$. Für $\vec{x}, \vec{h} \in \mathbb{R}^n$ gilt

$$f(\vec{x} + \vec{h}) - f(\vec{x}) - B\vec{h} = B(\vec{x} + \vec{h}) - B\vec{x} - B\vec{h} = B\vec{x} + B\vec{h} - B\vec{x} - B\vec{h} = \vec{0}.$$

Wir wählen also $A := B$ in der Definition und erhalten, dass f auf $\mathbb{R}^n$ differenzierbar ist mit $f'(\vec{x}) = B$ ($\vec{x} \in \mathbb{R}^n$). Insbesondere ist $f' : \mathbb{R}^n \rightarrow \mathbb{R}^{m \times n}$ konstant.

19.10. Kriterium für Differenzierbarkeit, stetige Differenzierbarkeit: Sei $D \subseteq \mathbb{R}^n$ offen und $f = (f_1, \dots, f_m) : D \rightarrow \mathbb{R}^m$ eine Funktion.

Definition: f heißt in D *partiell differenzierbar*, falls alle partiellen Ableitungen

$$\frac{\partial f_j}{\partial x_k} \quad (k \in \{1, \dots, n\}, j \in \{1, \dots, m\})$$

auf D existieren. f heißt *stetig partiell differenzierbar* in D , falls alle partiellen Ableitungen zusätzlich auf D stetig sind.

Satz: Es sei f in D partiell differenzierbar und $\vec{x}_0 \in D$. Sind alle partiellen Ableitungen $\frac{\partial f_j}{\partial x_k}$ stetig in $\vec{x}_0$, so ist f in $\vec{x}_0$ differenzierbar. (Ohne Beweis.)

Folgerung: Eine stetig partiell differenzierbare Funktion $f : D \rightarrow \mathbb{R}^m$ ist also insbesondere in D differenzierbar und

$$f' : D \rightarrow \mathbb{R}^{m \times n}$$

ist stetig. Daher heißt f dann *stetig differenzierbar*, geschrieben $f \in C^1(D, \mathbb{R}^m)$.

Beispiel: Für die Funktion f aus Beispiel 19.8 (2) gilt

$$f_x(x, y) = \frac{y(x^2 + y^2) - 2x^2y}{(x^2 + y^2)^2}, \quad f_y(x, y) = \frac{x(x^2 + y^2) - 2xy^2}{(x^2 + y^2)^2}$$

auf $\mathbb{R}^2 \setminus \{(0, 0)\}$. Also ist $f \in C^1(\mathbb{R}^2 \setminus \{(0, 0)\}, \mathbb{R})$.

19.11. Partielle Ableitungen höherer Ordnung: Sei $D \subseteq \mathbb{R}^n$ offen und $f : D \rightarrow \mathbb{R}$. Existiert die partielle Ableitung $f_{x_k} = \partial f / \partial x_k$ in D , so kann $f_{x_k} : D \rightarrow \mathbb{R}$ wieder partiell differenzierbar sein. Man gelangt so gegebenenfalls zu *partiellen Ableitungen zweiter Ordnung*

$$f_{x_k x_l} := (f_{x_k})_{x_l}.$$

Entsprechend werden (falls vorhanden) partielle Ableitungen höherer Ordnung definiert.

Schreibweisen: Z.B. sind

$$\frac{\partial^2 f}{\partial x_l \partial x_k} = f_{x_k x_l}, \quad \frac{\partial^3 f}{\partial^2 x \partial y} = f_{yxx}, \quad \frac{\partial^5 f}{\partial^2 x \partial z \partial^2 y} = f_{yyzxx}$$

partielle Ableitungen der Ordnung 2, 3 und 5.

Definition: Sei $k \in \mathbb{N}$. Eine Funktion $f : D \rightarrow \mathbb{R}$ heißt *k-mal stetig (partiell) differenzierbar* in D , falls alle partiellen Ableitungen von f der Ordnung $\leq k$ auf D existieren und dort stetig sind. Bezeichnung in diesem Fall: $f \in C^k(D, \mathbb{R})$.

Eine Funktion $f = (f_1, \dots, f_m) : D \rightarrow \mathbb{R}^m$ heißt *k-mal stetig differenzierbar* in D , falls gilt:

$$f_j \in C^k(D, \mathbb{R}) \quad (j = 1, \dots, m).$$

Bezeichnung in diesem Fall: $f \in C^k(D, \mathbb{R}^m)$. Gilt $f \in C^k(D, \mathbb{R}^m)$ für alle $k \in \mathbb{N}$ so schreiben wir $f \in C^\infty(D, \mathbb{R}^m)$. Für jedes $k \in \mathbb{N} \cup \{\infty\}$ ist $C^k(D, \mathbb{R}^m)$ ein reeller Vektorraum.

Beispiel: Betrachte $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, $f(x, y, z) = x \sin(xy) + z$. Dann ist $f \in C^\infty(\mathbb{R}^3, \mathbb{R})$ und z.B.

$$f_x(x, y, z) = \sin(xy) + xy \cos(xy), \quad f_{xy}(x, y, z) = 2x \cos(xy) - x^2 y \sin(xy),$$

$$f_y(x, y, z) = x^2 \cos(xy), \quad f_{yx}(x, y, z) = 2x \cos(xy) - x^2 y \sin(xy),$$

$$f_z(x, y, z) = 1, \quad f_{zx}(x, y, z) = 0, \quad f_{xz}(x, y, z) = 0.$$

Satz von Schwarz: Ist $k \in \mathbb{N}$ und $f \in C^k(D, \mathbb{R})$, so sind partielle Ableitungen einer Ordnung $\leq k$ unabhängig von der Reihenfolge der Differentiationen. (Ohne Beweis.)

19.12. Der Gradient: Ist $D \subseteq \mathbb{R}^n$ offen und $f : D \rightarrow \mathbb{R}$ partiell differenzierbar in D , so ist der *Gradient von f in $\vec{x}_0 \in D$* der Spaltenvektor der partiellen Ableitungen in $\vec{x}_0$, also

$$\text{grad } f(\vec{x}_0) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\vec{x}_0) \\ \frac{\partial f}{\partial x_2}(\vec{x}_0) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\vec{x}_0) \end{pmatrix}.$$

Man schreibt statt $\text{grad } f(\vec{x}_0)$ auch $\nabla f(\vec{x}_0)$ (“Nabla f ”), wobei der “Vektor”

$$\nabla := \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}$$

als *Differentialoperator* verstanden wird, der auf die reellwertige Funktion f wirkt und daraus die vektorwertige Funktion $\nabla f : D \rightarrow \mathbb{R}^n$ macht.

Bemerkung (Zusammenhang mit der Ableitung): Ist $f : D \rightarrow \mathbb{R}$ in D differenzierbar, so gilt für $\vec{x}_0 \in D$:

$$\text{grad } f(\vec{x}_0) = f'(\vec{x}_0)^T \in \mathbb{R}^{n \times 1} = \mathbb{R}^n$$

(beachte $f'(\vec{x}_0) \in \mathbb{R}^{1 \times n}$). **Beachte:** Der Gradient kann an einer Stelle $\vec{x}_0$ vorhanden sein obwohl f in $\vec{x}_0$ nicht differenzierbar ist. In Beispiel 19.8(2) ist f in $\mathbb{R}^2$ partiell differenzierbar mit $\text{grad } f(\vec{0}) = \vec{0}$, aber f ist in $\vec{0}$ nicht differenzierbar (f ist in $\vec{0}$ sogar unstetig).

Satz: Sei $f : D \rightarrow \mathbb{R}$ in D differenzierbar und $\vec{x}_0 \in D$. Dann gilt für jeden Vektor $\vec{v} \in \mathbb{R}^n$ mit $\|\vec{v}\| = 1$:

$$-\|\text{grad } f(\vec{x}_0)\| \leq \frac{\partial f}{\partial \vec{v}}(\vec{x}_0) \leq \|\text{grad } f(\vec{x}_0)\|.$$

Ist zusätzlich $\text{grad } f(\vec{x}_0) \neq \vec{0}$, so gilt Gleichheit rechts genau dann, wenn

$$\vec{v} = \frac{\text{grad } f(\vec{x}_0)}{\|\text{grad } f(\vec{x}_0)\|}$$

ist (d.h. **der Gradient zeigt in Richtung des stärksten Anstiegs von f**).

Beweis. Es gilt

$$\begin{aligned} \left| \frac{\partial f}{\partial \vec{v}}(\vec{x}_0) \right| &\stackrel{19.9}{=} |f'(\vec{x}_0)\vec{v}| = |(\text{grad } f(\vec{x}_0))^T \vec{v}| = |(\text{grad } f(\vec{x}_0)) \cdot \vec{v}| \\ &\stackrel{CSU}{\leq} \|\text{grad } f(\vec{x}_0)\| \|\vec{v}\| = \|\text{grad } f(\vec{x}_0)\|. \end{aligned}$$

□

Außerdem steht $\text{grad } f(\vec{x}_0)$ senkrecht auf der *Niveaulinie/-fläche*

$$\{\vec{x} \in \mathbb{R}^n : f(\vec{x}) = f(\vec{x}_0)\}.$$

19.13. Kettenregel: Sei $D \subseteq \mathbb{R}^n$ offen und $f : D \rightarrow \mathbb{R}^m$ differenzierbar in $\vec{x}_0 \in D$. Sei $G \subseteq \mathbb{R}^m$ offen mit $f(D) \subseteq G$, und sei $g : G \rightarrow \mathbb{R}^p$ differenzierbar in $\vec{y}_0 := f(\vec{x}_0)$. Dann ist $g \circ f : D \rightarrow \mathbb{R}^p$ differenzierbar in $\vec{x}_0$, und es gilt:

$$\underbrace{(g \circ f)'(\vec{x}_0)}_{\in \mathbb{R}^{p \times n}} = \underbrace{g'(f(\vec{x}_0))}_{\in \mathbb{R}^{p \times m}} \underbrace{f'(\vec{x}_0)}_{\in \mathbb{R}^{m \times n}}$$

(“äußere Ableitung mal innere Ableitung”; Reihenfolge beachten). Ohne Beweis.

Beispiele: (1) Seien $F : \mathbb{R}^n \rightarrow \mathbb{R}$ und $\phi = (\phi_1, \phi_2, \dots, \phi_n) : I \rightarrow \mathbb{R}^n$ differenzierbar, wobei $I \subseteq \mathbb{R}$ ein Intervall ist. Dann ist $F \circ \phi : I \rightarrow \mathbb{R}$, $t \mapsto F(\phi(t))$ differenzierbar, und es gilt

$$(F \circ \phi)'(t) = F'(\phi(t))\phi'(t) = \sum_{j=1}^n \frac{\partial F}{\partial x_j}(\phi(t))\phi'_j(t) \quad (t \in I).$$

Ist z.B. $n = 3$, $F : \mathbb{R}^3 \rightarrow \mathbb{R}$ differenzierbar, und

$$\phi : \mathbb{R} \rightarrow \mathbb{R}^3, \quad \phi(t) = \begin{pmatrix} t^2 \\ t \\ \sin(t) \end{pmatrix},$$

so gilt für $t \mapsto \psi(t) := (F \circ \phi)(t) = F(t^2, t, \sin(t))$:

$$\psi'(t) = F_x(t^2, t, \sin(t)) \cdot 2t + F_y(t^2, t, \sin(t)) \cdot 1 + F_z(t^2, t, \sin(t)) \cdot \cos(t) \quad (t \in \mathbb{R}).$$

(2) Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ differenzierbar, $B \in \mathbb{R}^{n \times m}$ und $F : \mathbb{R}^m \rightarrow \mathbb{R}$, $F(\vec{x}) := f(B\vec{x})$. Dann gilt

$$F'(\vec{x}) = f'(B\vec{x})B \quad (x \in \mathbb{R}^m),$$

also

$$\underbrace{\nabla F(\vec{x})}_{\in \mathbb{R}^m} = F'(\vec{x})^T = B^T (f'(B\vec{x}))^T = \underbrace{B^T}_{\in \mathbb{R}^{m \times n}} \underbrace{(f'(B\vec{x}))^T}_{\in \mathbb{R}^n}.$$

19.14. Der Umkehrsatz: Sei $D \subseteq \mathbb{R}^n$ offen, $f \in C^1(D, \mathbb{R}^n)$, $\vec{x}_0 \in D$ und $f'(\vec{x}_0) \in \mathbb{R}^{n \times n}$ **regulär**. Dann gibt es offene Mengen U und V mit $\vec{x}_0 \in U \subseteq D$, $\vec{y}_0 := f(\vec{x}_0) \in V$ derart, dass $f : U \rightarrow V$ bijektiv ist und die Umkehrabbildung $f^{-1} : V \rightarrow U$ stetig differenzierbar ist mit

$$(f^{-1})'(\vec{y}) = \left(f'(f^{-1}(\vec{y})) \right)^{-1} \quad (\vec{y} \in V).$$

Ohne Beweis.

Bemerkung: Die Eigenschaft *regulär* ersetzt hier für $n > 1$ die Bedingung $f'(x_0) \neq 0$, die im Fall $n = 1$ vorausgesetzt werden muss. Die Formel für die Ableitung von f^{-1} erhält man auch aus der Kettenregel, wenn man die Gleichung

$$\vec{x} = f^{-1}(f(\vec{x})) \quad (\vec{x} \in U),$$

nach $\vec{x}$ ableitet. Das ergibt

$$I_n = (f^{-1})'(f(\vec{x}))f'(\vec{x}), \quad \text{also} \quad (f^{-1})'(f(\vec{x})) = (f'(\vec{x}))^{-1} \quad (\vec{x} \in U).$$

Achtung: Der Umkehrsatz besagt nur, dass es **lokal** eine Umkehrfunktion zu f gibt. Auf ganz D muss dies nicht gelten, selbst dann, wenn $D = \mathbb{R}^n$ und die Voraussetzung “ $f'(\vec{x})$ regulär” für jedes $\vec{x} \in \mathbb{R}^n$ erfüllt ist (eine Funktion $f \in C^1(\mathbb{R}, \mathbb{R})$ mit $f'(x) \neq 0$ ($x \in \mathbb{R}$) ist auf $\mathbb{R}$ injektiv).

Beispiele: (1) Wir betrachten die komplexe Exponentialfunktion $z \mapsto e^z$ für $z = x + iy$ als Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, also

$$f(x, y) = \begin{pmatrix} e^x \cos y \\ e^x \sin y \end{pmatrix}.$$

Es gilt $f \in C^1(\mathbb{R}^2, \mathbb{R}^2)$ mit

$$f'(x, y) = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix} \quad ((x, y) \in \mathbb{R}^2),$$

und $\det(f'(x, y)) = e^{2x} > 0$. Also ist $f'(x, y) \in \mathbb{R}^{2 \times 2}$ für jedes $(x, y) \in \mathbb{R}^2$ regulär und f ist nach dem Umkehrsatz lokal injektiv. Aber $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ist *nicht* injektiv (was wir für die komplexe Exponentialfunktion ja schon wissen). Es gilt

$$f(x, y + 2k\pi) = f(x, y) \quad ((x, y) \in \mathbb{R}^2, k \in \mathbb{Z}).$$

(2) Wir betrachten

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad f(x, y, z) = \begin{pmatrix} yz \\ xz \\ xy \end{pmatrix}.$$

Hier ist $f \in C^1(\mathbb{R}^3, \mathbb{R}^3)$ mit

$$f'(x, y, z) = \begin{pmatrix} 0 & z & y \\ z & 0 & x \\ y & x & 0 \end{pmatrix} \quad ((x, y, z) \in \mathbb{R}^3).$$

Wir betrachten $\vec{x}_0 = (1, 1, 1)$. Dann ist

$$f'(\vec{x}_0) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad \text{und} \quad \det(f'(\vec{x}_0)) = 2 \neq 0.$$

Also ist f in einer Umgebung von $\vec{x}_0$ invertierbar und für die Umkehrfunktion gilt (beachte $f(\vec{x}_0) = \vec{x}_0$):

$$(f^{-1})'(\vec{x}_0) = (f'(\vec{x}_0))^{-1} = \frac{1}{2} \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}.$$

19.15. Der Satz über implizit definierte Funktionen:

Motivation: Wir betrachten die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = 2x^3 + y$. Es gilt $f(x, y) = 0 \iff y = -2x^3$. Setzt man $g(x) := -2x^3$, so gilt:

$$\forall x \in \mathbb{R} : f(x, g(x)) = 0.$$

Man sagt:

Die Gleichung $f(x, y) = 0$ kann nach y aufgelöst werden in der Form $y = g(x)$, oder durch die Gleichung $f(x, y) = 0$ wird eine Funktion g definiert mit $f(x, g(x)) = 0$.

Betrachtet man hingegen

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = y + xy^2 - e^{xy},$$

so gelingt keine formelmäßige (also explizite) Auflösung der Gleichung $f(x, y) = 0$ nach y . Dennoch kann manchmal die Existenz einer implizit definierten Funktion g gesichert werden, also die Existenz einer Funktion g mit $f(x, g(x)) = 0$, zumindest lokal, d.h. in einer Umgebung einer Nullstelle (x_0, y_0) von f . Es gilt z.B. $f(0, 1) = 0$. Unten werden wir sehen: Es gibt $\delta, \eta > 0$ und genau eine stetig differenzierbare Funktion $g : (-\delta, \delta) \rightarrow (1 - \eta, 1 + \eta)$ mit:

$$f(x, g(x)) = 0 \quad (x \in (-\delta, \delta)) \quad \text{und} \quad g(0) = 1.$$

Satz: Sei $n > m$, $p := n - m$, $D \subseteq \mathbb{R}^n$ offen und $f = (f_1, f_2, \dots, f_m) \in C^1(D, \mathbb{R}^m)$. Wir schreiben die Vektoren in $\mathbb{R}^n$ als $(\vec{x}, \vec{y})$ mit $\vec{x} = (x_1, \dots, x_p) \in \mathbb{R}^p$ und $\vec{y} = (y_1, \dots, y_m) \in \mathbb{R}^m$, sowie

$$f'(\vec{x}, \vec{y}) = \left(\begin{array}{ccc|ccc} \frac{\partial f_1}{\partial x_1}(\vec{x}, \vec{y}) & \dots & \frac{\partial f_1}{\partial x_p}(\vec{x}, \vec{y}) & \frac{\partial f_1}{\partial y_1}(\vec{x}, \vec{y}) & \dots & \frac{\partial f_1}{\partial y_m}(\vec{x}, \vec{y}) \\ \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\vec{x}, \vec{y}) & \dots & \frac{\partial f_m}{\partial x_p}(\vec{x}, \vec{y}) & \frac{\partial f_m}{\partial y_1}(\vec{x}, \vec{y}) & \dots & \frac{\partial f_m}{\partial y_m}(\vec{x}, \vec{y}) \end{array} \right) \in \mathbb{R}^{m \times n},$$

wobei wir den linken Block als $\frac{\partial f}{\partial \vec{x}}(\vec{x}, \vec{y}) \in \mathbb{R}^{m \times p}$ und den rechten Block als $\frac{\partial f}{\partial \vec{y}}(\vec{x}, \vec{y}) \in \mathbb{R}^{m \times m}$ bezeichnen.

Ist $(\vec{x}_0, \vec{y}_0) \in D$ mit $f(\vec{x}_0, \vec{y}_0) = \vec{0}$ und $\frac{\partial f}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0)$ **regulär**, also

$$\det \frac{\partial f}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0) \neq 0,$$

so gibt es offene Umgebungen U von $\vec{x}_0$ und V von $\vec{y}_0$, sowie eine stetig differenzierbare Funktion $g : U \rightarrow V$ derart, dass für alle $(\vec{x}, \vec{y}) \in U \times V$ gilt:

$$(i) \quad \det \frac{\partial f}{\partial \vec{y}}(\vec{x}, \vec{y}) \neq 0$$

und

$$(ii) \quad f(\vec{x}, \vec{y}) = \vec{0} \iff \vec{y} = g(\vec{x}).$$

Bemerkungen: (1) Durch die Funktion $g : U \rightarrow V$ ist (lokal um $(\vec{x}_0, \vec{y}_0)$) eine eindeutige Auflösung der Gleichung $f(\vec{x}, \vec{y}) = \vec{0}$ nach $\vec{y}$ gegeben. Es gilt $g(\vec{x}_0) = \vec{y}_0$ und

$$f(\vec{x}, g(\vec{x})) = 0 \quad (\vec{x} \in U)$$

und

$$\{(\vec{x}, \vec{y}) \in U \times V : f(\vec{x}, \vec{y}) = \vec{0}\}$$

ist der Graph der Funktion $g : U \rightarrow V$.

(2) Ableitungen von g kann man nach der Kettenregel aus

$$f(\vec{x}, g(\vec{x})) = \vec{0} \quad (\vec{x} \in U),$$

berechnen: Es gilt

$$\vec{0} = \frac{\partial f}{\partial \vec{x}}(\vec{x}, g(\vec{x})) + \frac{\partial f}{\partial \vec{y}}(\vec{x}, g(\vec{x}))g'(\vec{x}) \quad (\vec{x} \in U),$$

also

$$g'(\vec{x}) = -\left(\frac{\partial f}{\partial \vec{y}}(\vec{x}, g(\vec{x}))\right)^{-1} \frac{\partial f}{\partial \vec{x}}(\vec{x}, g(\vec{x})) \quad (\vec{x} \in U).$$

Beispiele: (1) $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = y + xy^2 - e^{xy}$, $(x_0, y_0) = (0, 1)$. Es gilt $f(0, 1) = 0$ und

$$f_y(x, y) = 1 + 2xy - xe^{xy}, \quad f_y(0, 1) = 1 \neq 0.$$

Also existiert eine implizit definierte C^1 -Funktion $g : U \rightarrow V$, wobei wir $U = (-\delta, \delta)$ und $V = (1 - \eta, 1 + \eta)$ annehmen können, mit

$$g(0) = 1 \quad \text{und} \quad 0 = f(x, g(x)) \quad (x \in U).$$

Berechnung von $g'(0)$: Ableiten obiger Gleichung liefert

$$0 = f_x(x, g(x)) + f_y(x, g(x))g'(x) \quad (x \in U),$$

also (für $x = 0$)

$$0 = f_x(0, 1) + f_y(0, 1)g'(0) = f_x(0, 1) + g'(0).$$

Wegen $f_x(x, y) = y^2 - ye^{xy}$ ist $f_x(0, 1) = 0$, also $g'(0) = 0$.

(2) Betrachte

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}^2, \quad f(x, y, z) = \begin{pmatrix} \sinh(yz) + (z - x)^2 - 1 \\ \cos^2(\pi y) + z - x^2/2 \end{pmatrix}.$$

Hier ist $n = 3$, $p = 1$, $m = 2$. Sei $(x_0, y_0, z_0) = (2, 0, 1)$. Es gilt und

$$f(2, 0, 1) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{und} \quad \det \frac{\partial f}{\partial (y, z)}(2, 0, 1) = \det \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} = 1 \neq 0.$$

Also ist in einer Umgebung von $(2, 0, 1)$ eine Auflösung des Gleichungssystems

$$f(x, y, z) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

nach y und z (jeweils als Funktion von x) möglich. D.h. es gibt Funktionen $g_1, g_2 : U \rightarrow \mathbb{R}$, definiert auf einer offenen Umgebung U von 2, und eine offenen Umgebung $V \subseteq \mathbb{R}^2$ von $(0, 1)$ derart, dass für alle $(x, y, z) \in U \times V$ gilt:

$$f(x, y, z) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} g_1(x) \\ g_2(x) \end{pmatrix}.$$

Die Funktion $g = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} : U \rightarrow V$ ist dabei C^1 mit

$$g'(2) = \begin{pmatrix} g'_1(2) \\ g'_2(2) \end{pmatrix} = - \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}^{-1} \frac{\partial f}{\partial x}(2, 0, 1) = - \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

19.16. Der Satz von Taylor: Sei $D \subseteq \mathbb{R}^n$ offen, $l \in \mathbb{N}_0$, $f \in C^{l+1}(D, \mathbb{R})$ und $\vec{x}_0 \in D$.

Für $\vec{h} = (h_1, h_2, \dots, h_n)^T \in \mathbb{R}^n$ schreiben wir

$$\begin{aligned}\vec{h} \cdot \nabla &:= h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2} + \dots + h_n \frac{\partial}{\partial x_n} = \sum_{j=1}^n h_j \frac{\partial}{\partial x_j}, \\ (\vec{h} \cdot \nabla) f(\vec{x}_0) &:= h_1 \frac{\partial f}{\partial x_1}(\vec{x}_0) + \dots + h_n \frac{\partial f}{\partial x_n}(\vec{x}_0) = \sum_{j=1}^n h_j \frac{\partial f}{\partial x_j}(\vec{x}_0) = \text{grad } f(\vec{x}_0) \cdot \vec{h}.\end{aligned}$$

und

$$(\vec{h} \cdot \nabla)^2 := \left(\sum_{j=1}^n h_j \frac{\partial}{\partial x_j} \right)^2 = \sum_{j,k=1}^n h_j h_k \frac{\partial^2}{\partial x_j \partial x_k}, \quad (\vec{h} \cdot \nabla)^2 f(\vec{x}_0) := \sum_{j,k=1}^n h_j h_k \frac{\partial^2 f}{\partial x_j \partial x_k}(\vec{x}_0).$$

Allgemein definiert man für $k = 1, 2, \dots, l+1$:

$$(\vec{h} \cdot \nabla)^k := \left(\sum_{j=1}^n h_j \frac{\partial}{\partial x_j} \right)^k, \quad (\vec{h} \cdot \nabla)^k f(\vec{x}_0) := \sum_{j_1, j_2, \dots, j_k=1}^n h_{j_1} h_{j_2} \dots h_{j_k} \frac{\partial^k f}{\partial x_{j_1} \partial x_{j_2} \dots \partial x_{j_k}}(\vec{x}_0),$$

und

$$(\vec{h} \cdot \nabla)^0 f(\vec{x}_0) := f(\vec{x}_0).$$

Beispiel: Für $n = 2$, $\vec{h} = (h_1, h_2)$ und $f \in C^2(D, \mathbb{R})$ hat man also

$$(\vec{h} \cdot \nabla)^2 f = h_1^2 f_{xx} + h_1 h_2 f_{xy} + h_2 h_1 f_{yx} + h_2^2 f_{yy} = h_1^2 f_{xx} + 2h_1 h_2 f_{xy} + h_2^2 f_{yy}.$$

Für eine Darstellung von $(\vec{h} \cdot \nabla)^2 f(\vec{x}_0)$ als quadratische Form definieren wir die *Hesse-Matrix von f in $\vec{x}_0$* :

$$H_f(\vec{x}_0) := \begin{pmatrix} f_{x_1 x_1}(\vec{x}_0) & \dots & f_{x_1 x_n}(\vec{x}_0) \\ f_{x_2 x_1}(\vec{x}_0) & \dots & f_{x_2 x_n}(\vec{x}_0) \\ \vdots & & \vdots \\ f_{x_n x_1}(\vec{x}_0) & \dots & f_{x_n x_n}(\vec{x}_0) \end{pmatrix} = \left(\frac{\partial^2 f}{\partial x_j \partial x_k}(\vec{x}_0) \right)_{j,k=1}^n.$$

Damit ist

$$(\vec{h} \cdot \nabla)^2 f(\vec{x}_0) = \vec{h}^T H_f(\vec{x}_0) \vec{h}.$$

Beachte: Für $f \in C^2(D, \mathbb{R})$ ist nach dem Satz von Schwarz $H_f(\vec{x}_0)$ symmetrisch.

Für $\vec{x}, \vec{y} \in \mathbb{R}^n$ bezeichne

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

die *Verbindungsstrecke* von $\vec{x}$ und $\vec{y}$.

Satz von Taylor: Unter den obigen Voraussetzungen seien $\vec{x}_0 \in D$ und $\vec{h} \in \mathbb{R}^n$ mit $S[\vec{x}_0, \vec{x}_0 + \vec{h}] \subseteq D$. Dann gibt es ein $\vec{\xi} \in S[\vec{x}_0, \vec{x}_0 + \vec{h}]$ mit

$$f(\vec{x}_0 + \vec{h}) = \sum_{k=0}^l \frac{(\vec{h} \cdot \nabla)^k f(\vec{x}_0)}{k!} + \frac{(\vec{h} \cdot \nabla)^{l+1} f(\vec{\xi})}{(l+1)!}.$$

Bemerkungen: (a) Der Ausdruck

$$T_l(f; \vec{x}_0)(\vec{x}) := \sum_{k=0}^l \frac{((\vec{x} - \vec{x}_0) \cdot \nabla)^k f(\vec{x}_0)}{k!}$$

heißt *l-tes Taylorpolynom von f in $\vec{x}_0$* . Hier ist $\vec{h} = \vec{x} - \vec{x}_0$.

(b) Für $l = 0$, $f \in C^1(D, \mathbb{R})$ erhält man unter obigen Voraussetzungen den mehrdimensionalen Mittelwertsatz:

$$\exists \vec{\xi} \in S[\vec{x}_0, \vec{x}_0 + \vec{h}] : f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) = \text{grad } f(\vec{\xi}) \cdot \vec{h}.$$

(c) Für $l = 1$, $f \in C^2(D, \mathbb{R})$ erhalten wir, wenn $\vec{h} \in \mathbb{R}^n$ mit $S[\vec{x}_0, \vec{x}_0 + \vec{h}] \subseteq D$:

$$\exists \vec{\xi} \in S[\vec{x}_0, \vec{x}_0 + \vec{h}] : f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + \text{grad } f(\vec{x}_0) \cdot \vec{h} + \frac{1}{2} \vec{h}^T H_f(\vec{\xi}) \vec{h}.$$

Schreibt man $\vec{x} - \vec{x}_0$ statt $\vec{h}$, so ist $\vec{x}_0 + \vec{h} = \vec{x}$.

Beispiel: Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = e^{xe^y-2}$ und $(x_0, y_0) = (2, 0)$. Dann ist $f \in C^2(\mathbb{R}^2, \mathbb{R})$ und

$$\begin{aligned} f_x(x, y) &= e^y f(x, y), \quad f_y(x, y) = x e^y f(x, y), \quad f_{xx}(x, y) = e^{2y} f(x, y), \\ f_{xy}(x, y) &= e^y (f(x, y) + f_y(x, y)) = (e^y + x e^{2y}) f(x, y), \\ f_{yy}(x, y) &= x e^y (f(x, y) + f_y(x, y)) = (x e^y + x^2 e^{2y}) f(x, y). \end{aligned}$$

Wir erhalten

$$f(2, 0) = 1, \quad f_x(2, 0) = 1, \quad f_y(2, 0) = 2, \quad f_{xx}(2, 0) = 1, \quad f_{xy}(2, 0) = 3, \quad f_{yy}(2, 0) = 6,$$

und damit ist das zweite Taylorpolynom von f in $(2, 0)$ gegeben durch

$$\begin{aligned} T_2(f; (2, 0))(x, y) &= f(2, 0) + \text{grad } f(2, 0) \cdot \begin{pmatrix} x-2 \\ y \end{pmatrix} + \frac{1}{2} \begin{pmatrix} x-2 \\ y \end{pmatrix}^T H_f(2, 0) \begin{pmatrix} x-2 \\ y \end{pmatrix} \\ &= 1 + (x-2) + 2y + \frac{1}{2}(x-2)^2 + 3(x-2)y + 3y^2. \end{aligned}$$

19.17. Lokale Extremstellen: Ist $D \subseteq \mathbb{R}^n$ (beliebig) und $f : D \rightarrow \mathbb{R}$, so definiert man lokale und globale Extremstellen wörtlich wie in HM I 11.5. Wir wiederholen die Definition der lokalen Extremstellen in diesem Fall:

Definition: f hat in $\vec{x}_0 \in D$ ein *lokales Maximum* [bzw. *lokales Minimum*], falls es ein $\delta > 0$ so gibt, dass

$$\forall \vec{x} \in K(\vec{x}_0, \delta) \cap D : f(\vec{x}) \leq f(\vec{x}_0) \text{ [bzw. } f(\vec{x}) \geq f(\vec{x}_0)].$$

Im Fall, dass D offen ist kann dann stets ein $\delta > 0$ mit $K(\vec{x}_0, \delta) \subseteq D$ gewählt werden.

Ein *lokales Extremum* ist ein lokales Maximum oder ein lokales Minimum. Ein lokales Extremum heißt *strikt*, wenn für $\vec{x} \neq \vec{x}_0$ jeweils “<” bzw. “>” gilt.

Satz über lokale Extremstellen: Es sei $D \subseteq \mathbb{R}^n$ offen, $f : D \rightarrow \mathbb{R}$ und $\vec{x}_0 \in D$.

(a) Hat f in $\vec{x}_0$ ein lokales Extremum und ist f in $\vec{x}_0$ partiell differenzierbar, so ist $\text{grad } f(\vec{x}_0) = \vec{0}$.

(b) Sei $f \in C^2(D, \mathbb{R})$ und $\text{grad } f(\vec{x}_0) = \vec{0}$. Dann gilt:

- (i) Ist $H_f(\vec{x}_0)$ positiv definit, so hat f in $\vec{x}_0$ ein striktes lokales Minimum.
- (ii) Ist $H_f(\vec{x}_0)$ negativ definit, so hat f in $\vec{x}_0$ ein striktes lokales Maximum.
- (iii) Ist $H_f(\vec{x}_0)$ indefinit, so hat f in $\vec{x}_0$ **kein** lokales Extremum, sondern einen *Sattelpunkt*.

Bemerkungen: (1) Nullstellen des Gradienten heißen auch *kritische Punkte*.

(2) Trifft in (b) keiner der Fälle (i), (ii) oder (iii) zu, so ist $H_f(\vec{x}_0)$ (positiv oder negativ) **semidefinit** (vgl. 18.9). In diesem Fall ist keine allgemeine Aussage möglich.

Beweisidee zu (b): Wegen $f \in C^2(D, \mathbb{R})$ existiert ein $\delta > 0$ mit $K(\vec{x}_0, \delta) \subseteq D$ und so, dass $H_f(\vec{x})$ für $\vec{x} \in K(\vec{x}_0, \delta)$ dieselben Definitheitseigenschaften wie $H_f(\vec{x}_0)$ hat.

Für $\vec{x} \in K(\vec{x}_0, \delta)$ gilt nun nach Bemerkung 19.16(c):

$$f(\vec{x}) = f(\vec{x}_0) + (\vec{x} - \vec{x}_0)^T H_f(\vec{\xi})(\vec{x} - \vec{x}_0)$$

für ein $\vec{\xi} \in S[\vec{x}_0, \vec{x}] \subseteq K(\vec{x}_0, \delta)$. Ist z.B. $H_f(\vec{x}_0)$ positiv definit, so ist auch $H_f(\vec{\xi})$ positiv definit und damit $f(\vec{x}) > f(\vec{x}_0)$ ($\vec{x} \in K(\vec{x}_0, \delta) \setminus \{\vec{x}_0\}$). $\square$

Beispiel für einen Sattelpunkt: Betrachte $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = xy$. Dann ist

$$\text{grad } f(x, y) = \begin{pmatrix} y \\ x \end{pmatrix} \quad \text{und} \quad \text{grad } f(0, 0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Die Hessematrix

$$H_f(0, 0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

hat eine negative Determinante und ist daher indefinit (siehe 18.9). Somit hat f in $(0, 0)$ einen Sattelpunkt. Für (x, y) im ersten oder dritten Quadranten ist $f(x, y) = xy > 0$, für (x, y) im zweiten oder vierten Quadranten ist $f(x, y) = xy < 0$.

Beispiel: Betrachte $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) = x^3 - 12xy + 8y^3$. Hier gilt

$$f_x(x, y) = 3x^2 - 12y, \quad f_y(x, y) = -12x + 24y^2.$$

Wir formen äquivalent um:

$$\begin{aligned} \text{grad } f(x, y) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} &\iff \begin{aligned} 3x^2 - 12y &= 0 \\ -12x + 24y^2 &= 0 \end{aligned} \iff \begin{aligned} x^2 &= 4y \\ 2y^2 &= x \end{aligned} \iff \begin{aligned} y^4 &= y \\ 2y^2 &= x \end{aligned} \iff \\ &\iff \begin{aligned} y &\in \{0, 1\} \\ 2y^2 &= x \end{aligned} \iff \begin{pmatrix} x \\ y \end{pmatrix} \in \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\}. \end{aligned}$$

Wir berechnen die Hessematrix

$$H_f(x, y) = \begin{pmatrix} 6x & -12 \\ -12 & 48y \end{pmatrix}.$$

Es ist

$$\det H_f(0, 0) = \det \begin{pmatrix} 0 & -12 \\ -12 & 0 \end{pmatrix} = -144 < 0,$$

also ist $H_f(0, 0)$ indefinit und f hat in $(0, 0)$ **kein** lokales Extremum (sondern einen Sattelpunkt). Weiter ist

$$H_f(2, 1) = \begin{pmatrix} 12 & -12 \\ -12 & 48 \end{pmatrix}$$

wegen $\det H_f(2, 1) > 0$ und $12 > 0$ positiv definit (siehe 18.9). Also hat f in $(2, 1)$ ein lokales Minimum. f hat in $(2, 1)$ kein globales Minimum, denn $f(x, 0) = x^3 \rightarrow -\infty$ ($x \rightarrow -\infty$).

19.18. Extremwertaufgaben mit Nebenbedingungen: Es sei $\emptyset \neq D \subseteq \mathbb{R}^n$ offen, $f \in C^1(D, \mathbb{R})$, $p \in \mathbb{N}$ mit $p < n$ und $h = (h_1, h_2, \dots, h_p) \in C^1(D, \mathbb{R}^p)$. Es sei

$$S := \{\vec{x} \in D : h(\vec{x}) = \vec{0}\}.$$

Definition: Man sagt

f hat in $\vec{x}_0 \in D$ ein lokales Maximum [Minimum] unter der Nebenbedingung $h = \vec{0}$,

falls $\vec{x}_0 \in S$ gilt und es ein $\delta > 0$ gibt mit $K(\vec{x}_0, \delta) \subseteq D$ und

$$\forall \vec{x} \in K(\vec{x}_0, \delta) \cap S : f(\vec{x}) \leq f(\vec{x}_0) \quad [\text{bzw. } f(\vec{x}) \geq f(\vec{x}_0)].$$

Bemerkung: Ist S kompakt, so existieren stets (sogar globale) Extremstellen unter der Nebenbedingung $h = \vec{0}$.

Beispiel: Seien $n = 3$, $p = 2$ und $D = \mathbb{R}^3$, sowie

$$f(x, y, z) = x + y + z, \quad h(x, y, z) = \begin{pmatrix} x^2 + y^2 - 2 \\ x + z - 1 \end{pmatrix}.$$

Dann gilt $f \in C^1(D, \mathbb{R})$, $h \in C^1(D, \mathbb{R}^2)$ und

$$S = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 2, x + z = 1\}.$$

Die Menge S ist abgeschlossen: Für eine Folge $((x_k, y_k, z_k))$ in S mit $(x_k, y_k, z_k) \rightarrow (x_0, y_0, z_0) \in \mathbb{R}^3$ gilt $(x_0, y_0, z_0) \in S$.

Die Menge S ist beschränkt: Ist $(x, y, z) \in S$, so gilt $x^2 + y^2 = 2$ und somit $|x| \leq \sqrt{2}$ und $|y| \leq \sqrt{2}$. Also ist

$$|z| = |1 - x| \leq 1 + |x| \leq 1 + \sqrt{2}$$

und insgesamt

$$\|(x, y, z)\| \leq \sqrt{4 + (1 + \sqrt{2})^2} =: M.$$

Also ist S kompakt und es existieren $\vec{a}, \vec{b} \in S$ mit

$$\forall \vec{x} \in S : f(\vec{a}) \leq f(\vec{x}) \leq f(\vec{b}),$$

vgl. 19.2 und 19.3.

19.19. Multiplikatorenregel von Lagrange: Seien n, D, f, p, h und S wie in 19.18. Es sei $F : D \times \mathbb{R}^p \rightarrow \mathbb{R}$ definiert durch

$$F(\vec{x}, \lambda_1, \lambda_2, \dots, \lambda_p) := f(\vec{x}) + \lambda_1 h_1(\vec{x}) + \lambda_2 h_2(\vec{x}) + \dots + \lambda_p h_p(\vec{x}).$$

Satz: Hat f in $\vec{x}_0 \in D$ ein lokales Extremum unter der Nebenbedingung $h = \vec{0}$ und gilt

$$\text{Rang } \underbrace{h'(\vec{x}_0)}_{\in \mathbb{R}^{p \times n}} = p \quad [h'(\vec{x}_0) \text{ hat vollen, d.h. maximalen, Rang}],$$

so gibt es $\lambda_1^0, \lambda_2^0, \dots, \lambda_p^0 \in \mathbb{R}$ (*Lagrangemultiplikatoren*) mit

$$\text{grad } F(\vec{x}_0, \lambda_1^0, \lambda_2^0, \dots, \lambda_p^0) = \vec{0} \quad (\in \mathbb{R}^{n+p}).$$

Bemerkung: (a) Beachte, dass

$$h'(\vec{x}_0) = \begin{pmatrix} \text{grad } h_1(\vec{x}_0)^T \\ \text{grad } h_2(\vec{x}_0)^T \\ \vdots \\ \text{grad } h_p(\vec{x}_0)^T \end{pmatrix}.$$

Die Voraussetzung “Rang $h'(\vec{x}_0) = p$ ” bedeutet also, dass

$$\text{grad } h_1(\vec{x}_0), \text{grad } h_2(\vec{x}_0), \dots, \text{grad } h_p(\vec{x}_0)$$

linear unabhängig sind.

(b) Schreibt man $\vec{x}_0 = (x_1^0, x_2^0, \dots, x_n^0)$, so ist die Bedingung $\text{grad } F = \vec{0}$ ein Gleichungssystem mit $n + p$ Gleichungen für die $n + p$ Unbekannten $x_1^0, x_2^0, \dots, x_n^0, \lambda_1^0, \lambda_2^0, \dots, \lambda_p^0$, nämlich

$$(1) \quad \frac{\partial F}{\partial x_k}(\vec{x}_0) = \frac{\partial f}{\partial x_k}(\vec{x}_0) + \sum_{j=1}^p \lambda_j^0 \frac{\partial h_j}{\partial x_k}(\vec{x}_0) = 0 \quad (k = 1, 2, \dots, n)$$

und

$$(2) \quad \frac{\partial F}{\partial \lambda_j}(\vec{x}_0) = h_j(\vec{x}_0) = 0 \quad (j = 1, 2, \dots, p).$$

Die Gleichungen in (1) besagen, dass $\text{grad } f(\vec{x}_0)$ eine Linearkombination der Vektoren

$$\text{grad } h_1(\vec{x}_0), \dots, \text{grad } h_p(\vec{x}_0)$$

ist (mit Koeffizienten $-\lambda_1^0, \dots, -\lambda_p^0$). Die Gleichungen in (2) sind gleichbedeutend mit $\vec{x}_0 \in S$.

(c) Zur Bestimmung der Extrema versucht man, dieses Gleichungssystem zu lösen. Kann man den Satz anwenden (d.h. sind die Voraussetzungen erfüllt), so findet man die gesuchten Extremstellen unter den Lösungen dieses Gleichungssystems. Kann man den Satz **nicht** anwenden (weil es z.B. lokale Extremstellen $\vec{x}_0$ gibt, in denen $h'(\vec{x}_0)$ nicht vollen Rang hat), so ist dies **nicht sicher!**

Fortsetzung des Beispiels aus 19.18: Wir wissen schon, dass es $\vec{a}, \vec{b} \in S$ gibt mit: f hat in $\vec{a}$ [bzw. in $\vec{b}$] ein globales Minimum [bzw. Maximum] unter der Nebenbedingung $h = \vec{0}$.

Wir haben

$$h'(x, y, z) = \begin{pmatrix} 2x & 2y & 0 \\ 1 & 0 & 1 \end{pmatrix},$$

und $\text{Rang } h'(x, y, z) < 2$ ist äquivalent zu $x = y = 0$, was jedoch für $(x, y, z) \in S$ wegen $x^2 + y^2 = 2$ nicht vorkommt. Also ist

$$\text{Rang } h'(x, y, z) = 2 \quad ((x, y, z) \in S),$$

und die Voraussetzungen des Satzes sind insbesondere in $\vec{a}$ und $\vec{b}$ erfüllt. Hier ist

$$F(x, y, z, \lambda_1, \lambda_2) = x + y + z + \lambda_1(x^2 + y^2 - 2) + \lambda_2(x + z - 1)$$

und

$$F_x(x, y, z, \lambda_1, \lambda_2) = 1 + 2\lambda_1 x + \lambda_2, \quad F_y(\dots) = 1 + 2\lambda_1 y, \quad F_z(\dots) = 1 + \lambda_2,$$

$$F_{\lambda_1}(\dots) = x^2 + y^2 - 2, \quad F_{\lambda_2}(\dots) = x + z - 1.$$

Aus $\text{grad } F = \vec{0}$ erhält man also $\lambda_2 = -1$, $\lambda_1 \neq 0$, $x = 0$, $z = 1$ und $y = \pm\sqrt{2}$ (der genaue Wert von λ_1 ist nicht wichtig). Also existieren genau zu den Stellen

$$\vec{a} := (0, -\sqrt{2}, 1), \quad \vec{b} := (0, \sqrt{2}, 1)$$

Lagrangemultiplikatoren. Wegen $f(0, \pm\sqrt{2}, 1) = 1 \pm \sqrt{2}$ ist $\vec{a}$ die Minimal- und $\vec{b}$ die Maximalstelle. Der maximale Wert von f auf S ist $f(\vec{b}) = 1 + \sqrt{2}$ und der minimale Wert von f auf S ist $f(\vec{a}) = 1 - \sqrt{2}$.

19.20. Rotation, Divergenz, Laplace: Sei $\emptyset \neq D \subseteq \mathbb{R}^n$ offen. Eine Funktion $f : D \rightarrow \mathbb{R}$ heißt *Skalarfeld (auf D)* und eine Funktion $\vec{v} : D \rightarrow \mathbb{R}^n$ heißt *Vektorfeld (auf D)*.

Bemerkung: Ist $f \in C^1(D, \mathbb{R})$ ein Skalarfeld auf D , so ist $\vec{v} := \text{grad } f = \nabla f$ ein Vektorfeld auf D .

Partielle Ableitungen schreiben wir im folgenden auch als

$$\partial_j := \frac{\partial}{\partial x_j} \quad (j = 1, 2, \dots, n).$$

Definition: Sei $\vec{v} = (v_1, v_2, \dots, v_n) \in C^1(D, \mathbb{R}^n)$ ein Vektorfeld auf D . Dann definiert man die *Divergenz von $\vec{v}$* durch

$$\text{div } \vec{v} := \nabla \cdot \vec{v} := \partial_1 v_1 + \partial_2 v_2 + \dots + \partial_n v_n = \sum_{j=1}^n \partial_j v_j,$$

und im Fall $n = 3$ die *Rotation von $\vec{v}$* durch

$$\text{rot } \vec{v} := \nabla \times \vec{v} := \begin{pmatrix} \partial_2 v_3 - \partial_3 v_2 \\ \partial_3 v_1 - \partial_1 v_3 \\ \partial_1 v_2 - \partial_2 v_1 \end{pmatrix}.$$

Bemerkungen: (a) $\text{div } \vec{v}$ ist ein Skalarfeld auf D und $\text{rot } \vec{v}$ ist ein Vektorfeld auf D .

(b) Vektorfelder mit $\text{div } \vec{v} = 0$ heißen *quellenfrei* und Vektorfelder mit $\text{rot } \vec{v} = \vec{0}$ heißen *wirbelfrei*.

(c) Für Skalarfelder $f \in C^2(D, \mathbb{R})$ definiert man den *Laplaceoperator Δ* durch

$$\Delta f := \text{div grad } f := \nabla \cdot \nabla f := \sum_{j=1}^n \partial_j^2 f,$$

und für Vektorfelder $\vec{v} \in C^2(D, \mathbb{R}^n)$ durch

$$\Delta \vec{v} := \begin{pmatrix} \Delta v_1 \\ \Delta v_2 \\ \vdots \\ \Delta v_n \end{pmatrix}.$$

Beispiele: (1) Sei $\vec{v} : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $\vec{v}(\vec{x}) := \vec{x}$. Dann gilt

$$\text{div } \vec{v}(\vec{x}) = n \quad (\vec{x} \in \mathbb{R}^n).$$

Im Falle $n = 3$ gilt $\operatorname{rot} \vec{v} = \vec{0}$ auf $\mathbb{R}^3$.

(2) Es sei

$$\vec{v} : \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad \vec{v}(x, y, z) = \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}.$$

Dann ist

$$\operatorname{rot} \vec{v}(x, y, z) = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} \quad ((x, y, z) \in \mathbb{R}^3).$$

19.21. Rechenregeln:

Produktregeln: Seien $f, g \in C^1(D, \mathbb{R})$ und $\vec{v} \in C^1(D, \mathbb{R}^n)$. Dann gilt:

$$\begin{aligned} \nabla(fg) &= g\nabla f + f\nabla g \\ \nabla \cdot (f\vec{v}) &= f(\nabla \cdot \vec{v}) + (\nabla f) \cdot \vec{v} \\ \nabla \times (f\vec{v}) &= f(\nabla \times \vec{v}) + (\nabla f) \times \vec{v} \quad (n = 3) \\ \Delta(fg) &= (\Delta f)g + 2\nabla f \cdot \nabla g + f(\Delta g) \quad (f, g \in C^2(D, \mathbb{R})). \end{aligned}$$

Zum Nachrechnen wende man die eindimensionale Produktregel aus HM I auf die einzelnen partiellen Ableitungen an:

$$\partial_j(\phi\psi) = (\partial_j\phi)\psi + \phi(\partial_j\psi).$$

Orthogonalen Koordinatentransformation: Sei $f \in C^2(\mathbb{R}^n, \mathbb{R})$ und $A \in \mathbb{R}^{n \times n}$ orthogonal (also $A^T = A^{-1}$), sowie $g := f \circ A$, d.h. $g(\vec{x}) = f(A\vec{x})$. Dann gilt $\Delta g = (\Delta f) \circ A$.

Hintereinanderausführung: Sei $D \subseteq \mathbb{R}^3$. Sind $f \in C^2(D, \mathbb{R})$ und $\vec{v} \in C^2(D, \mathbb{R}^3)$, so gilt:

$$\begin{aligned} \operatorname{rot}(\operatorname{grad} f) &= \vec{0}, & \text{d.h.} \quad \nabla \times (\nabla f) &= \vec{0} \\ \operatorname{div}(\operatorname{rot} \vec{v}) &= 0, & \text{d.h.} \quad \nabla \cdot (\nabla \times \vec{v}) &= 0 \\ \operatorname{rot}(\operatorname{rot} \vec{v}) &= \nabla \times (\nabla \times \vec{v}) = \operatorname{grad} \operatorname{div} \vec{v} - \Delta \vec{v} = \nabla(\nabla \cdot \vec{v}) - \Delta \vec{v}. \end{aligned}$$

Die beiden ersten Regeln folgen aus dem Satz von Schwarz.

19.22. Potentialfelder: Sei $\emptyset \neq D \subseteq \mathbb{R}^n$ offen. Ein Vektorfeld $\vec{v} \in C(D, \mathbb{R}^n)$ heißt *Potentialfeld* (oder *Gradientenfeld*, *konservatives Feld*), falls ein Skalarfeld $f \in C^1(D, \mathbb{R})$ existiert mit

$$\vec{v}(\vec{x}) = \operatorname{grad} f(\vec{x}) \quad (\vec{x} \in D).$$

In diesem Fall heißt f ein *Potential* oder eine *Stammfunktion* von $\vec{v}$ auf D .

Bemerkung: Ist $\vec{v} \in C^1(D, \mathbb{R}^3)$ ein Potentialfeld, so ist wegen 19.21

$$\operatorname{rot}(\vec{v}) = \operatorname{rot}(\operatorname{grad} f) = \vec{0}$$

auf D , also ist $\vec{v}$ wirbelfrei in D .

Integrabilitätsbedingungen: Sei $\vec{v} = (v_1, v_2, \dots, v_n) \in C^1(D, \mathbb{R}^n)$ ein Potentialfeld. Dann gilt auf D :

$$\partial_j v_k = \partial_k v_j \quad (j, k = 1, \dots, n).$$

Beweis. Sei $f \in C^1(D, \mathbb{R})$ eine Stammfunktion von $\vec{v}$ auf D . Wegen $\vec{v} \in C^1(D, \mathbb{R}^n)$ gilt dann $f \in C^2(D, \mathbb{R})$. Wegen $\partial_j f = v_j$ ($j = 1, \dots, n$) und nach dem Satz von Schwarz gilt auf D :

$$\partial_j v_k = \partial_j \partial_k f = \partial_k \partial_j f = \partial_k v_j.$$

□

Beispiele: (1) Betrachte $\vec{v} = (v_1, v_2) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $\vec{v}(x, y) = \begin{pmatrix} -y \\ x \end{pmatrix}$. Dann gilt $\vec{v} \in C^1(\mathbb{R}^2, \mathbb{R}^2)$. Weiter ist

$$\partial_2 v_1(x, y) = -1 \neq 1 = \partial_1 v_2(x, y).$$

Also hat $\vec{v}$ auf $\mathbb{R}^2$ keine Stammfunktion.

(2) Betrachte

$$\vec{v} = (v_1, v_2) : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \vec{v}(x, y) = \begin{pmatrix} 2xy + y^3 \\ x^2 + 3xy^2 + 1 \end{pmatrix}.$$

Hier gilt

$$\partial_2 v_1(x, y) = 2x + 3y^2 = \partial_1 v_2(x, y) \quad ((x, y) \in \mathbb{R}^2).$$

Die Integrabilitätsbedingungen sind also erfüllt.

Berechnung einer Stammfunktion. Ansatz:

$$f_x(x, y) = 2xy + y^3 \Rightarrow f(x, y) = x^2 y + xy^3 + c(y) \Rightarrow f_y(x, y) = x^2 + 3xy^2 + c'(y),$$

also $c'(y) = 1$. Wähle z.B. $c(y) = y$. Damit ist

$$f(x, y) = x^2 y + xy^3 + y$$

eine Stammfunktion von $\vec{v}$ auf $\mathbb{R}^2$.

(3) Für eine Punktladung Q im Ursprung $\vec{0} \in \mathbb{R}^3$ ist das elektrische Potential φ im Punkt $\vec{x}$ gegeben durch

$$\varphi(\vec{x}) = \frac{Q}{4\pi\epsilon_0 \|\vec{x}\|} \quad (\vec{x} \in D := \mathbb{R}^3 \setminus \{\vec{0}\}).$$

Das elektrische Feld $\vec{E}$ erhält man als $\vec{E} = -\nabla\varphi$. Also ist $\vec{E}$ ein Potentialfeld auf D mit Potential $-\varphi$. Wegen

$$\partial_j \varphi(\vec{x}) = -\frac{Q}{4\pi\epsilon_0} \frac{2x_j}{2(x_1^2 + x_2^2 + x_3^2)^{3/2}} \quad (j = 1, 2, 3),$$

ist

$$\vec{E}(\vec{x}) = \frac{Q \vec{x}}{4\pi\epsilon_0 \|\vec{x}\|^3} \quad (\vec{x} \in D).$$

Nach der Bemerkung oben ist das elektrische Feld $\vec{E}$ wirbelfrei in D .

20 Kurvenintegrale

20.1. Kurvenintegrale von Skalarfeldern: Sei $D \subseteq \mathbb{R}^n$ offen und $f \in C(D, \mathbb{R})$. Für eine Kurve $\gamma : [a, b] \rightarrow D$ (d.h. $\gamma \in C^1([a, b], \mathbb{R}^n)$ mit $\gamma([a, b]) \subseteq D$, siehe 19.5) setzt man

$$\int_{\gamma} f \, ds := \int_{\gamma} f(\vec{x}) \, ds := \int_a^b f(\gamma(t)) \|\gamma'(t)\| \, dt$$

(beachte, dass $\gamma(t)$ und $\gamma'(t)$ Vektoren aus $\mathbb{R}^n$ sind).

Bemerkungen: (a) Ist $\tilde{\gamma}$ eine orientierungserhaltende Umparametrisierung von γ (siehe 19.5), so gilt

$$\int_{\gamma} f \, ds = \int_{\tilde{\gamma}} f \, ds.$$

(b) Für $f = 1$ ist $\int_{\gamma} 1 \, ds = L(\gamma)$ (die Länge von γ , vgl. 19.6). Ist

$$s(t) = \int_a^t \|\gamma'(\tau)\| \, d\tau \quad (t \in [a, b])$$

die Kurvenlängenfunktion, so gilt

$$\frac{ds}{dt} = s'(t) = \|\gamma'(t)\|, \text{ also } "ds = \|\gamma'(t)\| \, dt."$$

(c) Ist γ wie oben und setzt man $\rho : [a, b] \rightarrow D$, $\rho(t) := \gamma(a + b - t)$, so durchläuft ρ die Spur von γ in umgekehrter Richtung. Diese Kurve wird auch mit γ^- bezeichnet. Es gilt dann

$$\int_{\gamma^-} f \, ds = \int_{\gamma} f \, ds.$$

(d) Sind $\gamma_j : [a_{j-1}, a_j] \rightarrow D$ ($j = 1, 2, \dots, m$) Kurven mit $\gamma_j(a_j) = \gamma_{j+1}(a_j)$, so nennt man auch $\gamma : [a_0, a_m] \rightarrow D$, $\gamma(t) := \gamma_j(t)$ ($t \in [a_{j-1}, a_j]$) eine *Kurve*. In den Punkten $a_1, a_2, \dots, a_{m-1}$ muss γ nicht differenzierbar sein, d.h. die rechts- und linksseitigen Ableitungen in diesen Punkten müssen nicht übereinstimmen. Man schreibt $\gamma = \gamma_1 + \gamma_2 + \dots + \gamma_m$ und definiert

$$\int_{\gamma} f \, ds := \sum_{j=1}^m \int_{\gamma_j} f \, ds.$$

(e) Ist $\gamma : [a, b] \rightarrow D$ geschlossen, d.h. $\gamma(a) = \gamma(b)$, so schreibt man auch

$$\int_{\gamma} f \, ds = \oint_{\gamma} f \, ds.$$

Beispiele: (1) Sei $r > 0$,

$$\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2, \quad \gamma(t) = \begin{pmatrix} r \cos t \\ r \sin t \end{pmatrix}$$

und $f(x, y) = 2xy$. Dann ist $\|\gamma'(t)\| = r$ ($t \in [0, 2\pi]$), also

$$\int_{\gamma} f(x, y) ds = \int_0^{2\pi} f(r \cos t, r \sin t) \|\gamma'(t)\| dt = \int_0^{2\pi} 2r^3 \cos t \sin t dt = r^3 \int_0^{2\pi} \sin(2t) dt = 0.$$

(2) Betrachte

$$\gamma_1(t) = \begin{pmatrix} t \\ 0 \end{pmatrix} \quad (t \in [0, 1]), \quad \gamma_2(t) = \begin{pmatrix} 1 \\ t - 1 \end{pmatrix} \quad (t \in [1, 2]),$$

$\gamma = \gamma_1 + \gamma_2 : [0, 2] \rightarrow \mathbb{R}^2$ und $f(x, y) = x + y$. Dann:

$$\int_{\gamma} f(x, y) ds = \int_{\gamma_1} (x + y) ds + \int_{\gamma_2} (x + y) ds = \int_0^1 t \cdot 1 dt + \int_1^2 t \cdot 1 dt = \int_0^2 t dt = 2.$$

20.2. Kurvenintegrale von Vektorfeldern: Sei $D \subseteq \mathbb{R}^n$ offen und $\vec{v} \in C(D, \mathbb{R}^n)$. Für eine Kurve $\gamma : [a, b] \rightarrow D$ (siehe 19.5) setzt man

$$\int_{\gamma} \vec{v} \cdot d\vec{x} := \int_{\gamma} \vec{v}(\vec{x}) \cdot d\vec{x} := \int_a^b \vec{v}(\gamma(t)) \cdot \gamma'(t) dt$$

(beachte $\gamma(t), \gamma'(t) \in \mathbb{R}^n$ und $\cdot$ ist das Skalarprodukt in $\mathbb{R}^n$). Im Fall $\gamma = \gamma_1 + \gamma_2 + \dots + \gamma_m$ definiert man analog wie in 20.1

$$\int_{\gamma} \vec{v} \cdot d\vec{x} := \int_{\gamma} \vec{v}(\vec{x}) \cdot d\vec{x} := \sum_{j=1}^m \int_{\gamma_j} \vec{v}(\vec{x}) \cdot d\vec{x}$$

Bemerkung: Ist γ regulär so heißt

$$\vec{T}(\gamma(t)) := \frac{\gamma'(t)}{\|\gamma'(t)\|}$$

Tangenteneinheitsvektor an die Kurve γ im Punkt $\gamma(t)$ und es gilt

$$\int_{\gamma} \vec{v} \cdot d\vec{x} = \int_{\gamma} (\vec{v} \cdot \vec{T}) ds.$$

Hier betrachtet man also $\vec{T}$ als Funktion auf der Spur von γ , was jedenfalls geht, wenn γ injektiv ist; ansonsten muss man γ zusammensetzen wie in Bemerkung 20.1(d). Die

Gleichung bedeutet, dass von dem Vektorfeld $\vec{v}$ entlang γ nur der Tangentialanteil integriert wird.

Beispiele: (1) Es gilt

$$\int_{\gamma^-} \vec{v} \cdot d\vec{x} = - \int_{\gamma} \vec{v} \cdot d\vec{x},$$

d.h. das Kurvenintegral eines Vektorfeldes ändert bei Orientierungsumkehr der Kurve das Vorzeichen (vgl. aber mit 20.1 Bemerkung (c)).

(2) Ist $f \in C^1(D, \mathbb{R})$ und $\gamma : [a, b] \rightarrow D$ eine Kurve, so gilt

$$\int_{\gamma} \nabla f \cdot d\vec{x} = f(\gamma(b)) - f(\gamma(a)).$$

Beweis. Die Kettenregel liefert

$$\frac{d}{dt} f(\gamma(t)) = f'(\gamma(t)) \gamma'(t) = (\nabla f)(\gamma(t)) \cdot \gamma'(t)$$

also (nach dem Hauptsatz)

$$\int_{\gamma} \nabla f \cdot d\vec{x} = \int_a^b (\nabla f)(\gamma(t)) \cdot \gamma'(t) dt = \int_a^b \frac{d}{dt} f(\gamma(t)) dt = f(\gamma(b)) - f(\gamma(a)).$$

□

(3) Betrachte $D := \mathbb{R}^2 \setminus \{\vec{0}\}$,

$$\vec{v} : D \rightarrow \mathbb{R}^2, \quad \vec{v}(x, y) = \frac{1}{x^2 + y^2} \begin{pmatrix} -y \\ x \end{pmatrix}, \quad \gamma(t) := \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \quad (t \in [0, 2\pi]).$$

Nachrechnen: $\vec{v}$ erfüllt die Integrabilitätsbedingungen aus 19.22.

Weiter gilt

$$\int_{\gamma} \vec{v} \cdot d(x, y) = \int_0^{2\pi} \frac{1}{\cos^2 t + \sin^2 t} \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \cdot \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} dt = 2\pi.$$

Mit obigem Beispiel (2) folgt: $\vec{v}$ hat auf $\mathbb{R}^2 \setminus \{\vec{0}\}$ keine Stammfunktion. Verkleinert man den Definitionsbereich z.B. zu

$$D_0 := \{(x, y) \in \mathbb{R}^2 : x > 0\} \subseteq D$$

so hat $\vec{v}$ eine Stammfunktion auf D_0 , nämlich (nachrechnen)

$$f(x, y) = \arctan\left(\frac{y}{x}\right).$$

20.3. Gebiete und einfach zusammenhängende Mengen: Eine Teilmenge $K \subseteq \mathbb{R}^n$ heißt *konvex*, falls

$$\forall \vec{x}, \vec{y} \in K : S[\vec{x}, \vec{y}] \subseteq K,$$

mit $S[\vec{x}, \vec{y}]$ die Verbindungsstrecke von $\vec{x}$ und $\vec{y}$, siehe 19.16.

Beispiele: $K(\vec{0}, 1)$ ist konvex, $K(\vec{0}, 1) \setminus \{\vec{0}\}$ ist nicht konvex.

Definition: Sind $\vec{x}_0, \vec{x}_1, \dots, \vec{x}_m \in \mathbb{R}^n$, so heißt

$$S[\vec{x}_0, \vec{x}_1, \dots, \vec{x}_m] := S[\vec{x}_0, \vec{x}_1] \cup S[\vec{x}_1, \vec{x}_2] \cup \dots \cup S[\vec{x}_{m-1}, \vec{x}_m]$$

Strecken Zug durch $\vec{x}_0, \vec{x}_1, \dots, \vec{x}_m$.

Ein Streckenzug ist Bild einer Kurve im Sinne von 20.1 Bemerkung (d).

Definition: Eine offene Teilmenge $G \subseteq \mathbb{R}^n$ heißt ein *Gebiet*, falls es zu je zwei Punkten $\vec{x}, \vec{y} \in G$ einen Streckenzug $S[\vec{x}_0, \vec{x}_1, \dots, \vec{x}_m] \subseteq G$ gibt mit $\vec{x}_0 = \vec{x}$ und $\vec{x}_m = \vec{y}$, d.h. wenn man je zwei Punkte aus G durch einen in G verlaufenden Streckenzug verbinden kann.

Ohne Beweis: Eine offene Teilmenge $G \subseteq \mathbb{R}^n$ ist genau dann ein Gebiet falls es zu je zwei Punkten $\vec{x}, \vec{y} \in G$ eine Kurve $\gamma : [a, b] \rightarrow G$ gibt mit $\gamma(a) = \vec{x}$ und $\gamma(b) = \vec{y}$.

Beispiele: Ist G offen und konvex, so ist G ein Gebiet. $\mathbb{R}^n \setminus \{\vec{0}\}$ ist ein Gebiet. Die Menge

$$\{(x, y) \in \mathbb{R}^2 : x \neq 0\}$$

ist kein Gebiet.

Satz: Ist $G \subseteq \mathbb{R}^n$ ein Gebiet, $f \in C^1(D, \mathbb{R})$ und

$$\text{grad } f(\vec{x}) = \vec{0} \quad (\vec{x} \in G),$$

so ist f konstant.

Beweis. Nach dem mehrdimensionalen Mittelwertsatzes (vgl. 19.16 Bemerkung (b)) existieren für jeden Streckenzug $S[\vec{x}_0, \vec{x}_1, \dots, \vec{x}_m] \subseteq G$ Punkte $\xi_k \in S[\vec{x}_k, \vec{x}_{k+1}]$ ($k = 0, \dots, m-1$) mit

$$f(\vec{x}_{k+1}) - f(\vec{x}_k) = \text{grad } f(\xi_k) \cdot (\vec{x}_{k+1} - \vec{x}_k) = 0,$$

also $f(\vec{x}_0) = f(\vec{x}_1) = \dots = f(\vec{x}_m)$. □

Definition (nur anschaulich): Ein Gebiet $G \subseteq \mathbb{R}^n$ heißt *einfach zusammenhängend* $:\Leftrightarrow$ “Jede geschlossene Kurve in G kann in G zu einem Punkt in G zusammengezogen werden.”

Bemerkung: Für ein Gebiet $G \subseteq \mathbb{R}^2$ bedeutet dies anschaulich: “ G hat keine Löcher.”

Beispiele: (1) $\mathbb{R}^2 \setminus \{\vec{0}\}$ ist nicht einfach zusammenhängend.

(2) Gibt es in $G \subseteq \mathbb{R}^n$ einen Punkt $\vec{c} \in G$ mit $S[\vec{c}, \vec{x}] \subseteq G$ ($\vec{x} \in G$) (solche Gebiete heißen *sternförmig*), so ist G einfach zusammenhängend. Konvexe Mengen sind sternförmig.

(3) $\mathbb{R}^3 \setminus \{\vec{0}\}$ ist einfach zusammenhängend, aber $\mathbb{R}^3 \setminus \{(0, 0, z) : z \in \mathbb{R}\}$ ist **nicht** einfach zusammenhängend.

20.4. Kurvenintegrale und Potentialfelder:

Satz: Sei $G \subseteq \mathbb{R}^n$ ein Gebiet und $\vec{v} \in C(G, \mathbb{R}^n)$. Dann sind äquivalent:

- (i) $\vec{v}$ ist Potentialfeld in G .
- (ii) Für je zwei Punkte $\vec{y}_1, \vec{y}_2 \in G$ ist $\int_\gamma \vec{v} \cdot d\vec{x}$ unabhängig von der Kurve $\gamma : [a, b] \rightarrow G$ mit $\gamma(a) = \vec{y}_1$, $\gamma(b) = \vec{y}_2$.
- (iii) Für jede geschlossene Kurve $\gamma : [a, b] \rightarrow G$ gilt

$$\int_\gamma \vec{v} \cdot d\vec{x} = 0.$$

Ist zusätzlich G einfach zusammenhängend und $\vec{v} \in C^1(G, \mathbb{R}^n)$, so ist (i) (also auch (ii) und (iii)) äquivalent zu den Integrabilitätsbedingungen

- (iv) $\partial_j v_k = \partial_k v_j \quad (j, k = 1, \dots, n).$

Bemerkung: Für $G \subseteq \mathbb{R}^3$ gilt also: Ist G ein einfach zusammenhängendes Gebiet und ist $\text{rot } \vec{v} = \vec{0}$ in G , so ist $\vec{v}$ ein Potentialfeld in G .

Beispiel: In 20.2 Beispiel (3) ist $D = \mathbb{R}^2 \setminus \{\vec{0}\}$ nicht einfach zusammenhängend. Das Beispiel zeigt, dass (iv) $\Rightarrow$ (i) für beliebige Gebiete i.a. **nicht** gilt. Die Menge D_0 in diesem Beispiel ist konvex, also einfach zusammenhängend.

21 Das mehrdimensionale Riemann-Integral

Alle Sätze in diesem Kapitel geben wir ohne Beweis an.

21.1. Die Konstruktion des Integrals: Ist $[a, b] \subseteq \mathbb{R}$ ein Intervall mit $a \leq b$ und $t_0, \dots, t_N \in \mathbb{R}$ mit

$$a = t_0 < t_1 < t_2 < \dots < t_N = b,$$

so heißt $Z := \{t_0, \dots, t_N\}$ eine **Zerlegung von** $[a, b]$ und die Intervalle

$$[t_{k-1}, t_k] \quad (k = 1, \dots, N)$$

heißen die *Teilintervalle bezüglich Z* .

Sind $a_1, b_1, a_2, b_2, \dots, a_n, b_n \in \mathbb{R}$ mit $a_j \leq b_j$ ($j = 1, \dots, n$), so heißt die Menge

$$I = [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_n, b_n]$$

ein *kompaktes Intervall im $\mathbb{R}^n$* , und die Zahl

$$|I| = \prod_{k=1}^n (b_k - a_k)$$

heißt *Inhalt* (oder *Volumen*) von I . Beachte:

$$|I| = 0 \iff \exists j \in \{1, \dots, n\} : a_j = b_j.$$

Zu jedem $j \in \{1, \dots, n\}$ sei nun eine Zerlegung Z_j von $[a_j, b_j]$ gegeben. Dann heißt

$$Z := Z_1 \times Z_2 \times \dots \times Z_n$$

eine *Zerlegung von I* . Ist nun T_j jeweils ein Teilintervall bezüglich Z_j ($j = 1, \dots, n$), so heißt

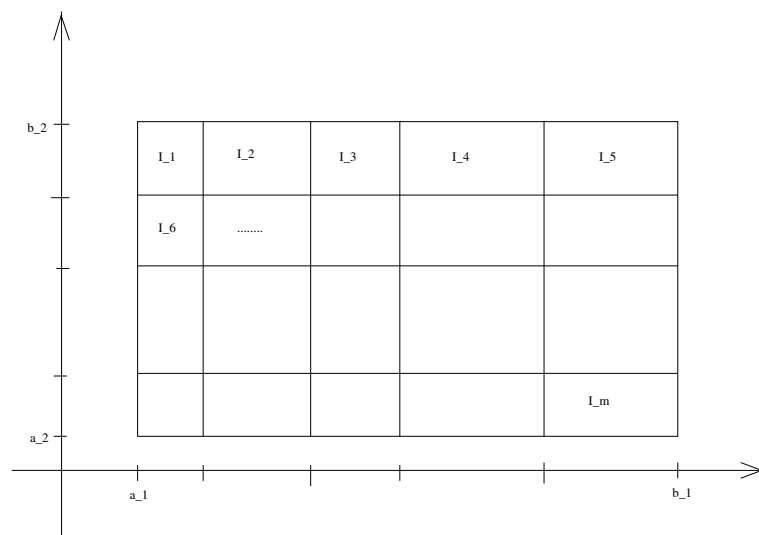
$$T_1 \times T_2 \times \dots \times T_n$$

ein *Teilintervall von I bezüglich Z* .

Es seien $I_1, \dots, I_m$ die Teilintervalle von I bzgl. Z . Dann gilt:

$$I = I_1 \cup I_2 \cup \dots \cup I_m, \quad |I| = |I_1| + \dots + |I_m|.$$

Wir bezeichnen mit $\mathcal{Z}(I)$ die Menge aller Zerlegungen von I .



Definition: Es sei $I \subseteq \mathbb{R}^n$ ein kompaktes Intervall, $f : I \rightarrow \mathbb{R}$ sei beschränkt und $Z \in \mathcal{Z}(I)$

sei eine Zerlegung mit den Teilintervallen $I_1, \dots, I_m$. Wir setzen:

$$m_j := \inf f(I_j), \quad M_j := \sup f(I_j) \quad (j = 1, \dots, m).$$

Die **Untersumme** bzw. die **Obersumme** von f bzgl. Z ist definiert durch

$$s_f(Z) := \sum_{j=1}^m m_j |I_j| \quad \text{bzw.} \quad S_f(Z) := \sum_{j=1}^m M_j |I_j|.$$

Bemerkung: Wie beim Riemann-Integral in HM I kann man zeigen: Sind $Z, \tilde{Z} \in \mathcal{Z}(I)$, so gilt:

1. Ist $Z \subseteq \tilde{Z}$, so ist $s_f(Z) \leq s_f(\tilde{Z})$, $S_f(Z) \geq S_f(\tilde{Z})$.
2. $(\inf f(I))|I| \leq s_f(Z) \leq S_f(\tilde{Z}) \leq (\sup f(I))|I|$.

Definition: Es sei $I \subseteq \mathbb{R}^n$ ein kompaktes Intervall und $f : I \rightarrow \mathbb{R}$ beschränkt. Wir setzen

$$s_f := \sup \{s_f(Z) : Z \in \mathcal{Z}(I)\},$$

$$S_f := \inf \{S_f(Z) : Z \in \mathcal{Z}(I)\}.$$

Mit obiger Bemerkung folgt $s_f \leq S_f$. Die Funktion f heißt *integrierbar über I* : $\iff s_f = S_f$. In diesem Fall heißt

$$\int_I f d\vec{x} := \int_I f(\vec{x}) d\vec{x} := s_f (= S_f)$$

das *Integral von f über I* und wir schreiben $f \in R(I) = R(I, \mathbb{R})$.

Wie im Fall $n = 1$ (vgl. HM I) gilt:

$$C(I) \subseteq R(I).$$

21.2. Der Satz von Fubini: Es seien $n_1, n_2 \in \mathbb{N}$, $n = n_1 + n_2$ (also $\mathbb{R}^n = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$). Es sei I_k ein kompaktes Intervall im $\mathbb{R}^{n_k}$ ($k = 1, 2$), und es sei $I := I_1 \times I_2$ und $f \in R(I)$. Punkte in I bezeichnen wir mit $(\vec{x}, \vec{y})$, wobei $\vec{x} \in I_1$ und $\vec{y} \in I_2$.

(1) Für jedes feste $\vec{y} \in I_2$ sei die Funktion $\vec{x} \mapsto f(\vec{x}, \vec{y})$ integrierbar über I_1 und es sei $g(\vec{y}) := \int_{I_1} f(\vec{x}, \vec{y}) d\vec{x}$. Dann gilt $g \in R(I_2)$ und

$$\int_I f(\vec{x}, \vec{y}) d(\vec{x}, \vec{y}) = \int_{I_2} g(\vec{y}) d\vec{y} = \int_{I_2} \left(\int_{I_1} f(\vec{x}, \vec{y}) d\vec{x} \right) d\vec{y}.$$

(2) Für jedes feste $\vec{x} \in I_1$ sei die Funktion $\vec{y} \mapsto f(\vec{x}, \vec{y})$ integrierbar über I_2 und es sei $g(\vec{x}) := \int_{I_2} f(\vec{x}, \vec{y}) d\vec{y}$. Dann gilt $g \in R(I_1)$ und

$$\int_I f(\vec{x}, \vec{y}) d(\vec{x}, \vec{y}) = \int_{I_1} g(\vec{x}) d\vec{x} = \int_{I_1} \left(\int_{I_2} f(\vec{x}, \vec{y}) d\vec{y} \right) d\vec{x}.$$

Bemerkung: Ist in obigem Satz $f \in C(I)$, so ist in (1) $\vec{x} \mapsto f(\vec{x}, \vec{y})$ stetig in I_1 für jedes $\vec{y} \in I_2$, also stets in $R(I_1)$, und es gilt $g \in C(I_2)$. Die analoge Aussage gilt in (2).

Folgerung: Es sei $I = [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_n, b_n]$ und $f \in C(I)$. Dann ist

$$\begin{aligned} \int_I f(\vec{x}) d\vec{x} &= \int_I f(x_1, \dots, x_n) d(x_1, \dots, x_n) \\ &= \int_{a_1}^{b_1} \left(\dots \int_{a_{n-1}}^{b_{n-1}} \left(\int_{a_n}^{b_n} f(x_1, \dots, x_n) dx_n \right) dx_{n-1} \dots \right) dx_1 \end{aligned}$$

wobei die Reihenfolge der Integrationen beliebig vertauscht werden darf.

Beispiele: (1) Es sei $I = [0, \frac{\pi}{2}] \times [0, \frac{\pi}{2}]$ und $f(x, y) = \sin(x + y)$.

$$\begin{aligned} \int_I \sin(x + y) d(x, y) &= \int_0^{\frac{\pi}{2}} \left(\int_0^{\frac{\pi}{2}} \sin(x + y) dy \right) dx \\ &= \int_0^{\frac{\pi}{2}} [-\cos(x + y)]_{y=0}^{y=\frac{\pi}{2}} dx = \int_0^{\frac{\pi}{2}} \left(-\cos\left(x + \frac{\pi}{2}\right) + \cos(x) \right) dx \\ &= \left[-\sin\left(x + \frac{\pi}{2}\right) + \sin x \right]_0^{\frac{\pi}{2}} \\ &= -\sin(\pi) + \sin\left(\frac{\pi}{2}\right) - \left(-\sin\left(\frac{\pi}{2}\right) + \sin 0 \right) = 1 + 1 = 2. \end{aligned}$$

(2) Es sei $I = [0, 2] \times [0, 1] \times [0, 1]$ und $f(x, y, z) = x^2 z + yxz$.

$$\begin{aligned} &\int_I (x^2 z + yxz) d(x, y, z) \\ &= \int_0^1 \left(\int_0^2 \left(\int_0^1 (x^2 z + yxz) dz \right) dx \right) dy \\ &= \int_0^1 \left(\int_0^2 \left[\frac{1}{2} x^2 z^2 + \frac{1}{2} yxz^2 \right]_{z=0}^{z=1} dx \right) dy \\ &= \int_0^1 \left(\int_0^2 \left(\frac{1}{2} x^2 + \frac{1}{2} yx \right) dx \right) dy \\ &= \int_0^2 \left(\int_0^1 \left(\frac{1}{2} x^2 + \frac{1}{2} yx \right) dy \right) dx \\ &= \int_0^2 \left[\frac{1}{2} x^2 y + \frac{1}{4} y^2 x \right]_{y=0}^{y=1} dx = \int_0^2 \left(\frac{1}{2} x^2 + \frac{1}{4} x \right) dx \end{aligned}$$

$$= \frac{1}{6}x^3 + \frac{1}{8}x^2 \Big|_0^2 = \frac{8}{6} + \frac{4}{8} = \frac{8}{6} + \frac{3}{6} = \frac{11}{6}.$$

(3) Es sei $I = [a_1, b_1] \times [a_2, b_2] \subseteq \mathbb{R}^2$, $f \in C([a_1, b_1])$ und $g \in C([a_2, b_2])$.

$$\begin{aligned} \int_I f(x)g(y)d(x, y) &= \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} f(x)g(y)dy \right) dx \\ &= \int_{a_1}^{b_1} f(x) \left(\int_{a_2}^{b_2} g(y)dy \right) dx \\ &= \left(\int_{a_1}^{b_1} f(x)dx \right) \left(\int_{a_2}^{b_2} g(y)dy \right). \end{aligned}$$

21.3. Das Volumen: Wir wollen nun bestimmten beschränkten Teilmengen von $\mathbb{R}^n$ ein Volumen zuordnen. Es sei also $B \subseteq \mathbb{R}^n$ beschränkt. Die Funktion $\chi_B : \mathbb{R}^n \rightarrow \mathbb{R}$,

$$\chi_B(\vec{x}) := \begin{cases} 1, & \vec{x} \in B \\ 0, & \vec{x} \notin B \end{cases}$$

heißt *die charakteristische Funktion von B*. Wähle ein kompaktes Intervall $I \subseteq \mathbb{R}^n$ mit $B \subseteq I$.

Es sei $Z \in \mathcal{Z}(I)$ mit den Teilintervallen $I_1, \dots, I_m$. Dann gilt:

$$\inf \chi_B(I_j) = \begin{cases} 1, & \text{falls } I_j \subseteq B \\ 0, & \text{falls } I_j \not\subseteq B \end{cases}.$$

Damit folgt:

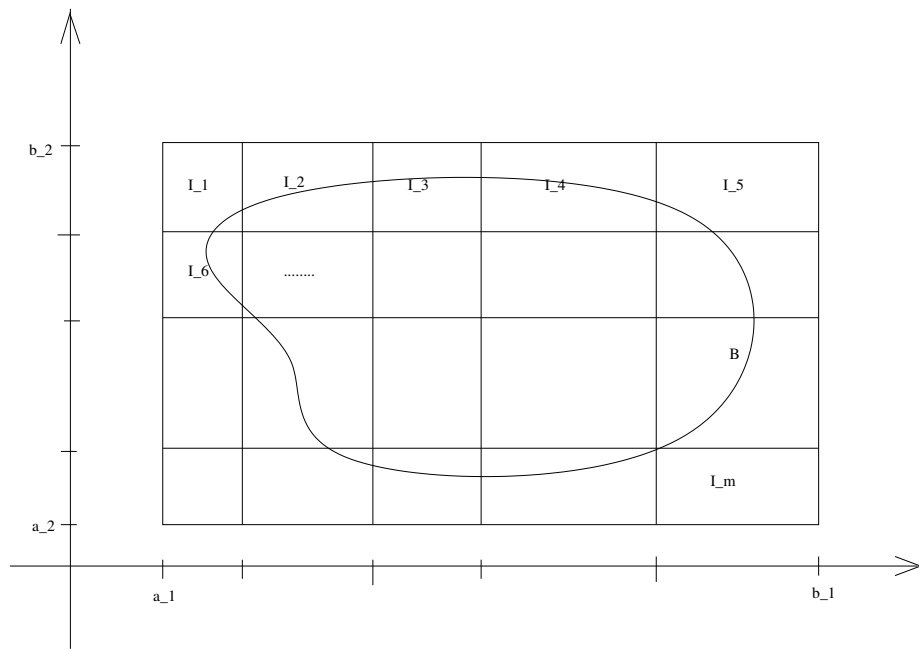
$$s_{\chi_B}(Z) = \sum_{j: I_j \subseteq B} |I_j|.$$

Weiter gilt:

$$\sup \chi_B(I_j) = \begin{cases} 1, & \text{falls } I_j \cap B \neq \emptyset \\ 0, & \text{falls } I_j \cap B = \emptyset \end{cases}.$$

Damit folgt:

$$S_{\chi_B}(Z) = \sum_{j: I_j \cap B \neq \emptyset} |I_j|.$$



Definition: Wir definieren den *inneren Inhalt* bzw. den *äußeren Inhalt* von B als

$$\underline{V}(B) := s_{\chi_B} \text{ bzw. } \overline{V}(B) := S_{\chi_B}.$$

Die Menge B heißt (*Jordan-*)*messbar* falls $\chi_B \in R(I)$. In diesem Fall ist

$$\underline{V}(B) = \overline{V}(B) = \int_I \chi_B(\vec{x}) d\vec{x}$$

und die Zahl

$$V(B) := \int_I \chi_B(\vec{x}) d\vec{x}$$

heißt das *Volumen* oder der *Inhalt* von B .

Bemerkungen: (1) Diese Definitionen sind unabhängig von der Wahl des Intervalls I , solange nur $B \subseteq I$ gilt.

(2) Ist $\overline{V}(B) = 0$, so gilt $0 \leq \underline{V}(B) \leq \overline{V}(B) = 0$. Dann ist B also messbar mit $V(B) = 0$. Wir nennen dann B auch eine (*Jordansche*) *Nullmenge*. Ist $B \subseteq \mathbb{R}^n$ beschränkt, so gilt

$$\overline{V}(B) = \underline{V}(B) + \overline{V}(\partial B),$$

d.h. B ist genau dann messbar, wenn ihr *Rand* ∂B eine Nullmenge ist.

Beispiele: (1) Betrachte $B = \emptyset$. Es sei I ein beliebiges kompaktes Intervall in $\mathbb{R}^n$. Dann gilt $\chi_B(\vec{x}) = 0$ ($\vec{x} \in I$). Also ist

$$s_{\chi_B}(Z) = S_{\chi_B}(Z) = 0 \quad (Z \in \mathcal{Z}(I)).$$

Somit ist $\emptyset$ messbar und $V(\emptyset) = 0$.

(2) Es sei $B \subseteq \mathbb{R}^n$ ein kompaktes Intervall. Wir haben nun **zwei** Definitionen für das Volumen von B , nämlich $V(B)$ und $|B|$. Diese Definitionen stimmen überein:

Wähle $I = B$. Für $Z \in \mathcal{Z}(I)$ mit den Teilintervallen $I_1, \dots, I_m$ gilt

$$s_{\chi_B}(Z) = S_{\chi_B}(Z) = \sum_{j=1}^m |I_j| = |I| = |B|.$$

Also ist B messbar und $V(B) = |B|$.

(3) Betrachte $B := [0, 1] \cap \mathbb{Q}$ und $I = [0, 1]$. Es gilt:

$$\chi_B(x) = \begin{cases} 1, & x \in [0, 1] \cap \mathbb{Q} \\ 0, & \text{sonst.} \end{cases}.$$

Aus HM I 12.2 ist bekannt: $\chi_B \notin R(I)$. Also ist B nicht messbar. In diesem Beispiel ist $\partial B = [0, 1]$, also ∂B keine Nullmenge.

21.4. Das Integral über messbare Mengen: Es sei $B \subseteq \mathbb{R}^n$ messbar und $f : B \rightarrow \mathbb{R}$ beschränkt. Es sei

$$f_B(\vec{x}) := \begin{cases} f(\vec{x}), & \vec{x} \in B \\ 0, & \vec{x} \notin B \end{cases},$$

und $I \subseteq \mathbb{R}^n$ ein kompaktes Intervall mit $B \subseteq I$. Gilt dann $f_B \in R(I)$, so heißt f *über B (Riemann-)integrierbar*. In diesem Fall schreiben wir $f \in R(B) = R(B, \mathbb{R})$ und die Zahl

$$\int_B f d\vec{x} := \int_B f(\vec{x}) d\vec{x} := \int_I f_B(\vec{x}) d\vec{x}$$

heißt das *(Riemann-)Integral von f über B* .

Bemerkungen: (1) Diese Definitionen sind unabhängig von der Wahl von I , solange nur $B \subseteq I$ gilt.

(2) Ist $B \subseteq \mathbb{R}^n$ messbar und speziell $f : B \rightarrow \mathbb{R}$, $f(\vec{x}) = 1$ ($\vec{x} \in B$), so ist $f_B = \chi_B$ und somit

$$V(B) = \int_B 1 d\vec{x}.$$

Satz (Rechenregeln): Es seien $A, B \subseteq \mathbb{R}^n$ messbar und $\alpha, \beta \in \mathbb{R}$. Dann gilt:

1. Ist $f \in C(B)$ und beschränkt, so ist $f \in R(B)$. (Beachte: Ist B nicht kompakt, so können Funktionen aus $C(B)$ unbeschränkt sein.)

2. Es seien $f, g \in R(B)$. Dann gilt:

$$\alpha f + \beta g, fg, |f| \in R(B),$$

$$\int_B (\alpha f + \beta g) d\vec{x} = \alpha \int_B f d\vec{x} + \beta \int_B g d\vec{x}, \quad \left| \int_B f d\vec{x} \right| \leq \int_B |f| d\vec{x}.$$

3. Sind $f, g \in R(B)$ und ist $f \leq g$ auf B , so ist

$$\int_B f d\vec{x} \leq \int_B g d\vec{x}.$$

4. Sind $f, g \in R(B)$ und existiert ein $c > 0$ mit $|g(\vec{x})| \geq c$ ($\vec{x} \in B$), so ist $\frac{f}{g} \in R(B)$.

5. $A \cup B$, $A \cap B$ und $A \setminus B$ sind messbar.

6. Aus $A \subseteq B$ folgt $V(A) \leq V(B)$.

7. $f \in R(A \cup B) \iff f \in R(A) \cap R(B)$. In diesem Fall ist

$$\int_{A \cup B} f d\vec{x} = \int_A f d\vec{x} + \int_B f d\vec{x} - \int_{A \cap B} f d\vec{x}.$$

Insbesondere gilt:

$$V(A \cup B) = V(A) + V(B) - V(A \cap B).$$

8. Ist $A^\circ \subseteq C \subseteq \overline{A}$, so ist C messbar und $V(C) = V(A)$.

9. Es seien $f, g \in R(B)$ und $g \leq f$ auf B . Weiter sei

$$M_{f,g} := \{(\vec{x}, y) \in \mathbb{R}^{n+1} : \vec{x} \in B, g(\vec{x}) \leq y \leq f(\vec{x})\}.$$

Dann ist $M_{f,g}$ messbar (im $\mathbb{R}^{n+1}$) und

$$V(M_{f,g}) = \int_B (f - g) d\vec{x}.$$

Ist speziell $g = 0$ auf B , so ist

$$V(M_{f,0}) = \int_B f d\vec{x}.$$

Beispiele: (1) Betrachte für $r > 0$ die Kreisscheibe

$$K := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq r^2\}.$$

Das Intervall $B := [-r, r] \subseteq \mathbb{R}$ ist messbar. Für $x \in B$ sei

$$f(x) := \sqrt{r^2 - x^2}, \quad g(x) := -\sqrt{r^2 - x^2}.$$

Dann gilt: $f, g \in C(B) \subseteq R(B)$ und $K = M_{f,g}$:

$$\begin{aligned} g(x) \leq y \leq f(x) &\iff -\sqrt{r^2 - x^2} \leq y \leq \sqrt{r^2 - x^2} \\ &\iff |y| \leq \sqrt{r^2 - x^2} \iff y^2 \leq r^2 - x^2. \end{aligned}$$

Also ist K messbar und

$$V(K) = \int_B (f - g) dx = \int_{-r}^r 2\sqrt{r^2 - x^2} dx \stackrel{HMI}{=} \pi r^2.$$

(2) Betrachte die Menge

$$K := \{(x, y, z) \in \mathbb{R}^3 : 0 \leq x \leq 1, 0 \leq y \leq 1, z \leq 1 - y^2, z \geq 0\},$$

und $B := [0, 1]^2$. B ist messbar. Für $(x, y) \in B$ sei

$$f(x, y) := 1 - y^2.$$

Dann gilt: $K = M_{f,0}$ und $f \in C(B) \subseteq R(B)$. Also ist K messbar und

$$\begin{aligned} V(K) &= \int_B f(x, y) d(x, y) = \int_0^1 \left(\int_0^1 (1 - y^2) dy \right) dx \\ &= \int_0^1 \left[y - \frac{1}{3} y^3 \right]_{y=0}^{y=1} dx = \int_0^1 \left(1 - \frac{1}{3} \right) dx = \frac{2}{3}. \end{aligned}$$

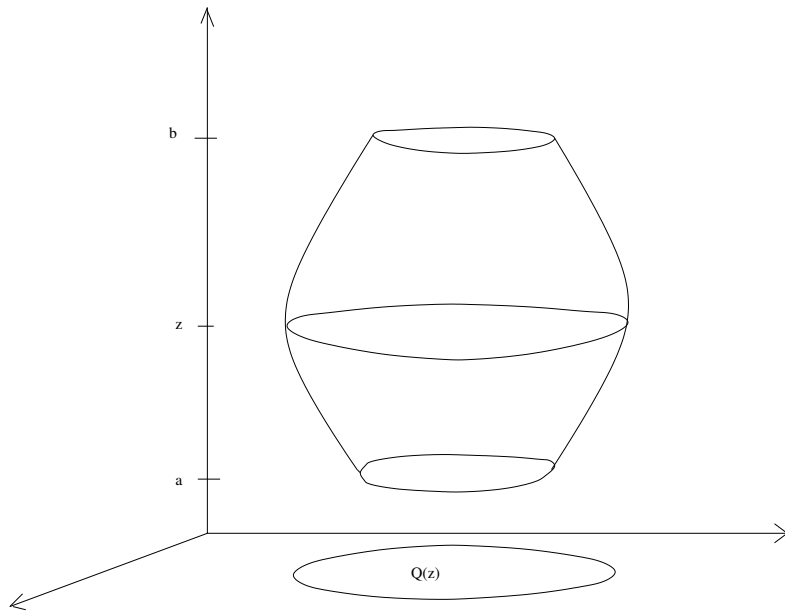
21.5. Das Prinzip von Cavalieri: Die Menge $B \subseteq \mathbb{R}^{n+1}$ sei messbar. Für Punkte im $\mathbb{R}^{n+1}$ schreiben wir $(\vec{x}, z)$ mit $\vec{x} \in \mathbb{R}^n$ und $z \in \mathbb{R}$. Es seien $a, b \in \mathbb{R}$ so, dass $a \leq z \leq b$ ($(\vec{x}, z) \in B$). Für $z \in [a, b]$ sei

$$Q(z) := \{\vec{x} \in \mathbb{R}^n : (\vec{x}, z) \in B\}.$$

Weiter sei $Q(z)$ messbar für jedes $z \in [a, b]$. Dann ist $z \mapsto V(Q(z))$ integrierbar über $[a, b]$ und

$$V(B) = \int_a^b V(Q(z)) dz.$$

Bemerkung: Die Koordinate bzgl. der im Prinzip von Cavalieri die *Querschnitte* $Q(\cdot)$ gebildet werden kann beliebig gewählt werden.



Beispiele: (1) Wir betrachten die abgeschlossene Kugel um $(0, 0, 0)$ mit Radius $r > 0$, also

$$B := \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq r^2\}.$$

Wähle $a = -r$, $b = r$. Für $z \in [-r, r]$ ist dann

$$Q(z) := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq r^2 - z^2\}$$

die Kreisscheibe um $(0, 0)$ mit Radius $\sqrt{r^2 - z^2}$. Es gilt

$$V(Q(z)) = \pi (r^2 - z^2).$$

Also ist

$$V(B) = \int_{-r}^r \pi (r^2 - z^2) dz = \frac{4}{3} \pi r^3.$$

(2) **Höherdimensionale Kugeln:** Wir betrachten die abgeschlossene Kugel um $(0, 0, 0, 0)$ mit Radius $r > 0$, also

$$B := \{(w, x, y, z) \in \mathbb{R}^4 : w^2 + x^2 + y^2 + z^2 \leq r^2\}.$$

Wähle $a = -r$, $b = r$. Für $z \in [-r, r]$ ist dann

$$Q(z) := \{(w, x, y) \in \mathbb{R}^3 : w^2 + x^2 + y^2 \leq r^2 - z^2\}$$

die Kugel um $(0, 0, 0)$ mit Radius $\sqrt{r^2 - z^2}$. Es gilt

$$V(Q(z)) = \frac{4}{3}\pi (r^2 - z^2)^{\frac{3}{2}}.$$

Also ist

$$V(B) = \int_{-r}^r \frac{4}{3}\pi (r^2 - z^2)^{\frac{3}{2}} dz = \frac{4\pi}{3} \frac{3\pi r^4}{8} = \frac{1}{2}\pi^2 r^4.$$

Bezeichnet $B_n(r)$ die abgeschlossene Kugel um $\vec{0}$ im $\mathbb{R}^n$ mit Radius r , so kann man mit obiger Methode induktiv zeigen:

$$V(B_{2n}(r)) = \frac{\pi^n}{n!} r^{2n}, \quad V(B_{2n+1}(r)) = \frac{2^{n+1}\pi^n}{1 \cdot 3 \cdot \dots \cdot (2n+1)} r^{2n+1}.$$

Bei festem r gilt insbesondere (Übung):

$$\lim_{n \rightarrow \infty} V(B_n(r)) = 0.$$

(3) **Rotationskörper:** Es sei $a < b$, $f \in C([a, b], \mathbb{R})$ und $f \geq 0$ auf $[a, b]$. Der Graph von f rotiere z.B. um die x -Achse, d.h. wir betrachten die Menge

$$B := \{(x, y, z) \in \mathbb{R}^3 : x \in [a, b], y^2 + z^2 \leq f(x)^2\}.$$

Mit dem Satz aus 21.4 kann man zeigen, dass B messbar ist. Für $x \in [a, b]$ bilden wir die Schnitte

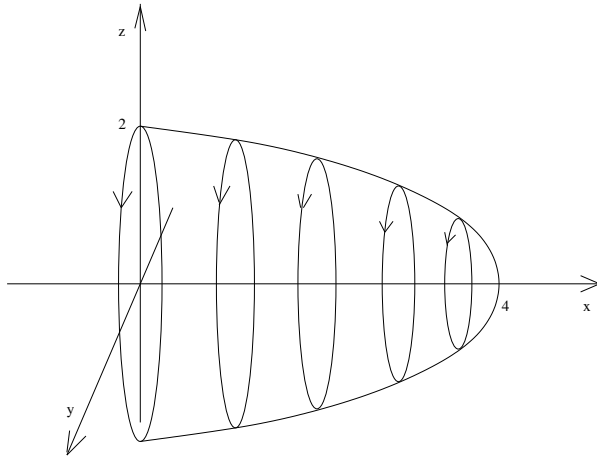
$$Q(x) := \{(y, z) \in \mathbb{R}^2 : y^2 + z^2 \leq f(x)^2\}.$$

Es gilt: $V(Q(x)) = \pi f(x)^2$ ($x \in [a, b]$) und somit

$$V(B) = \pi \int_a^b f(x)^2 dx.$$

Beispiel: Es sei $[a, b] = [0, 4]$ und $f(x) = \sqrt{4 - x}$. Dann ist B ein *Rotationsparaboloid* und es gilt:

$$V(B) = \pi \int_0^4 (4 - x) dx = 8\pi.$$



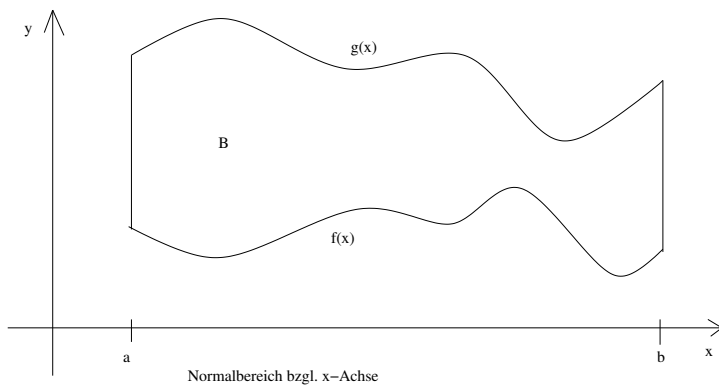
21.6. Normalbereiche: Es seien $a, b \in \mathbb{R}$, $a < b$, $f, g \in C([a, b])$ und $f \leq g$ auf $[a, b]$. Dann heißt die Menge

$$B := \{(x, y) \in \mathbb{R}^2 : x \in [a, b], f(x) \leq y \leq g(x)\}$$

bzw.

$$B := \{(x, y) \in \mathbb{R}^2 : y \in [a, b], f(y) \leq x \leq g(y)\}$$

ein *Normalbereich bzgl. der x-Achse* bzw. ein *Normalbereich bzgl. der y-Achse*. Nach dem Satz aus 21.4 ist in beiden Fällen B messbar.



Nun sei B ein Normalbereich bzgl. der x -Achse und $h \in C(B)$. Es sei

$$m := \min f([a, b]), \quad M := \max g([a, b]), \quad I := [a, b] \times [m, M].$$

Dann gilt:

$$\begin{aligned}\int_B h(x, y) d(x, y) &= \int_I h_B(x, y) d(x, y) \\ &\stackrel{\text{Fubini}}{=} \int_a^b \left(\int_m^M h_B(x, y) dy \right) dx = \int_a^b \left(\int_{f(x)}^{g(x)} h(x, y) dy \right) dx.\end{aligned}$$

Analog erhält man: Ist B ein Normalbereich bzgl. der y -Achse und $h \in C(B)$, so gilt:

$$\int_B h(x, y) d(x, y) = \int_a^b \left(\int_{f(y)}^{g(y)} h(x, y) dx \right) dy.$$

Beispiele: (1) Die Menge

$$B = \{(x, y) \in \mathbb{R}^2 : x \in [0, 1], \sqrt{x} \leq y \leq 2 - x\}$$

ist ein Normalbereich bzgl. der x -Achse. Somit gilt:

$$\begin{aligned}\int_B (x + y) d(x, y) &= \int_0^1 \left(\int_{\sqrt{x}}^{2-x} (x + y) dy \right) dx \\ &= \int_0^1 \left[xy + \frac{1}{2} y^2 \right]_{\sqrt{x}}^{2-x} dx \\ &= \int_0^1 \left(x(2 - x) + \frac{1}{2} (2 - x)^2 - x\sqrt{x} - \frac{1}{2} x \right) dx = \dots = \frac{71}{60}.\end{aligned}$$

(2) Die Menge

$$B = \{(x, y) \in \mathbb{R}^2 : y \in [0, 1], 0 \leq x \leq y^2\}$$

ist ein Normalbereich bzgl. der y -Achse ($f(y) = 0$, $g(y) = y^2$). Also gilt:

$$\begin{aligned}\int_B xy d(x, y) &= \int_0^1 \left(\int_0^{y^2} xy dx \right) dy \\ &= \int_0^1 \left[\frac{1}{2} x^2 y \right]_{x=0}^{x=y^2} dy = \int_0^1 \frac{1}{2} y^5 dy = \frac{1}{12}.\end{aligned}$$

Die Menge B ist aber auch Normalbereich bzgl. der x -Achse ($f(x) = \sqrt{x}$, $g(x) = 1$). Also gilt:

$$\begin{aligned}\int_B xy d(x, y) &= \int_0^1 \left(\int_{\sqrt{x}}^1 xy dy \right) dx = \int_0^1 \left[\frac{1}{2} xy^2 \right]_{y=\sqrt{x}}^{y=1} dx \\ &= \int_0^1 \left(\frac{1}{2} x - \frac{1}{2} x^2 \right) dx = \frac{1}{12}.\end{aligned}$$

Für die Formulierung von Integralsätzen im nächsten Kapitel definieren wir noch folgendes:

Satz und Definition: Sei $K \subseteq \mathbb{R}^2$ kompakt, und es seien $B_1, \dots, B_m \subseteq \mathbb{R}^2$ Normalbereiche bzgl. der x -Achse oder der y -Achse. Es gelte

$$K = \bigcup_{k=1}^m B_k \quad \text{und} \quad B_k^\circ \cap B_j^\circ = \emptyset \quad (k \neq j).$$

Dann heißt K *normalisierbar*, ist meßbar, und für jedes $f \in C(K, \mathbb{R})$ gilt

$$\int_K f(x, y) d(x, y) = \sum_{j=1}^m \int_{B_j} f(x, y) d(x, y)$$

Beispiel: Der Kreisring

$$K = \{(x, y) \in \mathbb{R}^2 : 1 \leq x^2 + y^2 \leq 4\}$$

ist kein Normalbereich, aber normalisierbar indem man z.B.

$$B_1 = \{(x, y) \in \mathbb{R}^2 : 1 \leq x^2 + y^2 \leq 4, y \geq 0\}, \quad B_2 = \{(x, y) \in \mathbb{R}^2 : 1 \leq x^2 + y^2 \leq 4, y \leq 0\}$$

wählt.

Normalbereiche können auch in höheren Dimensionen definiert werden. Wir behandeln dies nur anhand eines Beispiels:

Es sei $A \subseteq \mathbb{R}^2$ kompakt und messbar, $f, g : A \rightarrow \mathbb{R}$ seien stetig und es sei $f \leq g$ auf A . Wir setzen

$$B := \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in A, f(x, y) \leq z \leq g(x, y)\}.$$

Dann ist B messbar. Es sei $h \in C(B, \mathbb{R})$. Dann gilt:

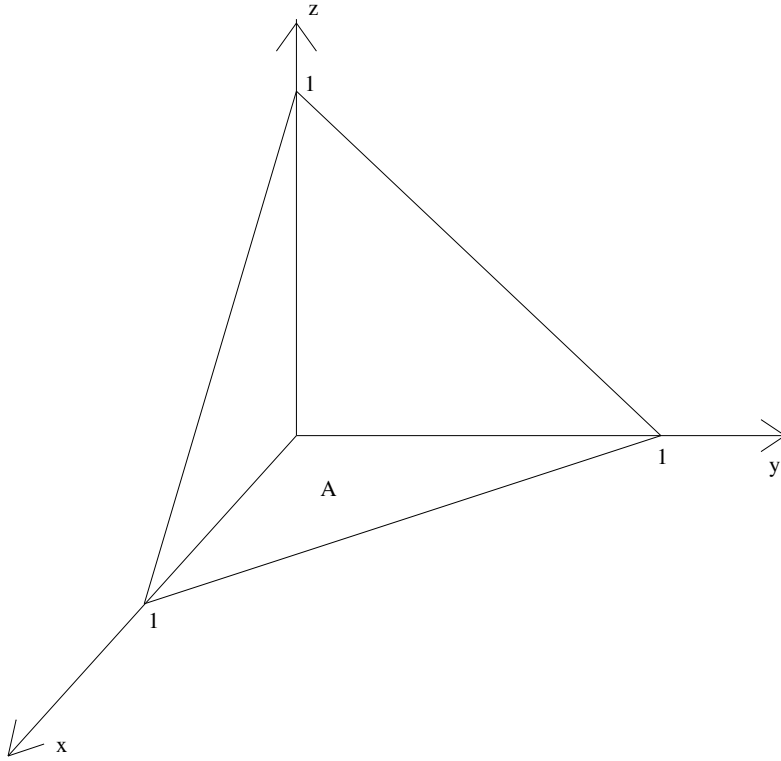
$$\int_B h(x, y, z) d(x, y, z) \stackrel{Fubini}{=} \int_A \left(\int_{f(x, y)}^{g(x, y)} h(x, y, z) dz \right) d(x, y).$$

Beispiel: Es sei $f(x, y) = 0$, $g(x, y) = 1 - (x + y)$, und

$$A = \{(x, y) \in \mathbb{R}^2 : x \geq 0, y \geq 0, x + y \leq 1\},$$

also

$$\begin{aligned} B &= \{(x, y, z) \in \mathbb{R}^3 : x, y, z \geq 0, x + y + z \leq 1\} \\ &= \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in A, 0 \leq z \leq g(x, y)\} \end{aligned}$$



Betrachte $h(x, y, z) = 2xyz$. Dann gilt:

$$\begin{aligned}
 \int_B 2xyz d(x, y, z) &= \int_A \left(\int_0^{1-(x+y)} 2xyz dz \right) d(x, y) \\
 &= \int_A [xyz^2]_{z=0}^{z=1-(x+y)} d(x, y) \\
 &= \int_A xy(1-(x+y))^2 d(x, y) \\
 &= \int_0^1 \left(\int_0^{1-x} xy(1-(x+y))^2 dy \right) dx = \dots = \frac{1}{360}.
 \end{aligned}$$

21.7. Die Substitutionsregel: Es sei $G \subseteq \mathbb{R}^n$ offen, $g \in C^1(G, \mathbb{R}^n)$ und $B \subseteq G$ kompakt und messbar. Weiter sei g auf dem Inneren B° von B injektiv und

$$\det g'(\vec{y}) \neq 0 \quad (\vec{y} \in B^\circ).$$

Ist dann $A := g(B)$ und $f \in C(A, \mathbb{R})$, so ist A kompakt und messbar und es gilt:

$$\int_A f(\vec{x}) d\vec{x} = \int_B f(g(\vec{y})) |\det g'(\vec{y})| d\vec{y}.$$

Bewegungsinvarianz des Volumens: Es sei $T \in \mathbb{R}^{n \times n}$ invertierbar, $\vec{w} \in \mathbb{R}^n$ und $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definiert durch

$$g(\vec{x}) = T\vec{x} + \vec{w}.$$

Dann ist g injektiv und $\det g'(\vec{x}) = \det T \neq 0$ ($\vec{x} \in \mathbb{R}^n$). Die Anwendung der Substitutionsregel in diesem Fall liefert die Formel

$$\int_A f(\vec{x}) d\vec{x} = |\det T| \int_B f(T\vec{y} + \vec{w}) d\vec{y},$$

(mit $B \subseteq \mathbb{R}^n$ kompakt und messbar, $A := g(B)$ und $f \in C(A, \mathbb{R})$). Insbesondere im Fall $f = 1$ auf A folgt

$$V(T(B) + \vec{w}) = V(A) = |\det T| V(B).$$

Ist T eine orthogonale Matrix ($T^T = T^{-1}$), so nennt man die affine Abbildung g auch eine *Bewegung*. Dann ist $|\det T| = 1$ und somit $V(T(B) + \vec{w}) = V(B)$. Typische Bewegungen sind Verschiebungen ($T = I_n$), Rotationen und Spiegelungen. Diese Bewegungsinvarianz des Volumens gilt allgemein für messbare Mengen:

Satz: Es sei $T \in \mathbb{R}^{n \times n}$ invertierbar, $\vec{w} \in \mathbb{R}^n$ und $B \subseteq \mathbb{R}^n$ messbar. Dann ist $T(B) + \vec{w}$ messbar, und es gilt

$$V(T(B) + \vec{w}) = |\det T| V(B).$$

Bemerkung: Im Unterschied zum Fall $n = 1$ wird im Fall $n > 1$ die Substitutionsregel i.a. nicht zur Vereinfachung der zu integrierenden Funktion sondern zur Vereinfachung des Integrationsbereiches benutzt, wie die folgenden Anwendungen zeigen.

Polarkoordinaten ($n = 2$): $x = r \cos \varphi$, $y = r \sin \varphi$, $r = \sqrt{x^2 + y^2}$. Wir setzen auf $G = \mathbb{R}^2$

$$g(r, \varphi) := \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix} \quad \text{mit} \quad \det g'(r, \varphi) = r.$$

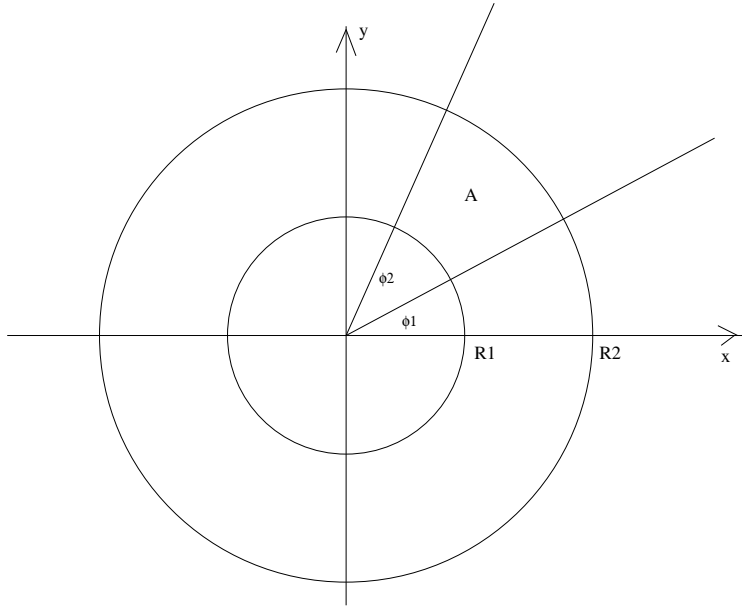
Betrachte $0 \leq \varphi_1 < \varphi_2 \leq 2\pi$, $0 \leq R_1 < R_2$ und

$$A := \left\{ \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix} : \varphi \in [\varphi_1, \varphi_2], r \in [R_1, R_2] \right\}.$$

Mit $B := [R_1, R_2] \times [\varphi_1, \varphi_2]$ ist $A = g(B)$. Auf $B^\circ = (R_1, R_2) \times (\varphi_1, \varphi_2)$ ist g injektiv und $\det g' \neq 0$. Ist nun $f \in C(A, \mathbb{R})$, so gilt:

$$\int_A f(x, y) d(x, y) = \int_B f(r \cos \varphi, r \sin \varphi) \cdot r dr d\varphi$$

$$\stackrel{\text{Fubini}}{=} \int_{\varphi_1}^{\varphi_2} \left(\int_{R_1}^{R_2} f(r \cos \varphi, r \sin \varphi) r dr \right) d\varphi.$$



Beispiele: (1) Es sei

$$A = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : 1 \leq x^2 + y^2 \leq 4 \right\}, \quad f(x, y) = x\sqrt{x^2 + y^2}.$$

Hier: $R_1 = 1$, $R_2 = 2$, $\varphi_1 = 0$, $\varphi_2 = 2\pi$, also $B = [1, 2] \times [0, 2\pi]$. Es gilt

$$\begin{aligned} \int_A x\sqrt{x^2 + y^2} d(x, y) &= \int_B (r \cos \varphi) r dr d(r, \varphi) \\ &= \int_0^{2\pi} \left(\int_1^2 r^3 \cos \varphi dr \right) d\varphi = \int_0^{2\pi} \left[\frac{1}{4} r^4 \cos \varphi \right]_{r=1}^{r=2} d\varphi \\ &= \int_0^{2\pi} \left(4 \cos \varphi - \frac{1}{4} \cos \varphi \right) d\varphi = \frac{15}{4} \int_0^{2\pi} \cos \varphi d\varphi = 0. \end{aligned}$$

(2) Es sei $R > 0$ und

$$A_R := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : x \geq 0, y \geq 0, x^2 + y^2 \leq R^2 \right\}.$$

Hier: $R_1 = 0$, $R_2 = R$, $\varphi_1 = 0$, $\varphi_2 = \frac{\pi}{2}$, also $B = [0, R] \times [0, \frac{\pi}{2}]$. Es gilt:

$$\int_{A_R} e^{-(x^2+y^2)} d(x, y) = \int_B e^{-r^2} r d(r, \varphi)$$

$$\begin{aligned}
&= \int_0^{\frac{\pi}{2}} \left(\int_0^R e^{-r^2} r dr \right) d\varphi = \frac{\pi}{2} \left[-\frac{1}{2} e^{-r^2} \right]_0^R \\
&= \frac{\pi}{2} \left(-\frac{1}{2} e^{-R^2} + \frac{1}{2} \right) = \frac{\pi}{4} (1 - e^{-R^2}) =: \alpha(R).
\end{aligned}$$

Weiter sei

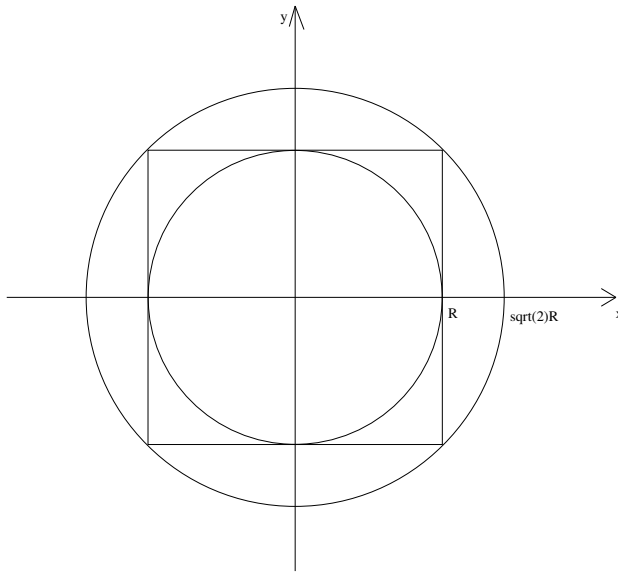
$$Q_R := [0, R] \times [0, R], \quad \beta(R) := \int_{Q_R} e^{-(x^2+y^2)} d(x, y).$$

Es ist $A_R \subseteq Q_R$ und $e^{-(x^2+y^2)} \geq 0$, also $\alpha(R) \leq \beta(R)$. Weiter ist

$$\beta(R) = \int_0^R \left(\int_0^R e^{-x^2} e^{-y^2} dy \right) dx = \left(\int_0^R e^{-x^2} dx \right)^2.$$

Setze $\rho := \sqrt{2}R$. Dann gilt $Q_R \subseteq A_\rho$ und somit

$$\beta(R) \leq \alpha(\rho) = \alpha(\sqrt{2}R).$$



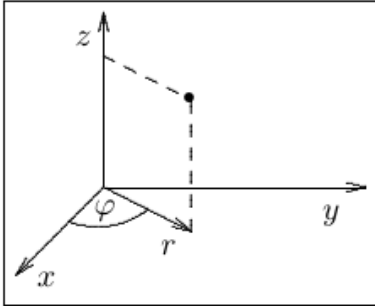
Fazit:

$$\forall R > 0 : \alpha(R) \leq \beta(R) \leq \alpha(\sqrt{2}R).$$

Damit folgt: $\frac{\pi}{4} = \lim_{R \rightarrow \infty} \beta(R)$. Also gilt: Das uneigentliche Integral

$$\int_0^\infty e^{-x^2} dx \text{ ist konvergent und } \int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{2}.$$

Zylinderkoordinaten ($n = 3$): $x = r \cos \varphi$, $y = r \sin \varphi$, $z = z$.



Wir setzen auf $G = \mathbb{R}^3$

$$g(r, \varphi, z) := \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \\ z \end{pmatrix} \quad \text{mit} \quad \det g'(r, \varphi, z) = r.$$

Es seien $A, B \subseteq \mathbb{R}^3$ wie in der Substitutionsregel und $f \in C(A, \mathbb{R})$. Dann gilt:

$$\int_A f(x, y, z) d(x, y, z) = \int_B f(r \cos \varphi, r \sin \varphi, z) \cdot r \, d(r, \varphi, z).$$

Beispiele: (1) Volumen eines Zylinders: Es seien $R, h > 0$ und

$$A := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : x^2 + y^2 \leq R^2, 0 \leq z \leq h \right\}.$$

Für $B := [0, R] \times [0, 2\pi] \times [0, h]$ ist $g(B) = A$. Also gilt:

$$\begin{aligned} V(A) &= \int_A 1 d(x, y, z) = \int_B r d(r, \varphi, z) \\ &= \int_0^h \left(\int_0^{2\pi} \left(\int_0^R r dr \right) d\varphi \right) dz = 2\pi h \left[\frac{1}{2} r^2 \right]_0^R = \pi R^2 h. \end{aligned}$$

(2) Es seien

$$A := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : x^2 + y^2 \leq 1, 0 \leq y \leq x, z \in [0, 1] \right\},$$

$$B := [0, 1] \times [0, \frac{\pi}{4}] \times [0, 1], \quad f(x, y, z) = x^2 + y^2 + z.$$

Es gilt $g(B) = A$, also

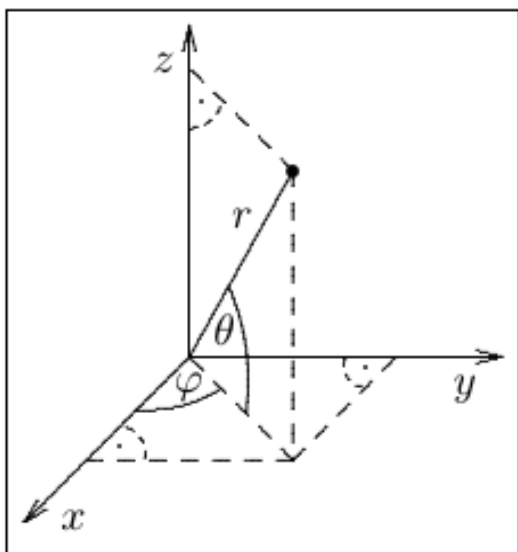
$$\begin{aligned} \int_A (x^2 + y^2 + z) d(x, y, z) &= \int_B (r^2 + z) r d(r, \varphi, z) \\ &= \int_0^{\frac{\pi}{4}} \left(\int_0^1 \left(\int_0^1 (r^3 + zr) dr \right) dz \right) d\varphi \\ &= \frac{\pi}{4} \int_0^1 \left[\frac{1}{4} r^4 + \frac{1}{2} z r^2 \right]_0^1 dz = \frac{\pi}{4} \int_0^1 \left(\frac{1}{4} + \frac{1}{2} z \right) dz = \frac{\pi}{8}. \end{aligned}$$

Kugelkoordinaten ($n = 3$):

$$x = r \cos \varphi \cos \vartheta, \quad y = r \sin \varphi \cos \vartheta, \quad z = r \sin \vartheta,$$

$$r = \sqrt{x^2 + y^2 + z^2}.$$

mit $\varphi \in [0, 2\pi]$, $\vartheta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.



Wir setzen auf $G = \mathbb{R}^3$

$$g(r, \varphi, \vartheta) := \begin{pmatrix} r \cos \varphi \cos \vartheta \\ r \sin \varphi \cos \vartheta \\ r \sin \vartheta \end{pmatrix},$$

mit

$$\det g'(r, \varphi, \vartheta) = r^2 \cos \vartheta.$$

Sind $A, B \subseteq \mathbb{R}^3$ wieder in der Substitutionsregel, so gilt für $f \in C(A, \mathbb{R})$:

$$\int_A f(x, y, z) d(x, y, z) = \int_B f(g(r, \varphi, \vartheta)) \cdot r^2 \cos \vartheta \, d(r, \varphi, \vartheta).$$

Beispiel: Es sei $f(x, y, z) = x\sqrt{x^2 + y^2 + z^2}$ und

$$A := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : x, y, z \geq 0, \, x^2 + y^2 + z^2 \leq 1 \right\}.$$

Für

$$B := \underbrace{[0, 1]}_r \times \underbrace{\left[0, \frac{\pi}{2}\right]}_\varphi \times \underbrace{\left[0, \frac{\pi}{2}\right]}_\vartheta$$

ist $g(B) = A$. Also gilt:

$$\begin{aligned} & \int_A x\sqrt{x^2 + y^2 + z^2} d(x, y, z) \\ &= \int_B (r \cos \varphi \cos \vartheta) r r^2 \cos \vartheta d(r, \varphi, \vartheta) \\ &= \int_B r^4 \cos^2 \vartheta \cos \varphi d(r, \varphi, \vartheta) \\ &= \int_0^1 \left(\int_0^{\frac{\pi}{2}} \left(\int_0^{\frac{\pi}{2}} r^4 \cos^2 \vartheta \cos \varphi d\varphi \right) d\vartheta \right) dr \\ &= \int_0^1 \left(\int_0^{\frac{\pi}{2}} r^4 \cos^2 \vartheta d\vartheta \right) dr = \frac{1}{5} \int_0^{\frac{\pi}{2}} \cos^2 \vartheta d\vartheta = \frac{\pi}{20}. \end{aligned}$$

22 Oberflächenintegrale und Integralsätze

22.1. Integralsätze im $\mathbb{R}^2$: Wir betrachten ein beschränktes Gebiet $G \subseteq \mathbb{R}^2$, welches folgende Eigenschaften besitzt, welche wir mit (V) bezeichnen:

- (1) $\overline{G}$ ist normalisierbar (also insbesondere kompakt und messbar, vgl. 21.6).
- (2) Es existieren reguläre Kurven $\gamma_1, \dots, \gamma_m$ so, dass für $\gamma := \gamma_1 + \dots + \gamma_m$ gilt:
 - (i) Spur von $\gamma = \partial G$,
 - (ii) γ ist geschlossen und doppelpunktfrei.
 - (iii) γ ist so orientiert, dass G “links von γ ” liegt.

Die Kurve γ durchläuft also ∂G nur einmal und im mathematisch positive Sinn. Unter diesen Voraussetzungen schreiben wir auch

$$\int_{\partial G} \dots \quad \text{für} \quad \int_{\gamma} \dots$$

Definition: Für die obige G umlaufende Kurve $\gamma : [a, b] \rightarrow \mathbb{R}^2$,

$$\gamma(t) = \begin{pmatrix} \gamma_1(t) \\ \gamma_2(t) \end{pmatrix}, \quad \gamma'(t) = \begin{pmatrix} \gamma'_1(t) \\ \gamma'_2(t) \end{pmatrix}$$

heißt

$$\vec{N}(\gamma(t)) := \frac{1}{\|\gamma'(t)\|} \begin{pmatrix} \gamma'_2(t) \\ -\gamma'_1(t) \end{pmatrix}$$

Normaleneinheitsvektor im Kurvenpunkt $\gamma(t)$. Eventuell ist γ an endlich vielen Stellen nicht differenzierbar. Dort kann man z.B. durch einseitige Grenzwerte $\vec{N}(\gamma(t))$ definieren. In solchen Punkten $\gamma(t)$ (und evtl. ebenso in $\gamma(a) = \gamma(b)$) ist $\vec{N}(\gamma(t))$ dann nicht eindeutig bestimmt.

Bemerkung: Der Tangenteneinheitsvektor war in 20.2 definiert als $\vec{T}(\gamma(t)) := \frac{\gamma'(t)}{\|\gamma'(t)\|}$. Für alle (bis auf evtl. endlich viele) $t \in [a, b]$ gilt:

$$(\vec{T}(\gamma(t)) | \vec{N}(\gamma(t))) = 0.$$

Integralsatz von Gauß im $\mathbb{R}^2$: Es seien G und γ wie in (V), es sei $D \subseteq \mathbb{R}^2$ offen mit $\overline{G} \subseteq D$ und $\vec{v} = (v_1, v_2) \in C^1(D, \mathbb{R}^2)$. Dann gilt

$$\int_{\partial G} \vec{v}(x, y) \cdot d(x, y) = \int_{\overline{G}} (\partial_1 v_2(x, y) - \partial_2 v_1(x, y)) d(x, y).$$

Anwendung (Flächenberechnung durch Kurvenintegrale): Es seien G und γ wie in (V), es sei $D = \mathbb{R}^2$ und $\vec{v}(x, y) = \begin{pmatrix} -y \\ x \end{pmatrix}$. Dann ist $(\partial_1 v_2)(x, y) = 1$ und $(\partial_2 v_1)(x, y) = -1$, also

$$(\partial_1 v_2)(x, y) - (\partial_2 v_1)(x, y) = 2.$$

Der Integralsatz von Gauß liefert

$$V(\overline{G}) = \frac{1}{2} \int_{\overline{G}} 2 d(x, y) = \frac{1}{2} \int_{\partial G} \begin{pmatrix} -y \\ x \end{pmatrix} \cdot d(x, y).$$

Beispiel: Sei $r > 0$ und

$$G := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : x^2 + y^2 < r^2 \right\}, \quad \text{also} \quad \overline{G} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : x^2 + y^2 \leq r^2 \right\},$$

und

$$\gamma(t) = \begin{pmatrix} r \cos t \\ r \sin t \end{pmatrix} \quad (t \in [0, 2\pi]).$$

Dann gilt

$$V(\overline{G}) = \frac{1}{2} \int_{\gamma} \begin{pmatrix} -y \\ x \end{pmatrix} \cdot d(x, y) = \frac{1}{2} \int_0^{2\pi} \begin{pmatrix} -r \sin t \\ r \cos t \end{pmatrix} \cdot \begin{pmatrix} -r \sin t \\ r \cos t \end{pmatrix} dt = \frac{1}{2} \int_0^{2\pi} r^2 dt = \pi r^2.$$

Divergenzsatz im $\mathbb{R}^2$: Es seien G und γ wie in (V), es sei $D \subseteq \mathbb{R}^2$ offen mit $\overline{G} \subseteq D$. Weiter sei $L := L(\gamma)$ und $\gamma : [0, L] \rightarrow \mathbb{R}^2$ sei in ihrer natürlichen Parametrisierung gegeben (vgl. 19.6). Für $\vec{v} \in C^1(D, \mathbb{R}^2)$ gilt dann

$$\int_{\partial G} \vec{v} \cdot \vec{N} \, ds = \int_{\overline{G}} \operatorname{div} \vec{v} \, d(x, y).$$

Bemerkung: Da $\gamma : [0, L] \rightarrow \mathbb{R}^2$ bzgl. seiner Weglänge parametrisiert ist gilt $\|\gamma'(s)\| = 1$ ($s \in [0, L]$), also

$$\int_{\gamma} \vec{v} \cdot \vec{N} \, ds = \int_0^L \vec{v}(\gamma(s)) \cdot \vec{N}(\gamma(s)) \, ds = \int_{\gamma} \begin{pmatrix} -v_2(x, y) \\ v_1(x, y) \end{pmatrix} \cdot d(x, y).$$

Weiter ist auch $\|\vec{N}(\gamma(s))\| = 1$ ($s \in [0, L]$). Die Vektoren $\vec{N}(\gamma(s))$, $\vec{T}(\gamma(s))$ bilden ein Rechtssystem im $\mathbb{R}^2$. Da G “links von γ ” liegt ist $\vec{N}(\gamma(s))$ senkrecht auf ∂G und nach außen gerichtet und wird dann auch *äußere Einheitsnormale* an G im Randpunkt $\gamma(s)$ genannt.

Beispiel: Sei $r > 0$ und

$$G := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : x^2 + y^2 < r^2 \right\}.$$

Wir parametrisieren $\partial G = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : x^2 + y^2 = r^2 \right\}$ durch

$$\gamma(s) = \begin{pmatrix} r \cos(s/r) \\ r \sin(s/r) \end{pmatrix} \quad (s \in [0, 2\pi r]).$$

Dann ist

$$\vec{N}(\gamma(s)) = \begin{pmatrix} \cos(s/r) \\ \sin(s/r) \end{pmatrix}.$$

Das Vektorfeld sei gegeben durch $\vec{v}(x, y) = \begin{pmatrix} xy \\ 2y \end{pmatrix}$. Dann gilt $\vec{v} \in C^1(\mathbb{R}^2, \mathbb{R}^2)$ und

$$\begin{aligned} \int_{\gamma} \vec{v} \cdot \vec{N} \, ds &= \int_{\overline{G}} \operatorname{div} \vec{v} \, d(x, y) = \int_{\overline{G}} (y + 2) \, d(x, y) \\ &= \int_{-r}^r \left(\int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} (y + 2) \, dy \right) dx = 4 \int_{-r}^r \sqrt{r^2 - x^2} \, dx. \end{aligned}$$

Wir substituieren $x = r\xi$, $dx = r d\xi$ und erhalten

$$\int_{\gamma} \vec{v} \cdot \vec{N} ds = 4r^2 \int_{-1}^1 \sqrt{1 - \xi^2} d\xi = 2\pi r^2.$$

Die Definition des Kurvenintegrals als Riemann-Integral liefert somit

$$\begin{aligned} 2\pi r^2 &= \int_{\gamma} \vec{v} \cdot \vec{N} ds = \int_0^{2\pi r} \begin{pmatrix} r^2 \cos(s/r) \sin(s/r) \\ 2r \sin(s/r) \end{pmatrix} \cdot \begin{pmatrix} \cos(s/r) \\ \sin(s/r) \end{pmatrix} ds \\ &= \int_0^{2\pi r} (r^2 \cos^2(s/r) \sin(s/r) + 2r \sin^2(s/r)) ds. \end{aligned}$$

Greensche Formeln Es seien G und γ wie in (V), es sei $D \subseteq \mathbb{R}^2$ offen mit $\overline{G} \subseteq D$. Weiter seien $f \in C^2(D, \mathbb{R})$, $g \in C^1(D, \mathbb{R})$.

Dann gilt die 1. Greensche Formel:

$$\int_{\partial G} g \frac{\partial f}{\partial \vec{N}} ds = \int_{\overline{G}} (g \Delta f + \nabla g \cdot \nabla f) d(x, y).$$

Zum Beweis wendet man den Divergenzsatz auf $\vec{v} := g \nabla f$ an. Man beachte

$$\vec{v} \cdot \vec{N} = g \nabla f \cdot \vec{N} = g \frac{\partial f}{\partial \vec{N}}$$

und

$$\operatorname{div} \vec{v} = \nabla \cdot \vec{v} = \nabla \cdot (g \nabla f) = \nabla g \cdot \nabla f + g(\nabla \cdot \nabla f) = \nabla g \cdot \nabla f + g(\Delta f)$$

(Produktregel aus 19.21).

Sind $f, g \in C^2(D, \mathbb{R})$, so gilt die 2. Greensche Formel:

$$\int_{\partial G} \left(g \frac{\partial f}{\partial \vec{N}} - f \frac{\partial g}{\partial \vec{N}} \right) ds = \int_{\overline{G}} (g \Delta f - f \Delta g) d(x, y).$$

Zum Beweis subtrahiere von der 1. Greenschen Formel die Formel, in der die Rollen von f und g vertauscht sind.

Beispiel: Sei $u \in C^2(D, \mathbb{R})$ in G harmonisch, d.h. $\Delta u = 0$ in G , und es sei $u = 0$ auf ∂G . Mit $f = g = u$ in der 1. Greenschen Formel erhält man

$$\int_{\overline{G}} \nabla u \cdot \nabla u d(x, y) = 0.$$

Da $\nabla u \cdot \nabla u = \|\nabla u\|^2 \geq 0$ stetig in $\overline{G}$ ist, folgt $\|\nabla u(x, y)\| = 0$ ($(x, y) \in \overline{G}$), also

$$\nabla u(x, y) = \vec{0} \quad ((x, y) \in \overline{G}).$$

Da G ein Gebiet ist, ist somit u auf G , also auch auf $\overline{G}$ konstant. Wegen $u = 0$ auf ∂G ist also $u = 0$ auf $\overline{G}$.

22.2. Flächen im $\mathbb{R}^3$: Es sei $\emptyset \neq U \subseteq \mathbb{R}^2$ ein Gebiet, es sei $\vec{g} \in C^1(U, \mathbb{R}^3)$ injektiv mit

$$\text{Rang } \vec{g}'(u, v) = \text{Rang}(\partial_u \vec{g}(u, v) \quad \partial_v \vec{g}(u, v)) = 2 \quad ((u, v) \in U).$$

Dann heißt $\mathcal{F} := \vec{g}(U)$ ein *reguläres Flächenstück im $\mathbb{R}^3$* und $\vec{g}$ eine *reguläre Parameterdarstellung* von $\mathcal{F}$.

Spezialfall (Explizite Darstellung): Sei $\vec{g} \in C^1(U, \mathbb{R}^3)$ von der Form

$$\vec{g}(x, y) = \begin{pmatrix} x \\ y \\ f(x, y) \end{pmatrix}$$

für eine Funktion $f \in C^1(U, \mathbb{R})$. Dann ist $\vec{g}$ injektiv und

$$\vec{g}'(x, y) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ f_x(x, y) & f_y(x, y) \end{pmatrix},$$

also $\text{Rang } \vec{g}'(x, y) = 2 \quad ((x, y) \in U)$. Dann heißt $\mathcal{F} := \vec{g}(U)$ ein *reguläres Flächenstück in expliziter Darstellung* und

$$\mathcal{F} = \left\{ \begin{pmatrix} x \\ y \\ f(x, y) \end{pmatrix} : \begin{pmatrix} x \\ y \end{pmatrix} \in U \right\}$$

ist der Graf von f .

Normaleneinheitsvektor: Ist $\mathcal{F}$ ein reguläres Flächenstück im $\mathbb{R}^3$, so gibt es in jedem Punkt auf der Fläche genau zwei Vektoren, die auf der Fläche senkrecht stehen und die Länge 1 haben. Sie sind entgegengesetzt gerichtet und heißen *Normaleneinheitsvektor*. Die Entscheidung für einen der beiden legt die *Orientierung des Flächenstücks* fest.

Häufig wird verlangt, dass $\vec{N}$ “nach außen” zeigt, was aber voraussetzt, dass es überhaupt “innen” und “außen” gibt. Das ist im allgemeinen nicht der Fall.

Im folgenden betrachten wir $\vec{N}$ als Abbildung $\mathcal{F} \rightarrow \mathbb{R}^3$. definiert durch

$$\vec{N}(\vec{g}(u, v)) = \frac{\partial_u \vec{g}(u, v) \times \partial_v \vec{g}(u, v)}{\|\partial_u \vec{g}(u, v) \times \partial_v \vec{g}(u, v)\|} \quad ((u, v) \in U).$$

Für die explizite Darstellung erhalten wir

$$\vec{N}(x, y, f(x, y)) = \frac{1}{\sqrt{1 + (\partial_x f(x, y))^2 + (\partial_y f(x, y))^2}} \begin{pmatrix} -\partial_x f(x, y) \\ -\partial_y f(x, y) \\ 1 \end{pmatrix} \quad ((x, y) \in U),$$

wegen

$$\partial_x \vec{g} = \begin{pmatrix} 1 \\ 0 \\ \partial_x f \end{pmatrix}, \quad \partial_y \vec{g} = \begin{pmatrix} 0 \\ 1 \\ \partial_y f \end{pmatrix}, \quad \partial_x \vec{g} \times \partial_y \vec{g} = \begin{pmatrix} -\partial_x f \\ -\partial_y f \\ 1 \end{pmatrix}.$$

22.3. Oberflächenintegral: Sei $U \subseteq \mathbb{R}^2$ ein Gebiet und $\vec{g} : U \rightarrow \mathbb{R}^3$ eine reguläre (insbesondere also injektive) Parametrisierung eines Flächenstücks.

Definition: Sei $B \subseteq U$ kompakt und messbar und $\mathcal{F}_0 := \vec{g}(B)$ (wir nennen dann $\mathcal{F}_0$ ein *kompaktes Flächenstück*). Für $f \in C(\mathcal{F}_0, \mathbb{R})$ definiert man das (*Oberflächen-*)*Integral von f über $\mathcal{F}_0$* durch

$$\int_{\mathcal{F}_0} f \, do := \int_B f(\vec{g}(u, v)) \|\partial_u \vec{g}(u, v) \times \partial_v \vec{g}(u, v)\| \, d(u, v),$$

und für $\vec{w} \in C(\mathcal{F}_0, \mathbb{R}^3)$ setzt man

$$\int_{\mathcal{F}_0} \vec{w} \cdot d\vec{o} := \int_B \vec{w}(\vec{g}(u, v)) \cdot (\partial_u \vec{g}(u, v) \times \partial_v \vec{g}(u, v)) \, d(u, v).$$

Mit $f = 1$ definiert man den *Flächeninhalt* von $\mathcal{F}_0$:

$$A(\mathcal{F}_0) := \int_{\mathcal{F}_0} 1 \, do = \int_B \|\partial_u \vec{g}(u, v) \times \partial_v \vec{g}(u, v)\| \, d(u, v).$$

Bemerkung: Nach Definition von $\vec{N}$ gilt

$$\int_{\mathcal{F}_0} \vec{w} \cdot d\vec{o} = \int_{\mathcal{F}_0} \vec{w} \cdot \vec{N} \, do,$$

das ist der **Fluss** des Vektorfelds $\vec{w}$ durch das mittels $\vec{N}$ orientierte Flächenstück $\mathcal{F}_0$.

Beispiel: $U = \mathbb{R}^2$, $f(x, y) = x^2 + y^2$,

$$\vec{g}(x, y) = \begin{pmatrix} x \\ y \\ x^2 + y^2 \end{pmatrix}, \quad B = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : x^2 + y^2 \leq 1 \right\}, \quad \mathcal{F}_0 := \vec{g}(B).$$

Dann ist

$$\vec{N}(\vec{g}(x, y)) = \frac{1}{\sqrt{4x^2 + 4y^2 + 1}} \begin{pmatrix} -2x \\ -2y \\ 1 \end{pmatrix},$$

also

$$\begin{aligned} A(\mathcal{F}_0) &= \int_{\mathcal{F}_0} 1 \, do = \int_B (4x^2 + 4y^2 + 1)^{1/2} \, d(x, y) \\ &= \int_0^{2\pi} \left(\int_0^1 (4r^2 + 1)^{1/2} r \, dr \right) d\varphi = 2\pi \int_0^1 (4r^2 + 1)^{1/2} r \, dr \end{aligned}$$

(Substitution $t = 4r^2 + 1$, $r \, dr = dt/8$)

$$= \frac{2\pi}{8} \int_1^5 \sqrt{t} \, dt = \frac{\pi}{6} (5\sqrt{5} - 1).$$

Bemerkung: Um einen Flächeninhalt wie oben zu definieren genügt es, wenn $\vec{g} \in C^1(U, \mathbb{R}^3)$ statt auf ganz U nur auf den Inneren von B injektiv ist.

Oberfläche von Rotationskörpern: Es sei $h \in C^1([a, b], \mathbb{R})$, $h > 0$ auf $[a, b]$ und

$$\vec{g}(u, v) = \begin{pmatrix} u \\ h(u) \cos(v) \\ h(u) \sin(v) \end{pmatrix} \quad ((u, v) \in B := [a, b] \times [0, 2\pi]).$$

Dann ist $\mathcal{F}_0 := \vec{g}(B)$ die Mantelfläche (ohne Deckel und Boden) des durch h erzeugten Rotationskörpers. Beachte: g ist nicht auf B aber auf $B^\circ = (a, b) \times (0, 2\pi)$ injektiv. Es gilt

$$\begin{aligned} A(\mathcal{F}_0) &= \int_B \|\partial_u \vec{g}(u, v) \times \partial_v \vec{g}(u, v)\| \, d(u, v). \\ &= \int_a^b \int_0^{2\pi} \left\| \begin{pmatrix} 1 \\ h'(u) \cos(v) \\ h'(u) \sin(v) \end{pmatrix} \times \begin{pmatrix} 0 \\ -h(u) \sin(v) \\ h(u) \cos(v) \end{pmatrix} \right\| \, dv du \\ &= \int_a^b \int_0^{2\pi} \left\| \begin{pmatrix} h(u)h'(u) \\ -h(u) \cos(v) \\ -h(u) \sin(v) \end{pmatrix} \right\| \, dv du = 2\pi \int_a^b h(u) \sqrt{1 + (h'(u))^2} \, du \end{aligned}$$

22.4. Integralsätze im $\mathbb{R}^3$:

Der Integralsatz von Stokes: Es sei $U \subseteq \mathbb{R}^2$ ein Gebiet und $\vec{g} : U \rightarrow \mathbb{R}^3$ eine reguläre Parametrisierung des Flächenstücks $\mathcal{F} := \vec{g}(U)$. Es seien $G \subseteq \mathbb{R}^2$ und γ wie in (V) mit $\overline{G} \subseteq U$ und $\mathcal{F}_0 := g(\overline{G})$. Es sei $\Gamma := g \circ \gamma$ (also ist die Spur von Γ die Menge $g(\partial G)$). Weiter sei $V \subseteq \mathbb{R}^3$ offen mit $\mathcal{F}_0 \subseteq V$ und $\vec{v} \in C^1(V, \mathbb{R}^3)$. Dann gilt

$$\int_{\Gamma} \vec{v}(x, y, z) \cdot d(x, y, z) = \int_{\mathcal{F}_0} \operatorname{rot} \vec{v} \cdot d\vec{o}.$$

Beispiel: Wir betrachten das Beispiel aus 22.3 mit $\overline{G} = B = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : x^2 + y^2 \leq 1 \right\}$ und

$$\gamma(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \quad (t \in [0, 2\pi]).$$

Dann ist

$$\Gamma(t) = g(\gamma(t)) = \begin{pmatrix} \cos t \\ \sin t \\ 1 \end{pmatrix} \quad (t \in [0, 2\pi]).$$

Betrachte

$$\vec{v}(x, y, z) := \begin{pmatrix} yz + y \\ xz \\ xy \end{pmatrix} \text{ auf } V = \mathbb{R}^3, \text{ mit } \operatorname{rot} \vec{v}(x, y, z) = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}.$$

Nach dem Satz von Stokes ist

$$\begin{aligned}\int_{\Gamma} \vec{v}(x, y, z) \cdot d(x, y, z) &= \int_{\mathcal{F}_0} \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \cdot d\vec{o} \\ &= \int_B \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} -2x \\ -2y \\ 1 \end{pmatrix} d(x, y) = - \int_B 1 d(x, y) = -V(B) = -\pi\end{aligned}$$

Wir können das hier auch direkt nachrechnen:

$$\begin{aligned}\int_{\Gamma} \vec{v}(x, y, z) \cdot d(x, y, z) &= \int_0^{2\pi} \vec{v}(\Gamma(t)) \cdot \Gamma'(t) dt = \int_0^{2\pi} \begin{pmatrix} 2 \sin t \\ \cos t \\ \sin t \cos t \end{pmatrix} \cdot \begin{pmatrix} -\sin t \\ \cos t \\ 0 \end{pmatrix} dt \\ &= \int_0^{2\pi} (\cos^2 t - 2 \sin^2 t) dt = -\pi.\end{aligned}$$

Der Divergenzsatz im $\mathbb{R}^3$: Sei $B \subseteq \mathbb{R}^3$ kompakt und messbar und $G := B \setminus \partial B$ ein Gebiet mit $\partial G = \partial B$ (dann ist $\overline{G} = B$). Der Rand ∂G lasse sich zerlegen in endlich viele kompakte Flächenstücke. Die Einheitsnormale $\vec{N}$ auf ∂G sei ins Äußere von G gerichtet.

Satz: Sei $V \subseteq \mathbb{R}^3$ offen mit $\overline{G} \subseteq V$ und $\vec{v} \in C^1(V, \mathbb{R}^3)$. Dann gilt

$$\int_{\overline{G}} \operatorname{div} \vec{v}(x, y, z) d(x, y, z) = \int_{\partial G} \vec{v} \cdot \vec{N} do.$$

Beispiele: (1) Für $\vec{v} \in C^2(V, \mathbb{R}^3)$ gilt wegen $\operatorname{div} \operatorname{rot} \vec{v} = 0$ in V :

$$\int_{\partial G} (\nabla \times \vec{v}) \cdot d\vec{o} = \int_{\partial G} (\operatorname{rot} \vec{v}) \cdot \vec{N} do = \int_{\overline{G}} \operatorname{div} \operatorname{rot} \vec{v}(x, y, z) d(x, y, z) = 0.$$

(2) Für $f \in C^2(V)$ gilt wegen $\Delta f = \operatorname{div} \nabla f$ in V :

$$\int_{\overline{G}} \Delta f(x, y, z) d(x, y, z) = \int_{\partial G} \nabla f \cdot d\vec{o} = \int_{\partial G} \nabla f \cdot \vec{N} do = \int_{\partial G} \frac{\partial f}{\partial \vec{N}} do.$$

(3) **Greensche Formeln:** Für $f, g \in C^2(V, \mathbb{R})$ und $h \in C^1(V, \mathbb{R})$ gilt:

$$\begin{aligned}\int_{\partial G} h \frac{\partial f}{\partial \vec{N}} do &= \int_{\overline{G}} \left(h \Delta f + \nabla h \cdot \nabla f \right) d(x, y, z), \\ \int_{\partial G} \left(g \frac{\partial f}{\partial \vec{N}} - f \frac{\partial g}{\partial \vec{N}} \right) do &= \int_{\overline{G}} \left(g \Delta f - f \Delta g \right) d(x, y, z).\end{aligned}$$

Bemerkung: In allen Integralsätzen können die Flächen- bzw. Volumenintegrale $\int_{\overline{G}}$ auch durch $\int_G$ ersetzt werden, da der Rand einer meßbaren Menge eine Nullmenge ist, vgl. 21.3.

23 Grundzüge der Funktionentheorie

23.1. Komplexe Differenzierbarkeit und Holomorphie: Der Raum $\mathbb{C}$ als reeller Vektorraum versehen mit dem komplexen Betrag $|\cdot|$ ist der Raum $\mathbb{R}^2$ versehen mit der Euklidischen Norm $\|\cdot\|$. Alle Begriffe die in $(\mathbb{R}^2, \|\cdot\|)$ definiert sind stehen damit in $(\mathbb{C}, |\cdot|)$ zur Verfügung (z.B. “offene, abgeschlossene, kompakte oder konvexe Mengen”, “stetige oder differenzierbare Funktionen” in Sinne von 19.3, 19.4 und 19.9, etc.).

Beispiel: Ist $I \subseteq \mathbb{R}$ ein Intervall und $g : I \rightarrow \mathbb{C}$ eine Funktion mit $u := \operatorname{Re} g : I \rightarrow \mathbb{R}$ und $v := \operatorname{Im} g : I \rightarrow \mathbb{R}$ (also $g = u + iv$), so gilt: g ist auf I differenzierbar genau dann wenn u und v auf I differenzierbar sind. In diesem Fall gilt

$$g'(t) = u'(t) + iv'(t) \quad (t \in I).$$

Hinzu kommt folgende

Definition: Eine Funktion $g : [a, b] \rightarrow \mathbb{C}$ heißt (Riemann-)integrierbar $:\Leftrightarrow u, v \in R([a, b], \mathbb{R})$. In diesem Fall schreiben wir $g \in R([a, b], \mathbb{C})$ und setzen

$$\int_a^b g(t) dt := \int_a^b u(t) dt + i \int_a^b v(t) dt.$$

Uneigentliche Integrale definiert man entsprechend.

Bemerkungen: (1) $R([a, b], \mathbb{C})$ ist ein komplexer Vektorraum und

$$g \mapsto \int_a^b g(t) dt$$

ist ein lineares Funktional. Sind $g_1, g_2 \in R([a, b], \mathbb{C})$, so ist $g_1 g_2 \in R([a, b], \mathbb{C})$.

(2) Es gelten die Integrationsregeln aus HM I (Hauptsatz, Substitutionsregel, partielle Integration, etc.)

(3) Ist $g \in R([a, b], \mathbb{C})$, so gilt $|g| \in R([a, b], \mathbb{R})$ und

$$\left| \int_a^b g(t) dt \right| \leq \int_a^b |g(t)| dt.$$

Beispiel: Für $w = x + iy \in \mathbb{C} \setminus \{0\}$ mit $x, y \in \mathbb{R}$ und $a < b$ gilt

$$\int_a^b e^{wt} dt = \frac{e^{wb} - e^{wa}}{w}.$$

Wegen

$$\begin{aligned} \frac{d}{dt} e^{wt} &= \frac{d}{dt} (e^{xt} (\cos(yt) + i \sin(yt))) \\ &= x e^{xt} (\cos(yt) + i \sin(yt)) + e^{xt} (-y \sin(yt) + iy \cos(yt)) \\ &= (x + iy) e^{xt} (\cos(yt) + i \sin(yt)) = w e^{wt} \end{aligned}$$

folgt das aus dem Hauptsatz.

Wir definieren nun einen **neuen** Differenzierbarkeitsbegriff:

Definition: Es sei $G \subseteq \mathbb{C}$ offen. Eine Funktion $f : G \rightarrow \mathbb{C}$ heißt in $z_0 \in G$ *komplex differenzierbar*, wenn der Limes

$$f'(z_0) := \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

in $\mathbb{C}$ existiert. In diesem Fall heißt $f'(z_0)$ die *komplexe Ableitung von f in z_0* .

Die Funktion f heißt *holomorph* in G , falls f in **jedem** $z_0 \in G$ komplex differenzierbar ist. Wir schreiben dann $f \in H(G)$. Funktionen aus $H(\mathbb{C})$ heißen *ganze Funktionen*.

Bemerkungen: (1) Ist f in G holomorph, so ist f in G stetig.

(2) Die Definition höherer Ableitungen erfolgt wie üblich.

(3) Sei $f : G \rightarrow \mathbb{C}$ eine Funktion. Ist $F \in H(G)$ und gilt $F' = f$ auf G , so heißt F eine *Stammfunktion* von f auf G .

23.2. Rechenregeln: Sei $G \subseteq \mathbb{C}$ offen, seien $f, g \in H(G)$ und $\alpha, \beta \in \mathbb{C}$. Dann gilt $\alpha f + \beta g, f \cdot g \in H(G)$ und

$$(\alpha f + \beta g)'(z) = \alpha f'(z) + \beta g'(z), \quad (f \cdot g)'(z) = f'(z)g(z) + f(z)g'(z) \quad (z \in G).$$

Insbesondere ist $H(G)$ ein komplexer Vektorraum.

Ist $g \neq 0$ in G , so ist auch $\frac{f}{g}$ holomorph in G und

$$\left(\frac{f}{g}\right)'(z) = \frac{f'(z)g(z) - f(z)g'(z)}{g(z)^2} \quad (z \in G).$$

Kettenregel: Ist $f \in H(G_1)$, $g \in H(G_2)$ und $f(G_1) \subseteq G_2$, so ist $g \circ f \in H(G_1)$ und

$$(g \circ f)'(z) = g'(f(z))f'(z) \quad (z \in G_1).$$

Die Beweise des reellen Falles lassen sich übertragen.

Beispiele: (1) $z \mapsto f(z) = z^n$, $n \in \mathbb{N}$ ist in $H(\mathbb{C})$ mit $f'(z) = nz^{n-1}$ ($z \in \mathbb{C}$). Damit folgt

$$\mathbb{C}[z] \subseteq H(\mathbb{C}).$$

(2) $z \mapsto f(z) = z^{-n}$, $n \in \mathbb{N}$ ist in $H(\mathbb{C} \setminus \{0\})$ mit $f'(z) = -nz^{-n-1}$ ($z \in \mathbb{C} \setminus \{0\}$). Damit folgt: Jede rationale Funktion

$$z \mapsto f(z) = \frac{p(z)}{q(z)} \quad \text{mit } p \in \mathbb{C}[z], \quad q \in \mathbb{C}[z] \setminus \{0\}$$

ist in $H(\mathbb{C} \setminus \{z \in \mathbb{C} : q(z) = 0\})$.

(3) $z \mapsto f(z) = e^z$ ist in $H(\mathbb{C})$ mit $f'(z) = e^z$ ($z \in \mathbb{C}$).

(4) $z \mapsto f(z) = e^{1/z}$ ist in $H(\mathbb{C} \setminus \{0\})$ mit $f'(z) = -z^{-2}e^{1/z}$ ($z \neq 0$).

23.3. Cauchy-Riemannsche Differentialgleichungen (CRD): Sei $G \subseteq \mathbb{C}$ offen und $f : G \rightarrow \mathbb{C}$. Wir betrachten G als offene Teilmenge von $\mathbb{R}^2$. Durch $u(x, y) := (\operatorname{Re} f)(x + iy)$ und $v(x, y) = (\operatorname{Im} f)(x + iy)$ für $x, y \in \mathbb{R}$ mit $x + iy \in G$ erhält man eine Funktion

$$G \rightarrow \mathbb{R}^2, \quad (x, y) \mapsto \begin{pmatrix} u(x, y) \\ v(x, y) \end{pmatrix},$$

die man mit den Methoden von Kapitel 19 (mehrdimensionale Differentialrechnung) behandeln kann.

Satz: Es sei $z_0 = x_0 + iy_0 \in G$. Dann sind äquivalent:

- (1) f ist in z_0 komplex differenzierbar;
- (2) u und v sind in (x_0, y_0) reell differenzierbar **und** es gelten die Cauchy-Riemannschen Differentialgleichungen

$$u_x(x_0, y_0) = v_y(x_0, y_0), \quad u_y(x_0, y_0) = -v_x(x_0, y_0).$$

In diesem Fall gilt:

$$f'(z_0) = u_x(x_0, y_0) + iv_x(x_0, y_0).$$

Beweis. Es sei $h = h_1 + ih_2 \in \mathbb{C} \setminus \{0\}$ ($h_1, h_2 \in \mathbb{R}$) so, dass $z_0 + h \in G$. Für $c = a + ib$ ($a, b \in \mathbb{R}$) sei

$$Q(h) := \frac{f(z_0 + h) - f(z_0) - ch}{|h|} = \frac{h}{|h|} \left(\frac{f(z_0 + h) - f(z_0)}{h} - c \right).$$

Dann gilt (nachrechnen):

$$R(h) := \operatorname{Re} Q(h) = \frac{u(x_0 + h_1, y_0 + h_2) - u(x_0, y_0) - (h_1 a - h_2 b)}{|h|},$$

$$I(h) := \operatorname{Im} Q(h) = \frac{v(x_0 + h_1, y_0 + h_2) - v(x_0, y_0) - (h_1 b + h_2 a)}{|h|}.$$

Damit folgt: f ist in z_0 komplex differenzierbar mit $f'(z_0) = c$

$$\iff Q(h) \rightarrow 0 \ (h \rightarrow 0) \iff R(h) \rightarrow 0 \text{ und } I(h) \rightarrow 0 \ (h \rightarrow 0)$$

$$\iff u \text{ und } v \text{ sind in } (x_0, y_0) \text{ reell differenzierbar mit}$$

$$u_x(x_0, y_0) = a, \quad u_y(x_0, y_0) = -b, \quad v_x(x_0, y_0) = b, \quad v_y(x_0, y_0) = a.$$

□

Beispiel: Betrachte $z = x + iy \mapsto f(z) = \bar{z} = x - iy$ ($z \in \mathbb{C}$). Hier ist $u(x, y) = x$ und $v(x, y) = -y$. Damit gilt für jedes $(x, y) \in \mathbb{R}^2$:

$$u_x(x, y) = 1 \neq -1 = v_y(x, y).$$

Also ist f in keinem $z \in \mathbb{C}$ komplex differenzierbar.

Bemerkung: Wir werden später sehen (23.8): Ist $f \in H(G)$, so ist f beliebig oft komplex differenzierbar. Aus den Cauchy-Riemannschen Differentialgleichungen folgt damit für u und v :

$$\Delta u = u_{xx} + u_{yy} = v_{yx} - v_{xy} = 0, \quad \Delta v = v_{xx} + v_{yy} = -u_{yx} + u_{xy} = 0,$$

d.h. Real- und Imaginärteil einer holomorphen Funktion sind harmonisch.

23.4. Potenzreihen: Es sei

$$\sum_{n=0}^{\infty} a_n (z - z_0)^n$$

eine Potenzreihe mit Konvergenzradius $R \in (0, \infty]$ und Entwicklungspunkt $z_0 \in \mathbb{C}$. Wir setzen

$$K(z_0, R) := \{z \in \mathbb{C} : |z - z_0| < R\}, \quad K(z_0, \infty) := \mathbb{C}.$$

Dann gilt:

(1) Die durch die Potenzreihe gegebene Funktion

$$f : K(z_0, R) \rightarrow \mathbb{C}, \quad z \mapsto f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n,$$

in $H(K(z_0, R))$, und es gilt

$$f'(z) = \sum_{n=1}^{\infty} n a_n (z - z_0)^{n-1} \quad (z \in K(z_0, R)).$$

Beachte: $\sum_{n=0}^{\infty} a_n (z - z_0)^n$ und $\sum_{n=1}^{\infty} n a_n (z - z_0)^{n-1}$ haben denselben Konvergenzradius.

(2) Durch induktive Anwendung von (1) folgt: f ist auf $K(z_0, R)$ beliebig oft komplex differenzierbar und

$$\forall n \in \mathbb{N}_0 : a_n = \frac{f^{(n)}(z_0)}{n!}$$

Beispiele: Die Funktionen

$$z \mapsto \exp z, \quad z \mapsto \sin z, \quad z \mapsto \cos z, \quad z \mapsto \sinh z, \quad z \mapsto \cosh z$$

sind in $H(\mathbb{C})$ (also ganze Funktionen).

23.5. Komplexe Kurvenintegrale: Wie in 20.1 verstehen wir unter einer *Kurve* eine stetige Abbildung $\gamma : [a, b] \rightarrow \mathbb{C}$, für die es $a = t_0 < t_1 < \dots < t_n = b$ so gibt, dass γ auf jedem Intervall $[t_{j-1}, t_j]$ ($j = 1, \dots, n$) stetig differenzierbar ist. Die Kurve γ heißt *geschlossen*, falls $\gamma(a) = \gamma(b)$ gilt und *doppelpunktfrei*, falls γ auf $[a, b]$ injektiv ist. Eine geschlossene doppelpunktfreie Kurve heißt *positiv orientiert*, wenn das von γ umlaufene Gebiet links von γ liegt.

Bemerkung: Sei $G \subseteq \mathbb{C}$ offen, $f : G \rightarrow \mathbb{C}$ holomorph und $\gamma : [a, b] \rightarrow G$ differenzierbar. Dann ist $f \circ \gamma : [a, b] \rightarrow \mathbb{C}$ differenzierbar und

$$(f \circ \gamma)'(t) = f'(\gamma(t))\gamma'(t) \quad (t \in [a, b]).$$

Definition: Sei $\gamma : [a, b] \rightarrow \mathbb{C}$ eine Kurve und $f \in C(\gamma([a, b]), \mathbb{C})$. Dann definiert man das Kurvenintegral

$$\int_{\gamma} f(z) dz := \sum_{j=1}^n \int_{t_{j-1}}^{t_j} f(\gamma(t))\gamma'(t) dt.$$

Ist $\gamma \in C^1([a, b], \mathbb{C})$, so ist also

$$\int_{\gamma} f(z) dz = \int_a^b f(\gamma(t))\gamma'(t) dt.$$

Das Integral ist invariant unter orientierungserhaltenden Umparametrisierungen und ändert das Vorzeichen bei Orientierungsumkehr.

Abschätzung: Sind f und γ wie in der Definition, so gilt

$$\left| \int_{\gamma} f(z) dz \right| \leq \int_a^b |f(\gamma(t))\gamma'(t)| dt \leq L(\gamma) \max\{|f(z)| : z \in \gamma([a, b])\},$$

wobei $L(\gamma) = \int_a^b |\gamma'(t)| dt$ die Länge von γ ist (siehe 19.6).

Wichtiges Beispiel: Sei $r > 0$ und $\gamma : [0, 2\pi] \rightarrow \mathbb{C}$, $\gamma(t) = re^{it}$. Dann ist γ eine geschlossene, doppelpunktfreie und positiv orientierte Kurve, und es gilt $\gamma'(t) = ire^{it}$ ($t \in [0, 2\pi]$). Sei $f(z) = z^n$, wobei $n \in \mathbb{Z}$. Dann gilt

$$\int_{\gamma} z^n dz = \int_0^{2\pi} (\gamma(t))^n \gamma'(t) dt = \int_0^{2\pi} r^n e^{itn} ir e^{it} dt = ir^{n+1} \int_0^{2\pi} e^{i(n+1)t} dt.$$

Für $n = -1$ ist das Kurvenintegral also $= 2\pi i$. Für $n \neq -1$ erhalten wir

$$\int_{\gamma} z^n dz = ir^{n+1} \left[\frac{e^{i(n+1)t}}{i(n+1)} \right]_0^{2\pi} = 0,$$

da $e^{2\pi i(n+1)} = 1$. Beachte: $f \in H(\mathbb{C})$ für $n \geq 0$ und $f \in H(\mathbb{C} \setminus \{0\})$ für $n < 0$.

Aus dem Hauptsatz erhalten wir:

Satz: Es sei $G \subseteq \mathbb{C}$ ein Gebiet, $f \in H(G)$ und $\gamma : [a, b] \rightarrow G$ eine Kurve. Hat f auf G eine Stammfunktion F , so gilt

$$\int_{\gamma} f(z) dz = F(\gamma(b)) - F(\gamma(a)).$$

Beweis. Es gilt:

$$\int_{\gamma} f(z) dz = \int_{\gamma} F'(z) dz = \int_a^b F'(\gamma(t))\gamma'(t) dt = \int_a^b (F \circ \gamma)'(t) dt = F(\gamma(b)) - F(\gamma(a)).$$

□

23.6. Cauchyscher Integralsatz (ohne Beweis): Es sei $G \subseteq \mathbb{C}$ ein einfach zusammenhängendes Gebiet (also z.B. konvex oder sternförmig) und $f \in H(G)$. Dann hat f auf G eine Stammfunktion.

Bemerkung: Ist also $G \subseteq \mathbb{C}$ ein einfach zusammenhängendes Gebiet und $\gamma : [a, b] \rightarrow G$ eine geschlossene Kurve, so gilt (vgl. den Satz in 23.5):

$$\forall f \in H(G) : \int_{\gamma} f(z) dz = 0.$$

Beispiel: Im Beispiel in 23.5 besitzen die Funktionen $f(z) = z^n$ eine Stammfunktion auf $\mathbb{C}$ falls $n \geq 0$ und auf $\mathbb{C} \setminus \{0\}$ falls $n \leq -2$ ist. Die Funktion

$$f(z) = \frac{1}{z} \text{ hat keine Stammfunktion auf } \mathbb{C} \setminus \{0\},$$

denn z.B. für die geschlossene Kurve $\gamma(t) = e^{it}$ ($t \in [0, 2\pi]$) ist

$$\int_{\gamma} \frac{1}{z} dz = 2\pi i \neq 0.$$

Beachte: $\mathbb{C} \setminus \{0\}$ ist nicht einfach zusammenhängend.

23.7. Umlaufzahl und Cauchysche Integralformel: Es sei $\gamma : [a, b] \rightarrow \mathbb{C}$ eine geschlossene Kurve. Wir setzen $\Omega := \mathbb{C} \setminus \gamma([a, b])$ und definieren $\text{ind}_{\gamma} : \Omega \rightarrow \mathbb{C}$ durch

$$\text{ind}_{\gamma}(z_0) := \frac{1}{2\pi i} \int_{\gamma} \frac{1}{z - z_0} dz \quad (z_0 \in \Omega).$$

Die Kurve γ zerlegt $\mathbb{C}$ in Gebiete, die sogenannten *Zusammenhangskomponenten* von Ω .

Satz: Es gilt $\text{ind}_\gamma(\Omega) \subseteq \mathbb{Z}$ und $\text{ind}_\gamma = 0$ auf der unbeschränkten Zusammenhangskomponente von Ω . Weiter ist ind_γ auf jeder Zusammenhangskomponente von Ω konstant.

Ohne Beweis.

Definition Die Zahl $\text{ind}_\gamma(z_0)$ heißt *Index* oder *Umlaufzahl* von γ in z_0 .

Bemerkung: $\text{ind}_\gamma(z_0)$ gibt an wie oft und mit welcher Orientierung die Kurve γ den Punkt z_0 umläuft. Ist γ eine geschlossene, doppelpunktfreie und positiv orientierte Kurve, so ist die Umlaufzahl 0 (auf der unbeschränkten Zusammenhangskomponente) und sonst 1.

Beispiele: (1) Sei $\gamma(t) = a + re^{it}$ ($t \in [0, 2\pi]$). Dann gilt

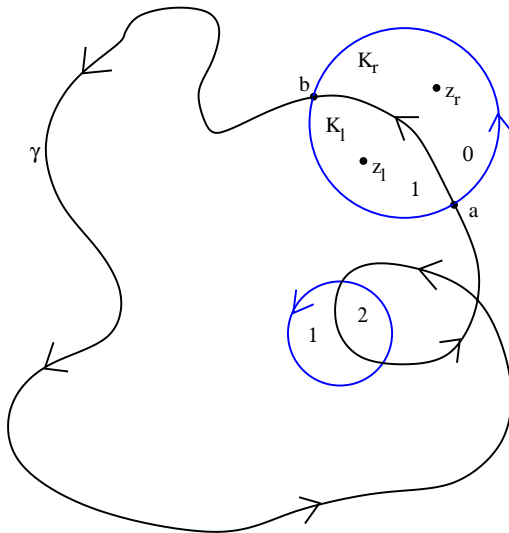
$$\text{ind}_\gamma(z_0) = \begin{cases} 1, & |z_0 - a| < r \\ 0, & |z_0 - a| > r \end{cases}.$$

Beweis: $\Omega = \mathbb{C} \setminus \{z : |z - a| = r\}$, und $\{z : |z - a| > r\}$ ist die unbeschränkte Zusammenhangskomponente von Ω . Für $|z_0 - a| < r$ gilt

$$\text{ind}_\gamma(z_0) = \text{ind}_\gamma(a) = \frac{1}{2\pi i} \int_\gamma \frac{dz}{z - a} = \frac{1}{2\pi i} \int_0^{2\pi} \frac{rie^{it}}{re^{it}} dt = 1.$$

(2) Graphische Bestimmung der Umlaufzahl (ohne Beweis):

Fährt man entlang der Kurve, so ist stets auf der linken Seite die Umlaufzahl um 1 größer als auf der rechten Seite.



Cauchysche Integralformel: Sei $G \subseteq \mathbb{C}$ ein einfach zusammenhängendes Gebiet, $f \in H(G)$ und $\gamma : [a, b] \rightarrow G$ eine geschlossene Kurve. Dann gilt:

$$\forall z_0 \in G \setminus \gamma([a, b]) : f(z_0) \text{ind}_\gamma(z_0) = \frac{1}{2\pi i} \int_\gamma \frac{f(z)}{z - z_0} dz.$$

Beweisidee: Man kann zeigen, dass 23.6 auch noch gilt, wenn die Funktion auf G stetig und in $G \setminus \{z_0\}$ holomorph ist. Diese Variante von 23.6 wendet man auf die Funktion

$$z \mapsto \begin{cases} \frac{f(z)-f(z_0)}{z-z_0}, & z \in G \setminus \{z_0\} \\ f'(z_0), & z = z_0 \end{cases}$$

an:

$$0 = \frac{1}{2\pi i} \int_{\gamma} \frac{f(z) - f(z_0)}{z - z_0} dz = \frac{1}{2\pi i} \int_{\gamma} \frac{f(z)}{z - z_0} dz - f(z_0) \underbrace{\frac{1}{2\pi i} \int_{\gamma} \frac{1}{z - z_0} dz}_{=\text{ind}_{\gamma}(z_0)}.$$

□

Beispiel: Es sei $\gamma(t) = 2e^{it}$ ($t \in [0, 2\pi]$). Berechne

$$\int_{\gamma} \frac{e^z - z}{z - 1} dz.$$

Setze $f(z) = e^z - z$ und beachte $\text{ind}_{\gamma}(1) = 1$. Also gilt

$$\int_{\gamma} \frac{e^z - z}{z - 1} dz = \int_{\gamma} \frac{f(z)}{z - 1} dz = 2\pi i f(1) = 2\pi i(e - 1).$$

23.8. Folgerungen: (Ohne Beweis)

(a) Es sei $G \subseteq \mathbb{C}$ offen und $f \in H(G)$. Dann gilt

$$f', f'', f''', \dots \in H(G),$$

d.h. f ist auf G beliebig oft komplex differenzierbar.

(b) Sind G , f und γ wie in 23.7, so gilt

$$\forall k \in \mathbb{N}_0 \quad \forall z_0 \in G \setminus \gamma([a, b]) : f^{(k)}(z_0) \text{ind}_{\gamma}(z_0) = \frac{k!}{2\pi i} \int_{\gamma} \frac{f(z)}{(z - z_0)^{k+1}} dz.$$

(c) Holomorphe Funktionen lassen sich **lokal in Potenzreihen entwickeln**. Ist G offen, $f \in H(G)$ und $z_0 \in G$, so gilt

$$f(z) = \sum_{k=0}^{\infty} \frac{f^{(k)}(z_0)}{k!} (z - z_0)^k \quad (z \in K(z_0, R)),$$

für jedes $R > 0$ mit $K(z_0, R) \subseteq G$. Die Reihe konvergiert dabei für jedes $z \in K(z_0, R)$ absolut und für jedes $\rho \in (0, R)$ auf $\{z \in \mathbb{C} : |z - z_0| \leq \rho\}$ gleichmäßig.

Beachte: Teil (c) ist in $\mathbb{R}$ falsch. Die Funktion

$$f(x) = \begin{cases} e^{-1/x^2}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

ist in $C^\infty(\mathbb{R})$ mit $f^{(k)}(0) = 0$ ($k \in \mathbb{N}_0$).

Beispiel: Betrachte $G = \mathbb{C} \setminus \{1\}$, $f(z) = \frac{1}{1-z}$ und $z_0 = i$. Es gilt $|z_0 - 1| = \sqrt{2}$. Damit gilt $K(i, \sqrt{2}) \subseteq G$ (dies ist der größte Kreis mit Mittelpunkt 1 der in G liegt). Es gilt

$$\begin{aligned} f(z) &= \frac{1}{1-z} = \frac{1}{1-i-(z-i)} = \frac{1}{1-i} \cdot \frac{1}{1-\frac{z-i}{1-i}} \\ &= \frac{1}{1-i} \sum_{k=0}^{\infty} \left(\frac{z-i}{1-i} \right)^k = \sum_{k=0}^{\infty} \frac{1}{(1-i)^{k+1}} (z-i)^k \\ &\text{für } \frac{|z-i|}{|1-i|} < 1, \text{ also für } |z-i| < |1-i| = \sqrt{2}. \end{aligned}$$

23.9. Satz von Liouville: Ist $f \in H(\mathbb{C})$ und ist f auf $\mathbb{C}$ beschränkt, so ist f konstant.

Beweis. Es gelte $|f(z)| \leq M$ für alle $z \in \mathbb{C}$. Wir schreiben nach 23.8:

$$f(z) = \sum_{k=0}^{\infty} a_k z^k \quad (z \in \mathbb{C}).$$

Sei $r > 0$ und $\gamma(t) = re^{it}$ ($t \in [0, 2\pi]$). Mit 23.8 (b) und der Abschätzung in 23.5 folgt (bea. $L(\gamma) = 2\pi r$):

$$|a_k| = \left| \frac{1}{2\pi i} \int_{\gamma} \frac{f(z)}{z^{k+1}} dz \right| = \frac{1}{2\pi} \left| \int_{\gamma} \frac{f(z)}{z^{k+1}} dz \right| \leq \frac{1}{2\pi} \cdot 2\pi r \cdot \frac{M}{r^{k+1}} = \frac{M}{r^k} \quad (k \in \mathbb{N}_0).$$

Für $r \rightarrow \infty$ erhält man $a_k = 0$ ($k \geq 1$). □

Bemerkung: $\sin$ und $\cos$ sind auf $\mathbb{R}$ beschränkt. Der Satz von Liouville zeigt, dass $\sin$ und $\cos$ auf $\mathbb{C}$ unbeschränkt sind. Tatsächlich gilt $\sin(\mathbb{C}) = \cos(\mathbb{C}) = \mathbb{C}$.

Anwendungen: (1) Jedes komplexe Polynom p vom Grad ≥ 1 hat eine Nullstelle in $\mathbb{C}$ (Fundamentalsatz der Algebra).

Beweis. Es gilt $|P(z)| \rightarrow \infty$ für $|z| \rightarrow \infty$. Hat p keine Nullstelle in $\mathbb{C}$, so ist $f : \mathbb{C} \rightarrow \mathbb{C}$, $f(z) := 1/p(z)$, holomorph mit $f(z) \rightarrow 0$ für $|z| \rightarrow \infty$. Insbesondere ist f beschränkt und damit konstant. Dann ist aber auch p konstant im Widerspruch zur Voraussetzung. □

(2) Sei $f \in H(\mathbb{C})$ und es gelte

$$|f(z)| \leq \frac{1}{e^x} \quad (z = x + iy \in \mathbb{C}).$$

Dann ist f von der Form $f(z) = ce^{-z}$ für ein $c \in \mathbb{C}$ mit $|c| \leq 1$.

Beweis. Sei $g(z) := e^z f(z)$. Dann ist $g \in H(\mathbb{C})$ und für $z = x + iy \in \mathbb{C}$ gilt:

$$|g(z)| \leq |e^z| \frac{1}{e^x} = e^x \frac{1}{e^x} = 1.$$

Also ist g beschränkt und somit konstant, d.h. $g(z) = g(0) =: c$ ($z \in \mathbb{C}$), also $f(z) = ce^{-z}$ ($z \in \mathbb{C}$) mit $|c| = |f(0)| \leq 1$. $\square$

23.10. Identitätssatz: Es sei $G \subseteq \mathbb{C}$ ein Gebiet und $f, g \in H(G)$. Hat die Menge

$$\{z \in G : f(z) = g(z)\}$$

einen Häufungspunkt in G , so gilt $f = g$ auf G .

Beweis. Ohne Einschränkung sei $g = 0$. Wir finden $z_0 \in G$ und eine Folge (z_n) in $G \setminus \{z_0\}$ mit $z_n \rightarrow z_0$ und $f(z_n) = 0$ ($n \in \mathbb{N}$). Wir schreiben

$$f(z) = \sum_{k=0}^{\infty} a_k (z - z_0)^k$$

in einer Umgebung von z_0 und nehmen an, es gäbe ein minimales $m \in \mathbb{N}_0$ mit $a_m \neq 0$. Dann gilt für fast alle $n \in \mathbb{N}$:

$$0 = f(z_n) = \frac{f(z_n)}{(z_n - z_0)^m} = \sum_{k=m}^{\infty} a_k (z_n - z_0)^{k-m} \longrightarrow a_m \quad (n \rightarrow \infty)$$

im Widerspruch zu $a_m \neq 0$. Folglich ist $a_k = 0$ ($k \in \mathbb{N}_0$), also $f = 0$ auf einer Kreisscheibe um z_0 .

Verbinde nun ein beliebiges $z \in G$ mit z_0 durch einen Streckenzug in G und wiederhole das Argument entlang des Streckenzugs. $\square$

Folgerungen: Es sei $G \subseteq \mathbb{C}$ ein Gebiet und $f \in H(G)$.

(1) Ist $f' = 0$ auf G , so ist f konstant auf G .

(2) Ist f auf G nicht konstant, $z_0 \in G$ und $f(z_0) = 0$, so existiert ein $\delta > 0$ mit

$$\forall z \in K(z_0, \delta) \setminus \{0\} : f(z) \neq 0,$$

d.h. Nullstellen von f liegen in G isoliert.

Beispiele: (1) Ist $f \in H(\mathbb{C})$ und $f(1/n) = 0$ ($n \in \mathbb{N}$), so ist $f = 0$ auf $\mathbb{C}$.

(2) Sei $G = \mathbb{C} \setminus \{0\}$ und $f(z) = \sin(\pi/z)$. Dann ist $f \in H(G)$ und

$$f(1/n) = \sin(\pi n) = 0 \quad (n \in \mathbb{N}).$$

Dies widerspricht nicht dem Identitätssatz, denn 0 ist zwar ein Häufungspunkt der Nullstellenmenge von f , aber $0 \notin G$.

(3) Betrachte $z \mapsto f(z) = \sin^2(z) + \cos^2(z)$. Dann ist $f \in H(\mathbb{C})$. Bekannt: $f(x) = 1$ ($x \in \mathbb{R}$). Also gilt

$$\mathbb{R} \subseteq \{z \in \mathbb{C} : f(z) = 1\}.$$

Jedes $x \in \mathbb{R}$ ist ein Häufungspunkt von $\{z \in \mathbb{C} : f(z) = 1\}$ in $\mathbb{C}$. Also folgt $f(z) = 1$ ($z \in \mathbb{C}$).

(4) Es gibt kein $f \in H(\mathbb{C})$ mit $f(1/n) = (-1)^n/n$ ($n \in \mathbb{N}$). (Übung.)

23.11. Isolierte Singularitäten und Laurententwicklung:

Definition: Es sei $G \subseteq \mathbb{C}$ offen, $z_0 \in G$ und $f \in H(G \setminus \{z_0\})$. Dann heißt z_0 eine *isolierte Singularität* von f . Eine isolierte Singularität z_0 von f heißt

(1) eine *hebbare Singularität* von f : $\iff$

$$\exists g \in H(G) : f(z) = g(z) \quad (z \in G \setminus \{z_0\}),$$

d.h. f lässt sich auf G holomorph fortsetzen. Die Fortsetzung g wird dann oft auch wieder mit f bezeichnet.

(2) ein *Pol* von f : $\iff$

$$|f(z)| \rightarrow \infty \quad (z \rightarrow z_0).$$

(3) eine *wesentliche Singularität* von f : $\iff$ z_0 ist keine hebbare Singularität und kein Pol von f .

Beispiele: In folgenden Beispielen sei stets $G = \mathbb{C}$ und $z_0 = 0$.

(1) Betrachte

$$f(z) = \frac{\sin z}{z}.$$

Für $z \neq 0$ ist

$$f(z) = \frac{1}{z} \left(z - \frac{z^3}{3!} + \frac{z^5}{5!} - \dots \right) = g(z) := 1 - \frac{z^2}{3!} + \frac{z^4}{5!} - \dots$$

Es gilt $g \in H(\mathbb{C})$ und $f = g$ auf $\mathbb{C} \setminus \{0\}$. Also hat f in 0 eine hebbare Singularität.

(2) Sei $n \in \mathbb{N}$ und

$$f(z) = \frac{1}{z^n}.$$

Dann hat f in 0 einen Pol.

(3) Betrachte

$$f(z) = \exp\left(\frac{1}{z}\right).$$

Es gilt $f(1/n) = e^n \rightarrow \infty$ ($n \rightarrow \infty$), also ist 0 nicht hebbar. Weiter gilt $|f(1/(in))| = 1$ ($n \in \mathbb{N}$), also ist 0 kein Pol. Somit hat f in 0 eine wesentliche Singularität.

Im folgenden sei $z_0 \in G$ eine isolierte Singularität von $f \in H(G \setminus \{z_0\})$.

Riemannscher Hebbarkeitssatz: Folgende Aussagen sind äquivalent:

- (1) f hat in z_0 eine hebbare Singularität.
- (2) Es gibt ein $R > 0$ mit $K(z_0, R) \subseteq G$ und f ist auf $K(z_0, R) \setminus \{z_0\}$ beschränkt.
- (3) $\lim_{z \rightarrow z_0} f(z)$ existiert in $\mathbb{C}$.

Beweis. (1) $\Rightarrow$ (2) und (3) $\Rightarrow$ (2) sind offensichtlich. Wir zeigen (2) $\Rightarrow$ (1),(3): O.B.d.A. betrachten wir den Fall $z_0 = 0$. Es sei $h : G \rightarrow \mathbb{C}$ definiert durch

$$h(z) = \begin{cases} z^2 f(z), & z \neq 0 \\ 0, & z = 0 \end{cases}.$$

Dann ist $h \in H(G \setminus \{0\})$ und da f in einem Kreis um 0 beschränkt ist folgt

$$\frac{h(z) - h(0)}{z - 0} = z f(z) \rightarrow 0 \quad (z \rightarrow 0).$$

Also gilt $h \in H(G)$ mit $h'(0) = 0$. Somit gilt

$$h(z) = \sum_{n=0}^{\infty} a_n z^n \quad (z \in K(0, R)),$$

mit $a_n = h^{(n)}(0)/n!$, also $a_0, a_1 = 0$. Es folgt

$$f(z) = a_2 + a_3 z + a_4 z^2 + \dots \quad (z \in K(0, R) \setminus \{0\})$$

Setzt man

$$g(z) = \begin{cases} f(z), & z \neq 0 \\ a_2, & z = 0 \end{cases},$$

so ist $g \in H(G)$ und $f = g$ auf $G \setminus \{0\}$. □

Charakterisierung von Polstellen (ohne Beweis): Es gilt: z_0 ist ein Pol von f

$$\Leftrightarrow \exists m \in \mathbb{N} : \lim_{z \rightarrow z_0} (z - z_0)^m f(z) \text{ existiert und ist } \neq 0.$$

Dieses m ist eindeutig bestimmt und heit *die Ordnung des Pols* z_0 .

Beispiele: (1) Sei $n \in \mathbb{N}$ und $f(z) = 1/z^n$. Dann gilt

$$z^n f(z) \rightarrow 1 \quad (z \rightarrow 0),$$

also hat f in 0 einen Pol der Ordnung n .

(2) Betrachte $G = K(0, 1)$ und

$$f(z) = \frac{e^z}{z \sin z} \quad (z \in K(0, 1) \setminus \{0\}).$$

Es gilt

$$z^2 f(z) = \frac{z}{\sin z} e^z \rightarrow 1 \cdot e^0 = 1 \quad (z \rightarrow 0),$$

also hat f in 0 einen Pol der Ordnung 2.

Definition: Es sei $(a_n)_{n \in \mathbb{Z}}$ eine Folge in $\mathbb{C}$. Beachte hierbei: Auch eine Abbildung mit Definitionsbereich $\mathbb{Z}$ nennen wir Folge.

Eine Reihe der Form

$$\sum_{n=-\infty}^{\infty} a_n (z - z_0)^n := \sum_{n=0}^{\infty} a_n (z - z_0)^n + \sum_{n=1}^{\infty} \frac{a_{-n}}{(z - z_0)^n},$$

heit eine *Laurentreihe*. Die Reihe

$$\sum_{n=1}^{\infty} \frac{a_{-n}}{(z - z_0)^n}$$

heit *Hauptteil* der Laurentreihe, und die Reihe

$$\sum_{n=0}^{\infty} a_n (z - z_0)^n$$

heit *Nebenteil* der Laurentreihe. Konvergenz einer Laurentreihe bedeutet also Konvergenz von Hauptteil **und** Nebenteil.

Satz und Definition (ohne Beweis): Sei $f \in H(G \setminus \{z_0\})$ und $R > 0$ mit $K(z_0, R) \subseteq G$. Dann existiert eine eindeutig bestimmte Folge $(a_n)_{n \in \mathbb{Z}}$ in $\mathbb{C}$ mit

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n \quad (z \in K(z_0, R) \setminus \{z_0\}).$$

Weiter gilt:

(1) z_0 ist eine hebbare Singularitt von $f \iff a_{-n} = 0 \ (n \in \mathbb{N})$.

- (2) z_0 ist ein Pol der Ordnung $m \in \mathbb{N} \iff a_{-m} \neq 0$ und $a_{-n} = 0$ ($n > m$).
 (3) z_0 ist eine wesentliche Singularität von $f \iff a_{-n} \neq 0$ für unendlich viele $n \in \mathbb{N}$.

Der Koeffizient a_{-1} heißt das *Residuum von f in z_0* und wird mit

$$\operatorname{res}(f, z_0) := a_{-1}$$

bezeichnet.

Beispiele: (1) Betrachte $n \in \mathbb{N}$ und $f(z) = z^{-n}$. Dann gilt $\operatorname{res}(z^{-n}, 0) = 1$ für $n = 1$ und $\operatorname{res}(z^{-n}, 0) = 0$ für $n \geq 2$.

(2) $f(z) = e^{2/z}$ hat in $z_0 = 0$ eine wesentliche Singularität. Für $z \in \mathbb{C} \setminus \{0\}$ ist

$$f(z) = \sum_{n=0}^{\infty} \frac{2^n}{n!} z^{-n} = 1 + \sum_{n=1}^{\infty} \frac{2^n}{n!} \frac{1}{z^n} \quad (= \text{Nebenteil} + \text{Hauptteil}).$$

Es gilt also $\operatorname{res}(f, 0) = 2$.

(3) Betrachte

$$f(z) = \frac{1}{z(z-1)} \quad (z \in \mathbb{C} \setminus \{0, 1\}).$$

Es gilt

$$\lim_{z \rightarrow 0} z f(z) = -1 \neq 0, \quad \lim_{z \rightarrow 1} (z-1) f(z) = 1 \neq 0,$$

also hat f in 0 und 1 jeweils einen Pol erster Ordnung. Wir bestimmen die Laurententwicklung in $z_0 = 1$: Für $0 < |z-1| < 1$ gilt

$$\begin{aligned} f(z) &= \frac{1}{z-1} \cdot \frac{1}{z} = \frac{1}{z-1} \cdot \frac{1}{1-(1-z)} = \frac{1}{z-1} \sum_{n=0}^{\infty} (1-z)^n \\ &= \frac{1}{z-1} \sum_{n=0}^{\infty} (-1)^n (z-1)^n = \frac{1}{z-1} + \sum_{n=1}^{\infty} (-1)^n (z-1)^{n-1} \\ &= \frac{1}{z-1} + \sum_{n=0}^{\infty} (-1)^{n+1} (z-1)^n \quad (= \text{Hauptteil} + \text{Nebenteil}). \end{aligned}$$

Es gilt also $\operatorname{res}(f, 1) = 1$.

Berechnung des Residuums bei Polen: f habe in z_0 einen Pol 1. Ordnung, also

$$f(z) = \frac{a_{-1}}{z-z_0} + \sum_{n=0}^{\infty} a_n (z-z_0)^n \quad (z \in K(z_0, R) \setminus \{z_0\}).$$

Dann gilt

$$(z-z_0)f(z) = a_{-1} + \sum_{n=0}^{\infty} a_n (z-z_0)^{n+1} \rightarrow a_{-1} \quad (z \rightarrow z_0),$$

also

$$\operatorname{res}(f, z_0) = \lim_{z \rightarrow z_0} (z - z_0) f(z).$$

Allgemein gilt: Besitzt f in z_0 einen höchstens n -fachen Pol ($n \in \mathbb{N}$), dann gilt

$$\operatorname{res}(f, z_0) = \lim_{z \rightarrow z_0} \frac{1}{(n-1)!} \left(\frac{d^{n-1}}{dz^{n-1}} \left((z - z_0)^n f(z) \right) \right).$$

23.12. Residuensatz: Es sei $G \subseteq \mathbb{C}$ ein einfach zusammenhängendes Gebiet. Ist $S \subseteq G$ eine Menge ohne Häufungspunkt in G (z.B. S endlich), $f \in H(G \setminus S)$ und $\gamma : [a, b] \rightarrow G \setminus S$ eine geschlossene Kurve, so ist die Menge $\{s \in S : \operatorname{ind}_\gamma(s) \neq 0\}$ endlich und es gilt

$$\frac{1}{2\pi i} \int_\gamma f(z) dz = \sum_{s \in S} \operatorname{res}(f, s) \operatorname{ind}_\gamma(s).$$

Ohne Beweis.

Beispiel: Sei $f(z) = \frac{1}{z^2+1}$. Da $z \mapsto z^2 + 1 = (z+i)(z-i)$ einfache Nullstellen in i und $-i$ hat, hat f einfache Polstellen in i und $-i$. Beachte: $\mathbb{C}$ ist einfach zusammenhängend und $f \in H(\mathbb{C} \setminus \{-i, i\})$. Wir erhalten

$$\operatorname{res}(f, i) = \lim_{z \rightarrow i} (z - i) f(z) = \lim_{z \rightarrow i} \frac{1}{z + i} = \frac{1}{2i} \quad \text{und} \quad \operatorname{res}(f, -i) = \lim_{z \rightarrow -i} (z + i) f(z) = -\frac{1}{2i}.$$

Ist z.B. $\gamma(t) = 2e^{it}$ ($t \in [0, 2\pi]$), so ist $\operatorname{ind}_\gamma(i) = \operatorname{ind}_\gamma(-i) = 1$, und nach dem Residuensatz ist

$$\int_\gamma \frac{1}{z^2+1} dz = 2\pi i (\operatorname{res}(f, i) + \operatorname{res}(f, -i)) = 0.$$

Für $\gamma(t) = i + e^{it}$ ($t \in [0, 2\pi]$) ist $\operatorname{ind}_\gamma(i) = 1$ aber $\operatorname{ind}_\gamma(-i) = 0$ und wir erhalten

$$\int_\gamma \frac{1}{z^2+1} dz = 2\pi i \operatorname{res}(f, i) = \pi.$$

Bemerkung: Die Cauchysche Integralformel ist ein Spezialfall des Residuensatzes, denn $g : G \setminus \{z_0\} \rightarrow \mathbb{C}$, $g(z) = f(z)/(z - z_0)$ hat in z_0 entweder eine hebbare Singularität (falls $f(z_0) = 0$) oder einen Pol 1. Ordnung. In beiden Fällen ist

$$\operatorname{res}(g, z_0) = \lim_{z \rightarrow z_0} (z - z_0) g(z) = f(z_0).$$

23.13. Satz von der Gebietstreue und Maximumprinzip:

(a) Ist $G \subseteq \mathbb{C}$ ein Gebiet und $f : G \rightarrow \mathbb{C}$ holomorph und nicht konstant, so ist $f(G) \subseteq \mathbb{C}$ wieder ein Gebiet. Ist zusätzlich f injektiv, so ist $f^{-1} \in H(f(G))$.

(b) Sei $G \subseteq \mathbb{C}$ ein Gebiet und $f : G \rightarrow \mathbb{C}$ holomorph. Hat $z \mapsto |f(z)|$ in $z_0 \in G$ ein lokales Maximum, so ist f konstant.

Bemerkung: Mit (b) folgt: Ist G ein beschränktes Gebiet, $f : \overline{G} \rightarrow \mathbb{C}$ stetig und $f \in H(G)$, so gilt

$$\max_{z \in \overline{G}} |f(z)| = \max_{z \in \partial G} |f(z)|.$$

Beweis. Sei O.B.d.A. f nicht konstant. Da $\overline{G}$ kompakt ist hat $|f|$ eine globale Maximalstelle z_0 in $\overline{G}$. Mit (b) folgt $z_0 \in \partial G$, sonst wäre f auf G also auf $\overline{G}$ konstant. $\square$

23.14. Konforme Abbildungen: Sei $G \subseteq \mathbb{C}$ ein Gebiet und $f \in H(G)$ injektiv, so heißt f *konform*.

Satz (Ohne Beweis): Sei $f \in H(G)$ konform. Dann gilt:

(a) $\forall z \in G : f'(z) \neq 0$.

(b) f ist *winkeltreu*. D.h. Sind $\gamma_1, \gamma_2 : [-1, 1] \rightarrow G$ stetig differenzierbare reguläre Kurven mit $\gamma_1(0) = z_0 = \gamma_2(0)$, die sich in z_0 im Winkel φ schneiden, so schneiden sich die Bildkurven $f \circ \gamma_1$ und $f \circ \gamma_2$ in $f(z_0)$ ebenfalls im Winkel φ .

Anwendung: In der zweidimensionalen Elektrostatik schneiden sich Feldlinien und Äquipotentiallinien im rechten Winkel. Diese Beziehung bleibt unter konformen Abbildungen erhalten. Durch konforme Abbildungen kann man die Berechnung von elektrischen Potentialen auf einfachere Situationen transformieren: Ist $f : G_1 \rightarrow G_2$ eine konforme surjektive Abbildung zwischen zwei Gebieten $G_1, G_2 \subseteq \mathbb{C}$ und ist $u : G_2 \rightarrow \mathbb{R}$ harmonisch, so ist auch $v : G_1 \rightarrow \mathbb{R}$, $v(x, y) = u(f(x + iy))$ harmonisch.

Beispiele: (1) Ist $G \subseteq \mathbb{C}$ ein konvexes Gebiet und

$$\operatorname{Re} f'(z) > 0 \quad (z \in G),$$

so ist f konform.

Beweis. Es seien $z, w \in G$, $z \neq w$ und $\gamma(t) = w + t(z - w)$ ($t \in [0, 1]$). Aus

$$f(z) - f(w) = \int_0^1 f'(\gamma(t)) \gamma'(t) dt = (z - w) \int_0^1 f'(\gamma(t)) dt$$

folgt:

$$\frac{f(z) - f(w)}{z - w} = \underbrace{\int_0^1 \operatorname{Re} f'(\gamma(t)) dt}_{>0} + i \int_0^1 \operatorname{Im} f'(\gamma(t)) dt \neq 0,$$

also $f(z) \neq f(w)$. $\square$

Betrachte z.B. $G = \{z \in \mathbb{C} : \operatorname{Im} z \in (0, \pi)\}$. Für $f(z) = z + e^z$ gilt ($z = x + iy$):

$$\operatorname{Re}(-if'(z)) = \operatorname{Im} f'(z) = e^x \sin y > 0 \quad (z \in G).$$

Also ist $-if$ und damit auch f konform auf G .

(2) **Möbiustransformationen:** Sind $a, b, c, d \in \mathbb{C}$ mit $ad - bc \neq 0$, so heißt

$$T(z) = \frac{az + b}{cz + d}$$

eine *Möbiustransformation*.

Es ist hier sinnvoll, zur komplexen Ebene $\mathbb{C}$ einen Punkt ∞ hinzuzunehmen und $\hat{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$ zu betrachten, wobei man setzt $\frac{1}{\infty} := 0$ und $\frac{1}{0} := \infty$. Außerdem bezeichnet man Kreise und Geraden als *verallgemeinerte Kreise*.

Eigenschaften:

- (1) Jede Möbiustransformation bildet $\hat{\mathbb{C}}$ bijektiv auf $\hat{\mathbb{C}}$ ab (mit $T(\infty) := a/c$).
- (2) Jede Möbiustransformation lässt sich schreiben als Hintereinanderausführung von Drehstreckungen $z \mapsto \alpha z$, Translationen $z \mapsto z + \beta$ und Inversion $z \mapsto 1/z$.
- (3) Die Menge der Möbiustransformationen ist eine Gruppe bzgl. der Komposition.
- (4) Jede Möbiustransformation bildet verallgemeinerte Kreise wieder auf verallgemeinerte Kreise ab.
- (5) Jede Möbiustransformation ist durch das Bild von drei verschiedenen Punkten aus $\hat{\mathbb{C}}$ eindeutig bestimmt.

Beispiel: Die Möbiustransformation T mit $T(1) = \infty$, $T(0) = 1$ und $T(-1) = 0$ ist gegeben durch $T(z) = \frac{1+z}{1-z}$. Es gilt $T(i) = i$. T bildet das Innere des Einheitskreises auf die rechte Halbebene und die Einheitskreislinie auf $(i\mathbb{R}) \cup \{\infty\}$ ab.

Bemerkung: Ordnet man invertierbaren Matrizen

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

die Möbiustransformation $T_A(z) = \frac{az+b}{cz+d}$ zu (beachte: $\det A = ad - bc \neq 0$), so gilt

$$T_{A_1} \circ T_{A_2} = T_{A_1 A_2}.$$

Damit kann (3) bewiesen werden.

Schwarzsches Lemma: Es sei $f \in H(K(0,1))$, $f(0) = 0$ und $|f(z)| \leq 1$ ($z \in K(0,1)$). Dann gilt

$$|f(z)| \leq |z| \quad (z \in K(0,1))$$

und $|f'(0)| \leq 1$. Weiter gilt $|f(z_0)| = |z_0|$ für ein $z_0 \in K(0,1) \setminus \{0\}$ g.d.w. $f(z) = cz$ mit $|c| = 1$. Ebenso gilt $|f'(0)| = 1$ g.d.w. $f(z) = cz$ mit $|c| = 1$.

Beweis. Wir können f als Potenzreihe schreiben:

$$f(z) = \sum_{k=0}^{\infty} a_k z^k$$

mit Konvergenzradius ≥ 1 und $a_0 = 0$. Also ist

$$g(z) = \frac{f(z)}{z} = \sum_{k=0}^{\infty} a_{k+1} z^k$$

in $K(0,1)$ holomorph (fortgesetzt mit $g(0) := a_1$). Es sei $z \in K(0,1)$ und dann $|z| < \rho < 1$. Dann gilt (Maximumprinzip)

$$|g(z)| \leq \max_{|\xi|=\rho} |g(\xi)| = \max_{|\xi|=\rho} \frac{|f(\xi)|}{|\xi|} \leq \frac{1}{\rho}$$

und $\rho \rightarrow 1-$ liefert $|g(z)| \leq 1$, also $|f(z)| \leq |z|$ und damit auch

$$\lim_{z \rightarrow 0} \left| \frac{f(z)}{z} \right| = |f'(0)| \leq 1.$$

Gilt $|f(z_0)| = |z_0|$ für ein $z_0 \in K(0,1) \setminus \{0\}$, so hat $|g|$ in z_0 ein Maximum. Also ist g konstant und somit ist f von der Form

$$f(z) = \frac{f(z_0)}{z_0} z \quad (z \in K(0,1)).$$

Ist $|f'(0)| = 1$, so ist $|g(0)| = 1$ und wie oben folgt

$$f(z) = f'(0)z \quad (z \in K(0,1)).$$

□

Mit dem Lemma von Schwarz können die konformen Abbildungen charakterisiert werden, die $K(0,1)$ auf sich abbilden:

Definition: Sei $a \in K(0,1)$. Die Möbiustransformation

$$\varphi_a(z) := \frac{z - a}{1 - \bar{a}z}$$

heißt ein *Blaschkefaktor*.

Übung: $\varphi_a : K(0, 1) \rightarrow K(0, 1)$ ist holomorph und bijektiv, und für die Umkehrfunktion gilt $\varphi_a^{-1} = \varphi_{-a}$.

Ist nun φ_a ein Blaschkefaktor und $c \in \mathbb{C}$ mit $|c| = 1$, so ist $f : K(0, 1) \rightarrow K(0, 1)$, $f(z) = c\varphi_a(z)$ holomorph und bijektiv.

Umgekehrt gilt: Ist $f : K(0, 1) \rightarrow K(0, 1)$ holomorph und bijektiv, so existieren $a \in K(0, 1)$ und $c \in \partial K(0, 1)$ mit $f = c\varphi_a$.

Beweis. Mit f ist auch $f^{-1} : K(0, 1) \rightarrow K(0, 1)$ holomorph und bijektiv.

1. Fall: Es sei $f(0) = 0$. Dann ist auch $f^{-1}(0) = 0$ und mit dem Lemma von Schwarz folgt

$$|f(z)| \leq |z|, \quad |f^{-1}(z)| \leq |z| \quad (|z| < 1).$$

Also gilt $|f(z)| = |z|$ ($|z| < 1$) und es folgt $f(z) = cz = c\varphi_0(z)$ für ein c mit $|c| = 1$.

2. Fall: Es sei $f(0) \neq 0$. Dann ist $a := f^{-1}(0) \neq 0$. Nun ist

$$g := f \circ \varphi_{-a} : K(0, 1) \rightarrow K(0, 1)$$

holomorph und bijektiv und $g(0) = f(\varphi_{-a}(0)) = f(a) = 0$. Nach dem 1. Fall ist $g(z) = cz$ für ein c mit $|c| = 1$, also

$$f(z) = c\varphi_{-a}^{-1}(z) = c\varphi_a(z) \quad (z \in K(0, 1)).$$

□

Der Riemannsche Abbildungssatz: Ist $G \subseteq \mathbb{C}$ ein einfach zusammenhängendes Gebiet mit $G \neq \mathbb{C}$, so existiert eine surjektive konforme Abbildung $f : G \rightarrow K(0, 1)$.

Ohne Beweis.

Bemerkungen: (1) Die Abbildung $f : G \rightarrow K(0, 1)$ ist also holomorph und bijektiv und $f^{-1} : K(0, 1) \rightarrow G$ ist dann ebenfalls holomorph und bijektiv. Ist $\tilde{G} \subseteq \mathbb{C}$, $\tilde{G} \neq \mathbb{C}$ ein weiteres einfach zusammenhängendes Gebiet, so existiert eine bijektive holomorphe Abbildung $g : \tilde{G} \rightarrow K(0, 1)$. Nun ist

$$g^{-1} \circ f : G \rightarrow \tilde{G}$$

eine holomorphe bijektive Abbildung von G auf $\tilde{G}$. Man sagt auch: Einfach zusammenhängende Gebiete ($\neq \mathbb{C}$) sind *konform äquivalent*.

(2) $G = \mathbb{C}$ kann nicht konform auf $K(0, 1)$ abgebildet werden: Eine holomorphe Funktion $f : \mathbb{C} \rightarrow K(0, 1)$ ist nach dem Satz von Liouville 23.9 konstant.

23.15. Logarithmus und allgemeine Potenz: Auf dem Streifen

$$G := \{z \in \mathbb{C} : |\operatorname{Im} z| < \pi\}$$

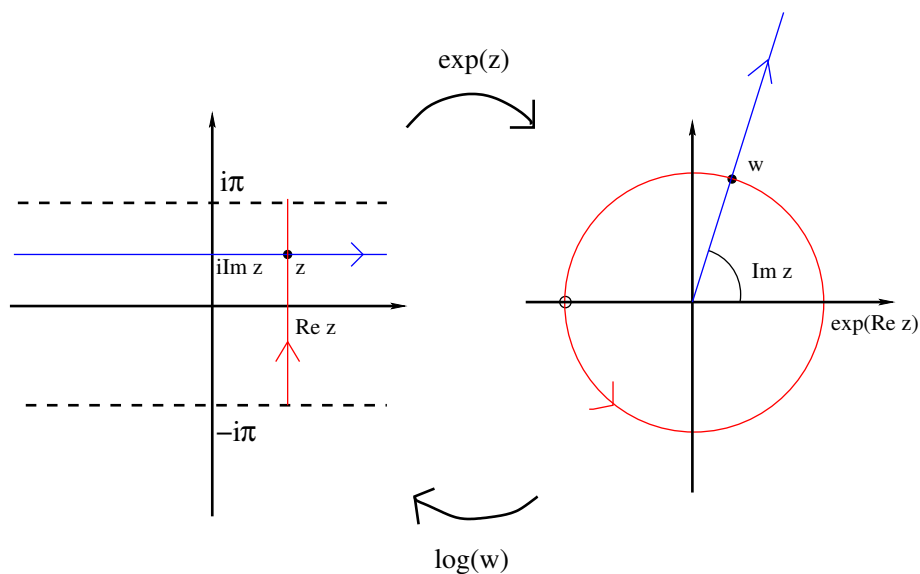
ist $\exp$ injektiv und $\exp(G) = \mathbb{C} \setminus (-\infty, 0]$.

Für die zugehörige Umkehrfunktion $\log : \mathbb{C} \setminus (-\infty, 0] \rightarrow \mathbb{C}$ gilt:

$$\log \in H(\mathbb{C} \setminus (-\infty, 0]), \quad \log z = \log |z| + i \arg z,$$

$$\log' z = \frac{1}{z} \quad (z \in \mathbb{C} \setminus (-\infty, 0]).$$

Bemerkung: $\arg$ bezeichnet den Hauptwert des Arguments (vgl. HM I), also $\arg z \in (-\pi, \pi]$ ($z \in \mathbb{C} \setminus \{0\}$), und die Funktion $\log : \mathbb{C} \setminus (-\infty, 0] \rightarrow \mathbb{C}$ heißt *Hauptwert des Logarithmus*. Der Logarithmus bildet also die geschlitzte Ebene $\mathbb{C} \setminus (-\infty, 0]$ konform auf den Streifen G ab.



Beweis. Aus $w = e^z$ ($w \neq 0$ gegeben z gesucht) folgt:

$$|w| = |e^z| = e^{\operatorname{Re} z} \Rightarrow \operatorname{Re} z = \log |w|$$

(reeller log). Wegen

$$\begin{aligned} w &= |w|(\cos(\arg w) + i \sin(\arg w)) = e^z \\ &= e^{\operatorname{Re} z}(\cos(\operatorname{Im} z) + i \sin(\operatorname{Im} z)) \end{aligned}$$

erfüllt $\operatorname{Im} z = \arg w$ das Verlangte, d.h. für $z = \log |w| + i \arg w$ gilt $e^z = w$.

Ist nun $w \in \mathbb{C} \setminus (-\infty, 0]$, so hat w in G das eindeutig bestimmte Urbild $z = \log |w| + i \arg w =: \log w$.

Auf $\mathbb{C} \setminus (-\infty, 0]$ ist $\log$ stetig und für $w_0, w_0 + h \in \mathbb{C} \setminus (-\infty, 0]$, $h \neq 0$, $z_0 := \log w_0$ und $z(h) := \log(w_0 + h)$ gilt:

$$\frac{\log(w_0 + h) - \log w_0}{h} = \frac{z(h) - z_0}{e^{z(h)} - e^{z_0}} \xrightarrow{h \rightarrow 0} \frac{1}{e^{z_0}} = \frac{1}{w_0},$$

also $\log \in H(\mathbb{C} \setminus (-\infty, 0])$ und

$$\log'(z) = \frac{1}{z} \quad (z \in \mathbb{C} \setminus (-\infty, 0]).$$

□

Es ist nun möglich Potenzen komplexer Zahlen zu definieren:

$$z^\alpha := e^{\alpha \log z} \quad (\alpha \in \mathbb{C}, \quad z \in \mathbb{C} \setminus (-\infty, 0]).$$

Vorsicht: Nicht alle aus dem Reellen bekannten Rechenregeln gelten für Potenzen komplexer Zahlen. Z.B.

$$(e^{2\pi i})^i = 1^i = 1 \neq e^{2\pi i^2} = e^{-2\pi}.$$

Die Funktion

$$\sqrt{z} := e^{\frac{1}{2} \log z} = \sqrt{|z|} \left(\cos\left(\frac{1}{2} \arg z\right) + i \sin\left(\frac{1}{2} \arg z\right) \right)$$

auf $\mathbb{C} \setminus (-\infty, 0]$ heißt *Hauptwert der Quadratwurzel*. Sie bildet $\mathbb{C} \setminus (-\infty, 0]$ konform auf die rechte Halbebene $\{z \in \mathbb{C} : \operatorname{Re} z > 0\}$ ab.

Mit dem Identitätssatz 23.10 kann man zeigen (vgl. die Formel für die komplexe Wurzel aus HM I):

$$\sqrt{z} = \sqrt{|z|} \frac{z + |z|}{|z + |z||} \quad (z \in \mathbb{C} \setminus (-\infty, 0]).$$

Beispiel: Gesucht ist eine konforme Abbildung die $G := \mathbb{C} \setminus (-\infty, -1]$ konform auf den Einheitskreis $K(0, 1)$ abbildet.

$f_1(z) := z + 1$ bildet G konform auf $\mathbb{C} \setminus (-\infty, 0]$ ab.

$f_2(z) := \sqrt{z}$ bildet $\mathbb{C} \setminus (-\infty, 0]$ konform auf $\{z \in \mathbb{C} : \operatorname{Re} z > 0\}$ ab.

$f_3(z) := \frac{z-1}{z+1}$ bildet $\{z \in \mathbb{C} : \operatorname{Re} z > 0\}$ konform auf $K(0, 1)$ ab. Beachte: f_3 ist die Umkehrabbildung des Beispiels $T(z) = \frac{1+z}{1-z}$ in 23.14.

Also bildet

$$f(z) := (f_3 \circ f_2 \circ f_1)(z) = \frac{\sqrt{z+1} - 1}{\sqrt{z+1} + 1}$$

G konform auf $K(0, 1)$ ab.

Holomorphe Logarithmen: Die Gleichung $e^v = w$ hat bei gegebenem $w \in \mathbb{C} \setminus \{0\}$ unendlich viele Lösungen $v \in \mathbb{C}$. Eine Möglichkeit allgemeine holomorphe Logarithmen zu definieren liefert der Cauchysche Integralsatz:

Satz: Es sei $G \subseteq \mathbb{C}$ einfach zusammenhängend, $f \in H(G)$ und $0 \notin f(G)$. Weiter sei $z_0 \in G$ und $w := f(z_0) = e^v$. Dann existiert genau ein $g \in H(G)$ mit

$$g(z_0) = v \quad \text{und} \quad f(z) = e^{g(z)} \quad (z \in G).$$

Die Funktion g heißt auch ein *holomorpher Logarithmus* von f .

Beweis. Es sei $h : G \rightarrow \mathbb{C}$ definiert durch $h := f'/f$. Nach 23.6 hat h eine Stammfunktion auf G . Es sei g die eindeutig bestimmte Stammfunktion von h mit $g(z_0) = v$ (Addition einer geeigneten Konstanten). Dann gilt

$$(e^{-g}f)' = -e^{-g}g'f + e^{-g}f' = -e^{-g}hf + e^{-g}f' = -e^{-g}f' + e^{-g}f' = 0$$

auf G , also $f = ce^g$ auf G für eine Konstante $c \in \mathbb{C}$. Wegen

$$e^{g(z_0)} = e^v = w = f(z_0) = ce^{g(z_0)}$$

ist $c = 1$. Also ist $e^g = f$ auf G . Ist $\tilde{g} \in H(G)$ eine weitere Funktion mit obigen Eigenschaften, so gilt

$$e^{g(z)-\tilde{g}(z)} = 1 \quad (z \in G), \quad g(z_0) - \tilde{g}(z_0) = 0.$$

Dann ist $g(z) - \tilde{g}(z) \in 2\pi i\mathbb{Z}$ ($z \in G$). Da $g - \tilde{g}$ holomorph (also insbesondere stetig) ist folgt $g - \tilde{g} = 0$ auf G . $\square$

Beispiel: Für $G = \mathbb{C} \setminus (-\infty, 0]$, $f(z) = z$, $z_0 = 1$ und $v = 0$ liefert obiger Satz: Es gibt genau ein $g \in H(\mathbb{C} \setminus (-\infty, 0])$ mit

$$e^{g(z)} = z \quad (z \in \mathbb{C} \setminus (-\infty, 0]) \quad \text{und} \quad g(1) = 0.$$

Dieses g ist der Hauptwert des Logarithmus: $g = \log$.

24 Fourierreihen

Ziel dieses Abschnitts ist es, periodische Funktionen $f : \mathbb{R} \rightarrow \mathbb{C}$, $t \mapsto f(t)$, der Periodenlänge $T > 0$ (man denke etwa an ein periodisches Audiosignal) als Überlagerung von T -periodischen “reinen” Schwingungen $t \mapsto e^{i\omega t}$ darzustellen. Die “Frequenzen” ω sind dann von der Form $\omega = \frac{2\pi k}{T}$ mit $k \in \mathbb{Z}$. Für $k = 0$ hat man einen konstanten Anteil, die anderen Frequenzen sind ganzzahlige Vielfache der Grundfrequenz $\frac{2\pi}{T}$ (Tonhöhe). Die Anteile der Oberschwingungen bestimmen die Klangfarbe.

Die Idee geht zurück auf Fourier (1822): “Jede 2π -periodische Funktion f lässt sich darstellen als Reihe $\sum_{k \in \mathbb{Z}} c_k e^{ikt}$.”

Es hat einige Zeit gedauert, diese Idee zu präzisieren (Funktionsbegriff, geeigneter Integralbegriff, verschiedene Konvergenzbegriffe etc). Wir kümmern uns hier um die Fragen, wie man die c_k aus f erhält und für welche f man eine punktweise Konvergenz der Fourierreihe hat.

24.1. T -periodische Funktionen: Sei $T > 0$. Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{C}$ heißt T -periodisch, falls $f(t+T) = f(t)$ ($t \in \mathbb{R}$).

Es gilt dann auch $f(t+kT) = f(t)$ ($t \in \mathbb{R}, k \in \mathbb{Z}$).

Bemerkungen: (a) Eine T -periodische Funktion $f : \mathbb{R} \rightarrow \mathbb{C}$ ist eindeutig bestimmt, wenn man ihre Werte auf einem Intervall $[a, a+T)$ kennt (hierbei ist $a \in \mathbb{R}$ beliebig). Jede Funktion $\tilde{f} : [a, a+T) \rightarrow \mathbb{C}$ lässt sich eindeutig zu einer T -periodischen Funktion $f : \mathbb{R} \rightarrow \mathbb{C}$ mit $f = \tilde{f}$ auf $[a, a+T)$ fortsetzen.

(b) Ist $f : \mathbb{R} \rightarrow \mathbb{C}$ eine T -periodische Funktion, so ist $g : \mathbb{R} \rightarrow \mathbb{C}$, definiert durch $g(t) := f(\frac{T}{2\pi}t)$, eine 2π -periodische Funktion:

$$\forall t \in \mathbb{R} : g(t+2\pi) = f\left(\frac{T}{2\pi}(t+2\pi)\right) = f\left(\frac{T}{2\pi}t + T\right) = f\left(\frac{T}{2\pi}t\right) = g(t).$$

Wir werden uns deshalb i.w. auf 2π -periodische Funktionen beschränken.

Beispiele: $\sin$, $\cos$ und $t \mapsto e^{it}$ sind 2π -periodische Funktionen. Die Funktion $t \mapsto \cos(2t)$ ist 2π -periodisch, aber auch π -periodisch.

Bemerkungen: (1) Ist $f : \mathbb{R} \rightarrow \mathbb{C}$ eine 2π -periodische Funktion und ist $f \in R([-\pi, \pi], \mathbb{C})$, so ist f über beschränkten Intervallen integrierbar, und es gilt für jedes $a \in \mathbb{R}$:

$$\int_a^{a+2\pi} f(t) dt = \int_{-\pi}^{\pi} f(t) dt.$$

(2) Für $l, k \in \mathbb{Z}$ gilt

$$\int_{-\pi}^{\pi} e^{ikt} e^{-ilt} dt = \int_{-\pi}^{\pi} e^{i(k-l)t} dt = \begin{cases} 2\pi, & k = l \\ 0, & k \neq l \end{cases}.$$

24.2. Trigonometrische Polynome: Sei $n \in \mathbb{N}$. Eine 2π -periodische Funktion $f : \mathbb{R} \rightarrow \mathbb{C}$ heißt *trigonometrisches Polynom vom Grad $\leq n \in \mathbb{N}$* , falls es Koeffizienten $c_k \in \mathbb{C}$ ($|k| \leq n$) gibt mit

$$f(t) = \sum_{|k| \leq n} c_k e^{ikt} \quad (t \in \mathbb{R}).$$

Für ein solches f und $l \in \mathbb{Z}$ gilt dann nach Bemerkung 24.1 (2):

$$\int_{-\pi}^{\pi} f(t) e^{-ilt} dt = \sum_{|k| \leq n} c_k \int_{-\pi}^{\pi} e^{ikt} e^{-ilt} dt = \begin{cases} 2\pi c_l, & |l| \leq n \\ 0, & |l| > n \end{cases}.$$

Bemerkung: Für $k \in \mathbb{N}$ gilt

$$c_k e^{ikt} + c_{-k} e^{-ikt} = (c_k + c_{-k}) \cos(kt) + i(c_k - c_{-k}) \sin(kt) \quad (t \in \mathbb{R}).$$

Man kann also statt mit e^{ikt} für $|k| \leq n$ auch mit $\sin(kt), \cos(kt)$ für $1 \leq k \leq n$ bzw. $0 \leq k \leq n$ arbeiten.

24.3. Fourierkoeffizienten einer Funktion: Sei $f \in R([-\pi, \pi], \mathbb{C})$. Für jedes $k \in \mathbb{Z}$ heißt

$$\widehat{f}(k) := c_k(f) := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ikt} dt$$

der k -te Fourierkoeffizient, und $(\widehat{f}(k))_{k \in \mathbb{Z}}$ heißt Folge der Fourierkoeffizienten von f .

Bemerkung: Man verwendet außerdem gelegentlich für $k \in \mathbb{N}_0$ bzw. $k \in \mathbb{N}$:

$$\begin{aligned} a_k(f) &:= c_k(f) + c_{-k}(f) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(kt) dt \\ b_k(f) &:= i(c_k(f) - c_{-k}(f)) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin(kt) dt. \end{aligned}$$

Ist f reellwertig, so sind $(a_k(f))_{k \in \mathbb{N}_0}, (b_k(f))_{k \in \mathbb{N}}$ reelle Folgen. Man hat dann (vergleiche Bemerkung in 24.2)

$$\sum_{|k| \leq n} c_k(f) e^{ikt} = \frac{a_0(f)}{2} + \sum_{k=1}^n (a_k(f) \cos(kt) + b_k(f) \sin(kt)),$$

und es gilt $c_0(f) = a_0(f)/2$, sowie für $k \in \mathbb{N}$:

$$c_k(f) = \frac{a_k(f) - ib_k(f)}{2}, \quad c_{-k}(f) = \frac{a_k(f) + ib_k(f)}{2}.$$

Beispiele: (1) Wir betrachten die 2π -periodische Funktion $f : \mathbb{R} \rightarrow \mathbb{C}$, die durch $f(t) := t$ für $t \in [-\pi, \pi)$ definiert ist. Es gilt $\widehat{f}(0) = 0$. Für $k \neq 0$ ist

$$\widehat{f}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} t e^{-ikt} dt = \frac{1}{2\pi} \frac{t}{-ik} e^{-ikt} \Big|_{-\pi}^{\pi} - \underbrace{\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{-ik} e^{-ikt} dt}_{=0} = \frac{(-1)^k}{k} i.$$

(2) Sei die 2π -periodische Funktion f auf $[-\pi, \pi)$ gegeben durch $f(t) = t^2$. Für $k = 0$ ist

$$\widehat{f}(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} t^2 dt = \frac{1}{2\pi} \frac{1}{3} t^3 \Big|_{-\pi}^{\pi} = \frac{\pi^2}{3}.$$

Für $k \neq 0$ ist

$$\widehat{f}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} t^2 e^{-ikt} dt = \underbrace{\frac{t^2}{2\pi} \frac{1}{-ik} e^{-ikt} \Big|_{-\pi}^{\pi}}_{=0} + \frac{2}{2\pi ik} \int_{-\pi}^{\pi} t e^{-ikt} dt = \frac{2}{ik} \frac{(-1)^k i}{k} = \frac{2(-1)^k}{k^2}.$$

24.4. Stückweise stetige und stückweise glatte Funktionen: Eine 2π -periodische Funktion $f : \mathbb{R} \rightarrow \mathbb{C}$ heißt

(a) *stückweise stetig*, falls es

$$-\pi = t_0 < t_1 < t_2 < \dots < t_n = \pi$$

so gibt, dass für jedes $j = 0, \dots, n-1$ gilt:

- (i) $f : (t_j, t_{j+1}) \rightarrow \mathbb{C}$ ist stetig,
- (ii) die einseitigen Grenzwerte

$$f(t_j+) = \lim_{t \rightarrow t_j+} f(t) \quad \text{und} \quad f(t_{j+1}-) = \lim_{t \rightarrow t_{j+1}-} f(t)$$

existieren in $\mathbb{C}$.

(b) *stückweise glatt*, falls es

$$-\pi = t_0 < t_1 < t_2 < \dots < t_n = \pi$$

so gibt, dass für jedes $j = 0, \dots, n-1$ gilt:

- (i) $f : (t_j, t_{j+1}) \rightarrow \mathbb{C}$ ist stetig differenzierbar,
- (ii) die einseitigen Grenzwerte $f(t_j+)$, $f(t_{j+1}-)$ und

$$f'(t_j+) = \lim_{t \rightarrow t_j+} f'(t), \quad f'(t_{j+1}-) = \lim_{t \rightarrow t_{j+1}-} f'(t)$$

existieren in $\mathbb{C}$.

In den Punkten t_j muss f nicht stetig sein, der Funktionswert $f(t_j)$ spielt keine Rolle.

Bemerkung: Ist f stetig, so ist f stückweise stetig. Ist f stetig differenzierbar, so ist f stückweise glatt. Ist f stückweise stetig, so existieren in jedem Punkt $t \in \mathbb{R}$ die einseitigen Grenzwerte $f(t+)$ und $f(t-)$.

Beispiele: (1) Die Funktionen in den Beispielen 24.3 (1) und (2) sind stückweise glatt.

(2) Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ 2π -periodisch mit

$$f(t) = \begin{cases} -1, & t \in [-\pi, 0) \\ 1, & t \in [0, \pi) \end{cases}$$

Dann ist f stückweise glatt.

24.5. Darstellungssatz für 2π -periodische Funktionen: Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ 2π -periodisch und stückweise glatt. Dann gilt für jedes $t \in \mathbb{R}$:

$$\sum_{k=-\infty}^{\infty} \widehat{f}(k) e^{ikt} := \lim_{n \rightarrow \infty} \sum_{k=-n}^n \widehat{f}(k) e^{ikt} = \frac{f(t+) + f(t-)}{2}.$$

Ist f stetig in t , so konvergiert die Fourierreihe also gegen $f(t)$.

Beweis. Es sei

$$s_n(f, t) := \sum_{k=-n}^n \widehat{f}(k) e^{ikt}$$

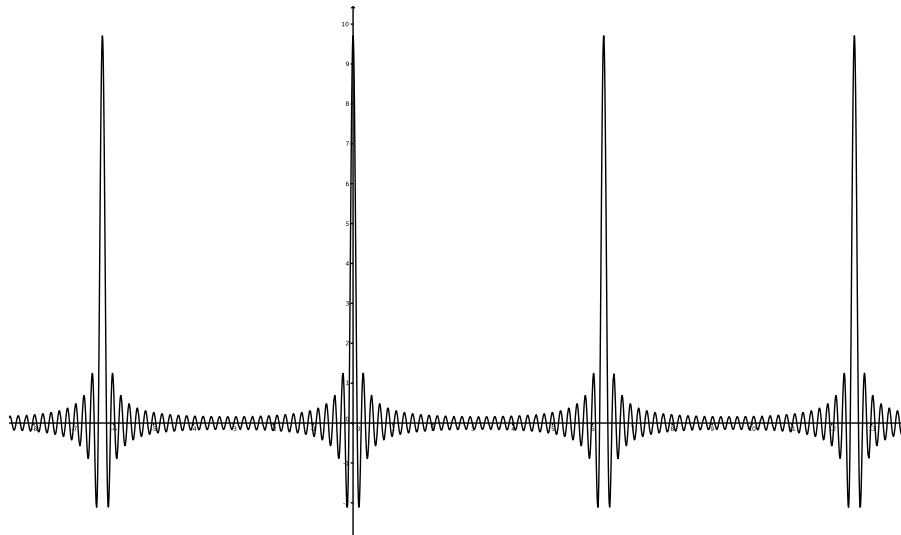
Die 2π -periodische Funktion $D_n : \mathbb{R} \rightarrow \mathbb{C}$

$$D_n(t) := \frac{1}{2\pi} \sum_{k=-n}^n e^{ikt}$$

heißt *Dirichlet-Kern* und besitzt u.a. folgende Eigenschaften:

$$D_n(-t) = D_n(t), \quad \int_0^\pi D_n(t) dt = \frac{1}{2}, \quad D_n(t) = \begin{cases} \frac{1}{2\pi} \frac{\sin((n+1/2)t)}{\sin(t/2)}, & t \notin 2\pi\mathbb{Z} \\ \frac{n+1/2}{\pi}, & t \in 2\pi\mathbb{Z} \end{cases}.$$

Folgendes Bild zeigt den Graph von D_{30} :



Für $t \notin 2\pi\mathbb{Z}$ ergibt sich $D_n(t)$ aus der geometrischen Summenformel:

$$\sum_{k=-n}^n e^{ikt} = e^{-int} \sum_{k=0}^{2n} (e^{it})^k = e^{-int} \frac{1 - e^{i(2n+1)t}}{1 - e^{it}}$$

$$= \frac{e^{i(n+1/2)t} - e^{-i(n+1/2)t}}{e^{it/2} - e^{-it/2}} = \frac{\sin((n+1/2)t)}{\sin(t/2)}$$

Mithilfe der Dirichlet-Kerne können die Teilsummen $s_n(f, t)$ wie folgt dargestellt werden:

$$\begin{aligned} s_n(f, t) &= \int_{-\pi}^{\pi} \left(f(s) \frac{1}{2\pi} \sum_{k=-n}^n e^{ik(t-s)} \right) ds = \int_{-\pi}^{\pi} f(s) D_n(t-s) ds = \int_{-\pi}^{\pi} f(t-s) D_n(s) ds \\ &= \int_0^{\pi} f(t+s) D_n(s) ds + \int_0^{\pi} f(t-s) D_n(s) ds. \end{aligned}$$

Für $\delta > 0$ hinreichend klein liefert die Voraussetzung “stückweise glatt” an die Funktion f dabei (für hinreichend kleines $\delta \in (0, 2\pi)$) eine Konstante L mit $|f(t \pm s) - f(t \pm)| \leq Ls$ ($s \in (0, \delta)$), also

$$|(f(t \pm s) - f(t \pm)) D_n(s)| \leq \frac{L}{\pi} \left| \frac{s/2}{\sin(s/2)} \right| \leq \frac{L}{\pi} \frac{\delta/2}{\sin(\delta/2)} \quad (s \in (0, \delta)),$$

und somit

$$\int_0^{\delta} |(f(t \pm s) - f(t \pm)) D_n(s)| ds \leq \frac{\delta L}{\pi} \frac{\delta/2}{\sin(\delta/2)}.$$

Beachte dabei:

$$s \mapsto \frac{s/2}{\sin(s/2)}$$

ist ≥ 0 und monoton wachsend auf $[0, 2\pi)$.

Weiter erhält man durch partielle Integration

$$\int_{\delta}^{\pi} (f(t \pm s) - f(t \pm)) D_n(s) ds \rightarrow 0 \quad (n \rightarrow \infty).$$

Damit folgt für $n \rightarrow \infty$:

$$\begin{aligned} \int_0^{\pi} f(t+s) D_n(s) ds - \frac{f(t+)}{2} &= \int_0^{\pi} (f(t+s) - f(t+)) D_n(s) ds \rightarrow 0, \\ \int_0^{\pi} f(t-s) D_n(s) ds - \frac{f(t-)}{2} &= \int_0^{\pi} (f(t-s) - f(t-)) D_n(s) ds \rightarrow 0. \end{aligned}$$

Wähle dazu zu gegebenem $\varepsilon > 0$ erst δ hinreichend klein und dann n hinreichend groß. $\square$

Beispiele: (1) Für die 2π -periodische Funktion f mit $f(t) = t$ ($t \in [-\pi, \pi)$) gilt nach Beispiel 24.3 (1)

$$\widehat{f}(0) = 0, \quad \widehat{f}(k) = \frac{(-1)^k}{k} i \quad (k \in \mathbb{Z} \setminus \{0\}).$$

Also ist

$$\sum_{|k| \leq n} \widehat{f}(k) e^{ikt} = \sum_{0 < |k| \leq n} \frac{(-1)^k i}{k} (\cos(kt) + i \sin(kt)) = \sum_{k=1}^n \frac{2(-1)^{k+1}}{k} \sin(kt),$$

also

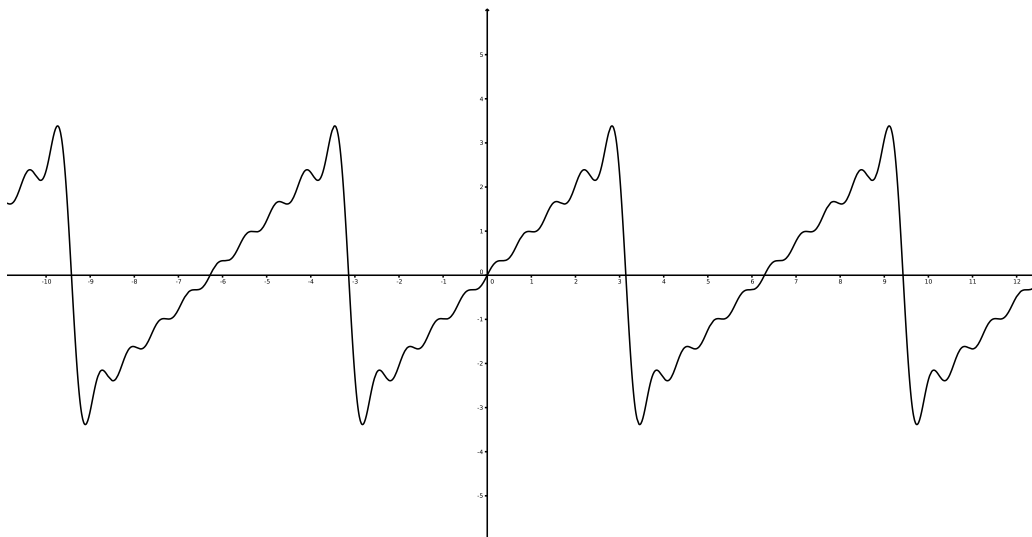
$$t = 2 \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \sin(kt) \quad (t \in (-\pi, \pi)),$$

da f in jedem $t \in (-\pi, \pi)$ stetig ist.

An der Stelle $t = \pi$ ist f unstetig. Setzt man $t = \pi$ in die Reihe rechts ein, so erhält man den Wert 0. In der Tat ist $f(\pi-) = \pi$, $f(\pi+) = f(-\pi+) = -\pi$, also

$$\frac{f(\pi+) + f(\pi-)}{2} = 0.$$

Folgendes Bild zeigt den Graph der Teilsumme $t \mapsto 2 \sum_{k=1}^9 \frac{(-1)^{k+1}}{k} \sin(kt)$.



Bemerkung: Die Fourierreihe in diesem Beispiel ist auf $\mathbb{R}$ nicht gleichmäßig konvergent. Das Überspringen der Teilsummen von Fourierreihen in der Nähe von Sprungstellen der Funktion f (vgl. obiges Bild) heißt *Gibbssches Phänomen*.

(2) Für die 2π -periodische Funktion f mit $f(t) = t^2$ ($t \in [-\pi, \pi)$) gilt nach Beispiel 24.3 (2):

$$\widehat{f}(0) = \frac{\pi^2}{3}, \quad \widehat{f}(k) = \frac{2(-1)^k}{k^2} \quad (k \in \mathbb{Z} \setminus \{0\}).$$

Da f stetig und stückweise glatt ist, gilt nach Satz 24.5:

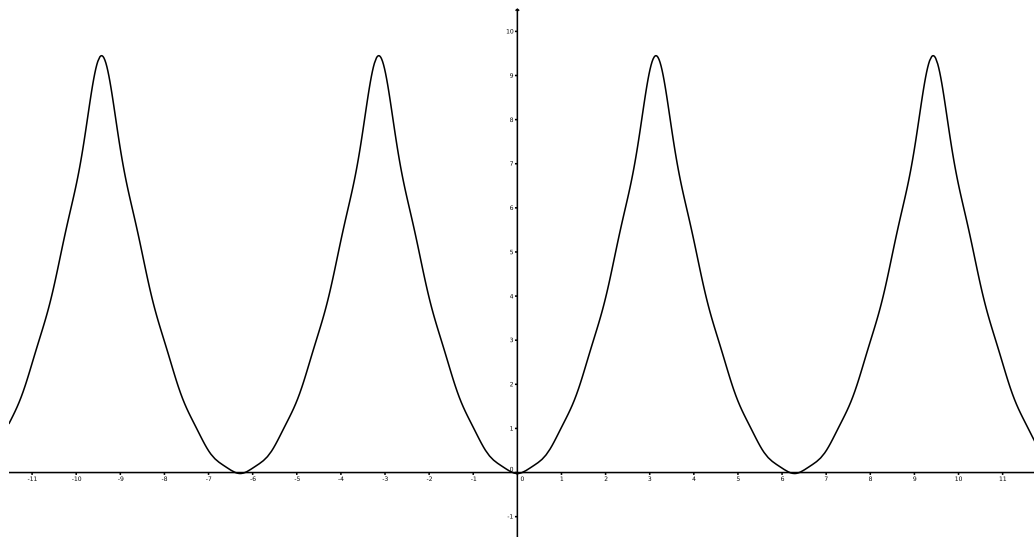
$$t^2 = \frac{\pi^2}{3} + 2 \sum_{k \in \mathbb{Z} \setminus \{0\}} \frac{(-1)^k}{k^2} e^{ikt} = \frac{\pi^2}{3} + 4 \sum_{k=1}^{\infty} \frac{(-1)^k}{k^2} \cos(kt) \quad (t \in [-\pi, \pi]).$$

Die Fourierreihe in diesem Beispiel ist auf $\mathbb{R}$ gleichmäßig konvergent.

Speziell für $t = \pi$ erhält man:

$$\frac{\pi^2}{6} = \sum_{k=1}^{\infty} \frac{1}{k^2}.$$

Folgendes Bild zeigt den Graph der Teilsumme $t \mapsto \frac{\pi^2}{3} + 4 \sum_{k=1}^9 \frac{(-1)^k}{k^2} \cos(kt)$.



Bemerkung: Der Darstellungssatz legt es nahe, für stückweise stetige (und damit insbesondere für stückweise glatte Funktionen) die folgende *Normalisierung* zu betrachten:

Ist $f : \mathbb{R} \rightarrow \mathbb{C}$ 2π -periodisch und stückweise stetig mit $-\pi = t_0 < t_1 < \dots < t_n = \pi$ wie in 24.4 (a), so heißt f *normalisiert*, wenn

$$f(t_j) = \frac{f(t_j+) + f(t_j-)}{2} \quad (j = 1, \dots, n)$$

gilt. Dann ist

$$f(t) = \frac{f(t+) + f(t-)}{2} \quad \text{für jedes } t \in \mathbb{R}.$$

Durch eventuelles Abändern der Funktionswerte $f(t_j)$ lässt sich jede stückweise stetige Funktion normalisieren. Die Fourierkoeffizienten ändern sich dabei nicht.

Ist also $f : \mathbb{R} \rightarrow \mathbb{C}$ 2π -periodisch, stückweise glatt und normalisiert, so gilt für jedes $t \in \mathbb{R}$:

$$\sum_{k=-\infty}^{\infty} \hat{f}(k) e^{ikt} = f(t).$$

Beispiel: Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ die 2π -periodische Funktion mit

$$f(t) = \begin{cases} 0, & t = -\pi \\ -1, & t \in (-\pi, 0) \\ 0, & t = 0 \\ 1, & t \in (0, \pi) \end{cases}$$

Dann ist f stückweise glatt und normalisiert. Es gilt $\widehat{f}(0) = 0$ und für $k \neq 0$ ist

$$\widehat{f}(k) = -\frac{1}{2\pi} \int_{-\pi}^0 e^{-ikt} dt + \frac{1}{2\pi} \int_0^{\pi} e^{-ikt} dt = \frac{i}{k\pi}((-1)^k - 1),$$

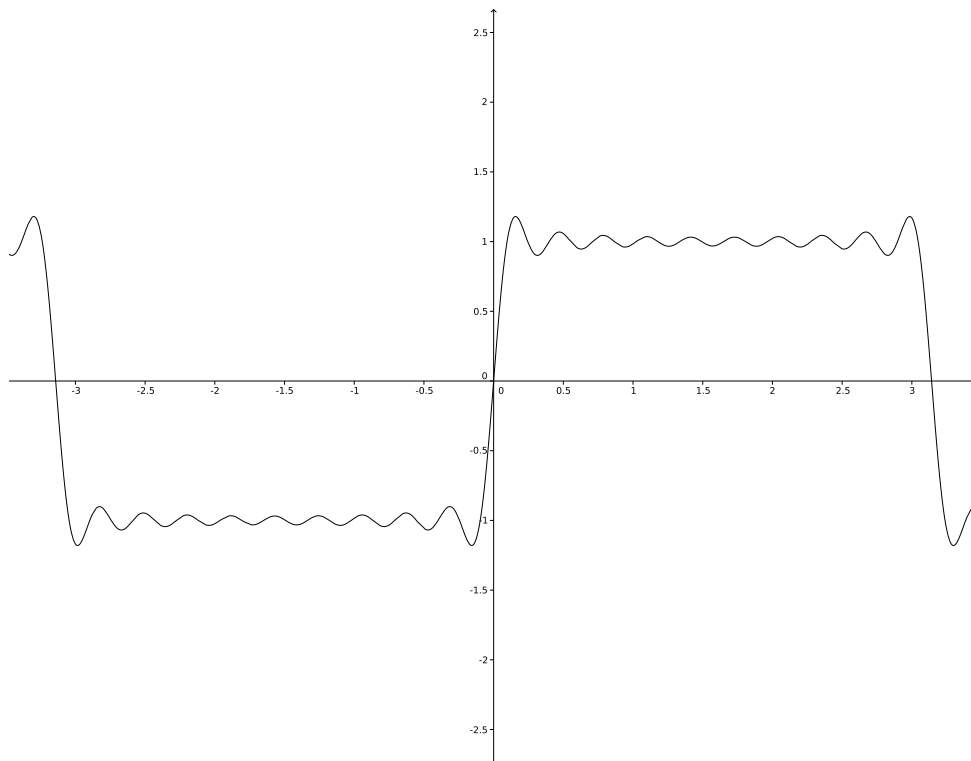
also

$$\forall k \in \mathbb{Z} : \quad \widehat{f}(2k) = 0, \quad \widehat{f}(2k+1) = -\frac{2i}{(2k+1)\pi}.$$

Nach Satz 24.5 gilt:

$$\forall t \in \mathbb{R} : \quad f(t) = \sum_{k=-\infty}^{\infty} \frac{-2i}{(2k+1)\pi} e^{i(2k+1)t} = \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{1}{2k+1} \sin((2k+1)t).$$

Folgendes Bild zeigt den Graph der Teilsumme $t \mapsto \frac{4}{\pi} \sum_{k=0}^9 \frac{1}{2k+1} \sin((2k+1)t)$.



24.6. Konvergenz im quadratischen Mittel: Es gibt stetige 2π -periodische Funktionen $f : \mathbb{R} \rightarrow \mathbb{C}$ deren Fourierreihe

$$\sum_{k=-\infty}^{\infty} \widehat{f}(k) e^{ikt}$$

nicht für jedes $t \in \mathbb{R}$ konvergiert. Insbesondere konvergiert die Fourierreihe dann nicht punktweise gegen f . Wir werden nun sehen, dass die Fourierreihe stetiger 2π -periodischer Funktionen $f : \mathbb{R} \rightarrow \mathbb{C}$ stets im quadratischen Mittel gegen die Funktion f konvergiert.

Durch

$$(f|g) := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \overline{g(t)} dt, \quad \|f\| := \sqrt{(f|f)} = \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t)|^2 dt \right)^{1/2}$$

wird auf $C([-\pi, \pi], \mathbb{C})$ ein Skalarprodukt und eine Norm definiert, also auch auf

$$C_{\text{per}}([-\pi, \pi], \mathbb{C}) := \{f \in C([-\pi, \pi], \mathbb{C}) : f(\pi) = f(-\pi)\}.$$

Wir können diesen Raum mit

$$\{f \in C(\mathbb{R}, \mathbb{C}) : f \text{ ist } 2\pi\text{-periodisch}\}$$

identifizieren (vergleiche mit 24.1 Bemerkung (a)).

Bemerkung: Konvergenz im quadratischen Mittel, also Konvergenz bzgl. obiger Norm $\|\cdot\|$, hat nicht punktweise Konvergenz zur Folge. Betrachte z.B. die Folge $(f_n)_{n \in \mathbb{N}}$ in $C_{\text{per}}([-\pi, \pi], \mathbb{C})$ mit $f_n(t) = (\cos(t))^n$. Dann gilt $f_n \rightarrow 0$ ($n \rightarrow \infty$) im quadratischen Mittel, denn (Übung)

$$\|f_n - 0\|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |\cos(t)|^{2n} dt \rightarrow 0 \quad (n \rightarrow \infty),$$

aber f_n ist nicht punktweise konvergent gegen die Nullfunktion:

Die Folge $(f_n(0))_{n \in \mathbb{N}} = (1)_{n \in \mathbb{N}}$ konvergiert gegen 1 und die Folgen

$$(f_n(\pi))_{n \in \mathbb{N}} = (f_n(-\pi))_{n \in \mathbb{N}} = ((-1)^n)_{n \in \mathbb{N}}$$

sind divergent. Es gibt Folgen in $C_{\text{per}}([-\pi, \pi], \mathbb{C})$ die im quadratischen Mittel aber in keinem Punkt $t \in [-\pi, \pi]$ konvergieren.

Bezeichnen wir für $k \in \mathbb{Z}$ die Funktion $t \mapsto e^{ikt}$ mit e_k , so bedeutet Bemerkung 24.1 (2), dass $(e_k)_{k \in \mathbb{Z}}$ ein *Orthonormalsystem* in $C_{\text{per}}([-\pi, \pi], \mathbb{C})$ ist. In folgendem allgemeinen Satz betrachten wir ein beliebiges Orthonormalsystem $(e_k)_{k \in \mathbb{N}}$. Die Aussagen gelten analog für ein beliebiges Orthonormalsystem $(e_k)_{k \in \mathbb{Z}}$.

Satz 1: Sei V ein $\mathbb{K}$ -Vektorraum mit Skalarprodukt und $\dim V = \infty$. Sei $(e_k)_{k \in \mathbb{N}}$ ein Orthonormalsystem in V , d.h. $(e_k | e_l) = \delta_{kl}$ ($k, l \in \mathbb{N}$). Dann gilt:

$$\forall v \in V : \sum_{k=1}^{\infty} |(v | e_k)|^2 \leq \|v\|^2 \quad (\text{Besselsche Ungleichung}),$$

Weiter sind äquivalent:

- (i) $\sum_{k=1}^{\infty} |(v | e_k)|^2 = \|v\|^2$ (Parsevalsche Gleichung),
- (ii) $\|v - \sum_{k=1}^n (v | e_k) e_k\| \rightarrow 0$ ($n \rightarrow \infty$) (d.h. $v = \sum_{k=1}^{\infty} (v | e_k) e_k$).

Gilt (i) (oder (ii)) für jedes $v \in V$, so heißt das Orthonormalsystem $(e_k)_{k \in \mathbb{N}}$ *vollständig* in V .

Beweis. Für jedes $v \in V$ gilt:

$$\begin{aligned} 0 \leq \|v - \sum_{k=1}^n (v | e_k) e_k\|^2 &= \|v\|^2 + \sum_{k=1}^n |(v | e_k)|^2 - 2 \operatorname{Re} (v | \sum_{k=1}^n (v | e_k) e_k) \\ &= \|v\|^2 + \sum_{k=1}^n |(v | e_k)|^2 - 2 \operatorname{Re} \sum_{k=1}^n (v | e_k) \overline{(v | e_k)} \\ &= \|v\|^2 - \sum_{k=1}^n |(v | e_k)|^2. \end{aligned}$$

Daraus folgt die Besselsche Ungleichung, und die Äquivalenz von (i) und (ii). □

Bemerkung (Ohne Beweis): Ist $(e_k)_{k \in \mathbb{N}}$ vollständig in V so gilt auch

$$(v | w) = \sum_{k=1}^{\infty} (v | e_k) \overline{(w | e_k)} \quad (v, w \in V).$$

Satz 2: Für jedes $f \in C_{\text{per}}([-\pi, \pi], \mathbb{C})$ gilt:

$$\sum_{k \in \mathbb{Z}} |\hat{f}(k)|^2 = \|f\|^2 \quad \text{und} \quad \|f - \sum_{|k| \leq n} \hat{f}(k) e_k\| \rightarrow 0 \quad (n \rightarrow \infty).$$

Insbesondere ist $(e_k)_{k \in \mathbb{Z}}$ ein vollständiges Orthonormalsystem in $C_{\text{per}}([-\pi, \pi], \mathbb{C})$.

Beispiel: Für die 2π -periodische Funktion f mit $f(t) = t^2$ ($t \in [-\pi, \pi]$) gilt nach Beispiel 24.3 (2):

$$\hat{f}(0) = \frac{\pi^2}{3}, \quad \hat{f}(k) = \frac{2(-1)^k}{k^2} \quad (k \in \mathbb{Z} \setminus \{0\}).$$

Es gilt also

$$\sum_{k \in \mathbb{Z}} |\widehat{f}(k)|^2 = \frac{\pi^4}{9} + \sum_{k \in \mathbb{Z} \setminus \{0\}} \frac{4}{k^4} = \frac{\pi^4}{9} + \sum_{k=1}^{\infty} \frac{8}{k^4}$$

und

$$\|f\|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} t^4 dt = \frac{1}{2\pi} \frac{2\pi^5}{5} = \frac{\pi^4}{5}.$$

Somit ist

$$\sum_{k=1}^{\infty} \frac{1}{k^4} = \frac{1}{8} \left(\frac{\pi^4}{5} - \frac{\pi^4}{9} \right) = \frac{\pi^4}{90}.$$

Beweis. 1) Sei zunächst $f \in C^1(\mathbb{R}, \mathbb{C})$ und 2π -periodisch. Dann gilt $f, f' \in C_{\text{per}}([-\pi, \pi], \mathbb{C})$ und mit partieller Integration folgt:

$$\widehat{f}(k) = \frac{1}{ik} \widehat{(f')}(k) \quad (k \in \mathbb{Z} \setminus \{0\}).$$

Insbesondere folgt mithilfe der Cauchy-Schwarz-Ungleichung

$$\sum_{k \in \mathbb{Z} \setminus \{0\}} |\widehat{f}(k)| \leq \left(\sum_{k \in \mathbb{Z} \setminus \{0\}} \frac{1}{k^2} \right)^{1/2} \left(\sum_{k \in \mathbb{Z} \setminus \{0\}} |\widehat{(f')}(k)|^2 \right)^{1/2} < \infty.$$

Somit konvergiert die Reihe $\sum_{k=-\infty}^{\infty} \widehat{f}(k) e_k$ auf $[-\pi, \pi]$ gleichmäßig gegen ein $g \in C_{\text{per}}([-\pi, \pi], \mathbb{C})$. Da f stetig differenzierbar ist (also insbesondere stetig und stückweise glatt), konvergiert diese Reihe nach 24.5 aber punktweise gegen f . Also ist $g = f$ und somit gilt

$$\|f - \sum_{|k| \leq n} \widehat{f}(k) e_k\| \leq \|f - \sum_{|k| \leq n} \widehat{f}(k) e_k\|_{\infty} \rightarrow 0 \quad (n \rightarrow \infty).$$

2) Ist nun $f \in C_{\text{per}}([-\pi, \pi], \mathbb{C})$ beliebig, so kann man Teil 1) anwenden auf f_{δ} , gegeben durch

$$f_{\delta}(t) := \frac{1}{\delta} \int_t^{t+\delta} f(s) ds \quad (t \in [-\pi, \pi]),$$

wobei $\delta > 0$. Für jedes $g \in C_{\text{per}}([-\pi, \pi], \mathbb{C})$ gilt

$$\|g\| = \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |g(t)|^2 dt \right)^{1/2} \leq \|g\|_{\infty} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} 1 dx \right)^{1/2} = \|g\|_{\infty},$$

also $\|f - f_{\delta}\| \leq \|f - f_{\delta}\|_{\infty}$. Weiter gilt $\|f - f_{\delta}\|_{\infty} \rightarrow 0$ für $\delta \rightarrow 0+$: Zu $\varepsilon > 0$ findet man wegen der gleichmäßigen Stetigkeit von f auf $[-2\pi, 2\pi]$ ein $\delta_0 \in (0, 1)$ mit $|f(t) - f(s)| \leq \varepsilon$ für alle $t, s \in [-2\pi, 2\pi]$ mit $|t - s| \leq \delta_0$. Für $\delta \in (0, \delta_0)$ ist dann

$$\begin{aligned} \|f - f_{\delta}\|_{\infty} &= \sup_{t \in [-\pi, \pi]} |f(t) - f_{\delta}(t)| = \sup_{t \in [-\pi, \pi]} \left| \frac{1}{\delta} \int_t^{t+\delta} (f(t) - f(s)) ds \right| \\ &\leq \sup_{t \in [-\pi, \pi]} \sup_{|t-s| \leq \delta} |f(t) - f(s)| \leq \varepsilon. \end{aligned}$$

Wir zeigen nun die Konvergenzaussage aus dem Satz für f . Dazu sei $\varepsilon > 0$ gegeben. Wir wählen $\delta > 0$ mit $\|f - f_\delta\| < \varepsilon/3$. Dann gilt

$$\begin{aligned} \|f - \sum_{|k| \leq n} \widehat{f}(k) e_k\| &\leq \|f - f_\delta\| + \|f_\delta - \sum_{|k| \leq n} \widehat{f}_\delta(k) e_k\| + \|\sum_{|k| \leq n} (\widehat{f} - \widehat{f}_\delta)(k) e_k\| \\ &\leq 2\|f - f_\delta\| + \|f_\delta - \sum_{|k| \leq n} \widehat{f}_\delta(k) e_k\| \leq 2\varepsilon/3 + \|f_\delta - \sum_{|k| \leq n} \widehat{f}_\delta(k) e_k\| \leq \varepsilon \end{aligned}$$

für $n \geq n_0$ und geeignetes n_0 (existiert nach Teil 1)). Damit ist der Satz gezeigt. $\square$

24.7. Hilberträume: Ein $\mathbb{K}$ -Vektorraum V mit Skalarprodukt heißt *Hilbertraum*, falls er bezüglich der zum Skalarprodukt gehörigen Norm $\|\cdot\|$ *vollständig* ist, d.h. also wenn jede Cauchyfolge $(v_n)_{n \in \mathbb{N}}$ in V einen Grenzwert hat.

Dabei heißt eine Folge $(v_n)_{n \in \mathbb{N}}$ *Cauchyfolge*, falls gilt (vgl. HM I):

$$\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n, m \geq n_0 : \|v_n - v_m\| \leq \varepsilon.$$

Beispiele: 1) Für jedes $n \in \mathbb{N}$ ist $\mathbb{K}^n$ ein Hilbertraum.

2) $l_{\mathbb{K}}^2(\mathbb{N})$ ist ein Hilbertraum. Sei $(x^{(n)})_{n \in \mathbb{N}}$ eine Cauchyfolge in $l_{\mathbb{K}}^2(\mathbb{N})$, wobei $x^{(n)} = (x_k^{(n)})_{k \in \mathbb{N}}$. Wegen

$$|x_k^{(n)} - x_k^{(m)}| \leq \|x^{(n)} - x^{(m)}\|$$

ist $(x_k^{(n)})_{n \in \mathbb{N}}$ für jedes k eine Cauchyfolge in $\mathbb{K}$, die gegen ein $x_k^{(0)} \in \mathbb{K}$ konvergiert.

Sei nun $\varepsilon > 0$ und n_0 gewählt mit $\|x^{(n)} - x^{(m)}\| \leq \varepsilon$ für alle $n, m \geq n_0$. Sei $n \geq n_0$. Dann gilt für jedes $N \in \mathbb{N}$:

$$\sum_{k=1}^N |x_k^{(n)} - x_k^{(0)}|^2 = \lim_{m \rightarrow \infty} \underbrace{\sum_{k=1}^N |x_k^{(n)} - x_k^{(m)}|^2}_{\leq \|x^{(n)} - x^{(m)}\|^2} \leq \varepsilon^2,$$

für $N \rightarrow \infty$ also $\|x^{(n)} - x^{(0)}\| \leq \varepsilon$. Daraus folgt $x^{(0)} \in l_{\mathbb{K}}^2(\mathbb{N})$ (mit der umgekehrten Dreiecksungleichung) und dann $\|x^{(n)} - x^{(0)}\| \rightarrow 0$ für $n \rightarrow \infty$.

24.8. Bemerkungen: (1) Setzt man $l_{\mathbb{C}}^2(\mathbb{Z}) := \{(c_k)_{k \in \mathbb{Z}} \in \mathbb{C}^{\mathbb{Z}} : \sum_{k=-\infty}^{\infty} |c_k|^2 < \infty\}$, so erhält die lineare Abbildung

$$C_{\text{per}}([-\pi, \pi], \mathbb{C}) \rightarrow l_{\mathbb{C}}^2(\mathbb{Z}), \quad f \mapsto (\widehat{f}(k))_{k \in \mathbb{Z}},$$

Norm und Skalarprodukt, ist also insbesondere injektiv. Sie ist aber nicht surjektiv! Betrachtet man den größeren Raum $L^2([-\pi, \pi], \mathbb{C})$ ($\rightarrow$ Lebesgue-Integral), so wird diese Abbildung surjektiv. $L^2([-\pi, \pi], \mathbb{C})$ ist ein Hilbertraum, nicht aber $C_{\text{per}}([-\pi, \pi], \mathbb{C})$.

(2) Neben der Konvergenz im quadratischen Mittel gibt es für stetige periodische Funktionen noch einen Satz von Fejér über Konvergenz im arithmetischen Mittel:

Satz: Es sei $f \in C_{\text{per}}([-\pi, \pi], \mathbb{C})$, und für $n, N \in \mathbb{N}_0$ sei

$$s_n(f, t) := \sum_{k=-n}^n \widehat{f}(k) e^{ikt}, \quad \sigma_N(f, t) := \frac{1}{N} \sum_{n=0}^{N-1} s_n(f, t).$$

Dann konvergiert $(\sigma_N(f, \cdot))$ gleichmäßig auf $\mathbb{R}$ gegen f .

Der Beweis geht ähnlich wie der Beweis des Darstellungssatzes in 24.5. Der wesentliche Unterschied ist, dass der Dirichlet-Kern $D_n(t)$ durch den *Fejér-Kern*

$$F_N(t) := \frac{1}{N} \sum_{n=0}^{N-1} D_n(t) = \begin{cases} \frac{1}{2\pi N} \left(\frac{\sin(Nt/2)}{\sin(t/2)} \right)^2, & t \notin 2\pi\mathbb{Z} \\ \frac{N}{2\pi}, & t \in 2\pi\mathbb{Z} \end{cases}$$

ersetzt wird. Im Unterschied zu den Dirichlet-Kernen sind die Fejér-Kerne stets ≥ 0 .

24.9. Anwendung: Wir betrachten einen dünnen Metallstab der Länge π und suchen eine Funktion $u : [0, \pi] \times [0, \infty) \rightarrow \mathbb{R}$, $(x, t) \mapsto u(x, t)$, welche die Temperaturverteilung beschreibt. Für u gilt dann die *Wärmeleitungsgleichung*

$$u_t(x, t) = cu_{xx}(x, t) \quad ((x, t) \in (0, \pi) \times (0, \infty) =: S)$$

mit einer Konstanten $c > 0$ (der Temperaturleitfähigkeit des Materials). Wir nehmen $c = 1$ an. Ferner nehmen wir an, dass zur Zeit $t = 0$ gilt $u(x, 0) = f(x)$ ($x \in [0, \pi]$) für eine Funktion $f \in C^1([0, \pi], \mathbb{R})$ mit $f(0) = f(\pi) = 0$, und dass der Stab an den Enden gekühlt wird (d.h. auf Temperatur 0 gehalten wird). Dies führt auf das *Anfangs-Randwertproblem*

$$(ARP) \quad \begin{cases} u_t(x, t) = u_{xx}(x, t) & ((x, t) \in S), \\ u(x, 0) = f(x) & (x \in [0, \pi]), \\ u(0, t) = u(\pi, t) = 0 & (t \geq 0). \end{cases}$$

Durch einen Separationsansatz (vgl. HM III) findet man Lösungen von $u_t(x, t) = u_{xx}(x, t)$ der Form

$$u_k(x, t) = e^{-k^2 t} \sin(kx) \quad (k \in \mathbb{N}).$$

Linearkombination solcher Lösungen sind wieder Lösungen der Wärmeleitungsgleichung. Dies führt auf die Idee die Lösung von (ARP) als Fourierreihe darzustellen:

Wir setzen f auf $[-\pi, \pi]$ ungerade fort durch $f(x) := -f(-x)$ ($x \in [-\pi, 0)$), und dann 2π -periodisch fort auf $\mathbb{R}$. Damit ist $f \in C^1(\mathbb{R}, \mathbb{R})$. Die reellen Fourierkoeffizienten (vgl. 24.3) sind

$$a_k(f) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(kx) dx = 0 \quad (k \in \mathbb{N}_0),$$

$$b_k(f) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(kx) dx \quad (k \in \mathbb{N}),$$

und wir erhalten aus dem Darstellungssatz 24.5

$$f(x) = \sum_{k=1}^{\infty} b_k(f) \sin(kx) \quad (x \in \mathbb{R}).$$

Setzt man nun

$$u(x, t) := \sum_{k=1}^{\infty} b_k(f) e^{-k^2 t} \sin(kx) \quad ((x, t) \in [0, \pi] \times [0, \infty)),$$

so ist u die Lösung von (ARP) (Eindeutigkeit in HM III). Wegen $f \in C^1(\mathbb{R}, \mathbb{R})$ gilt

$$\sum_{k=1}^{\infty} |b_k(f)| < \infty.$$

Die Reihe für f konvergiert also absolut und gleichmäßig auf $\mathbb{R}$ und u ist stetig auf $[0, \pi] \times [0, \infty)$. Gliedweises Ableiten der Reihe für u ist in $[0, \pi] \times (0, \infty)$ möglich nach Sätzen aus HM I. Weiter gilt für alle $x \in [0, \pi]$:

$$|u(x, t)| \leq e^{-t} \sum_{k=1}^{\infty} |b_k(f)| \rightarrow 0 \quad (t \rightarrow \infty)$$

d.h. der Stab kühlt asymptotisch gleichmäßig aus.

25 Fouriertransformation

25.1. Motivation: In vorigen Kapitel haben wir 2π -periodische Funktionen $f : \mathbb{R} \rightarrow \mathbb{C}$, die etwa stetig differenzierbar sind, in Fourierreihen $\sum_{k \in \mathbb{Z}} \hat{f}(k) e^{ik(\cdot)}$ entwickelt. Dabei kann man für $k \in \mathbb{Z}$ den Koeffizienten

$$\hat{f}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} dx$$

als Anteil der Frequenz k (Oberton der Grundfrequenz 1) im Gesamtsignal verstehen. Für $2T$ -periodische Funktionen lassen sich die Ergebnisse aus Kapitel 16 umrechnen.

Sei nun $a > 0$ und sei $f \in C^1(\mathbb{R}, \mathbb{C})$ mit $f(x) = 0$ ($x \notin [-a, a]$) (Impulssignal ohne jegliche Periodizität). Was kann man für eine solche Funktion machen?

Für $T > a$ definieren wir die $2T$ -periodische Funktion $f_T : \mathbb{R} \rightarrow \mathbb{C}$ durch $f_T := f$ auf $[-T, T)$ und wollen $T \rightarrow \infty$ betrachten. Die durch $g(x) := f_T(\frac{T}{\pi}x)$ definierte Funktion $g : \mathbb{R} \rightarrow \mathbb{C}$ ist 2π -periodisch und $g \in C^1(\mathbb{R}, \mathbb{C})$. Wir entwickeln g in eine Fourierreihe

mittels der Koeffizienten

$$\begin{aligned}
 c_k := \widehat{g}(k) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} g(x) e^{-ikx} dx \\
 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f_T\left(\frac{T}{\pi}x\right) e^{-ikx} dx \\
 &= \frac{1}{2T} \int_{-T}^T f(y) e^{-i\frac{k\pi}{T}y} dy \\
 &= \frac{1}{2T} \int_{-a}^a f(y) e^{-i\frac{k\pi}{T}y} dy \quad (k \in \mathbb{Z}),
 \end{aligned}$$

wobei wir $x = \frac{\pi}{T}y$ substituiert haben. Wir verzichten auf den Normierungsfaktor $\frac{1}{2T}$ und definieren

$$\widehat{f}(\xi) := \int_{-a}^a f(x) e^{-i\xi x} dx = \int_{-\infty}^{\infty} f(x) e^{-i\xi x} dx \quad (\xi \in \mathbb{R}).$$

Es ist dann

$$c_k = \frac{1}{2T} \widehat{f}\left(\frac{k\pi}{T}\right) \quad (k \in \mathbb{Z}).$$

Die Fouriersumme $\sum_{|k| \leq n} c_k e^{ikx}$ mit Limes $g(x) = f_T\left(\frac{T}{\pi}x\right)$ ($x \in \mathbb{R}$), erhält somit (wieder über $x = \frac{\pi}{T}y$) die Form

$$\sum_{|k| \leq n} \widehat{f}\left(\frac{k\pi}{T}\right) e^{i\frac{k\pi}{T}y} \frac{1}{2T} = \frac{1}{2\pi} \sum_{|k| \leq n} \widehat{f}\left(\frac{k\pi}{T}\right) e^{i\frac{k\pi}{T}y} \frac{\pi}{T}$$

mit Grenzwert

$$g\left(\frac{\pi}{T}y\right) = f_T(y) = f(y)$$

für $n \rightarrow \infty$ und $|y| \leq T$. Die Summe ist dabei eine Riemannsumme für das Integral

$$\frac{1}{2\pi} \int_{-n\pi/T}^{n\pi/T} \widehat{f}(\xi) e^{i\xi y} d\xi.$$

Man wird also hoffen, dass (oft)

$$f(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{f}(\xi) e^{i\xi y} d\xi \quad (y \in \mathbb{R}).$$

25.2. Absolut integrierbare Funktionen: Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ eine Funktion, die für jedes $R > 0$ über $[-R, R]$ integrierbar ist. Wir nennen die Funktion f *absolut integrierbar* (aib), falls das uneigentliche Integral

$$\int_{-\infty}^{\infty} f(x) dx$$

absolut konvergiert. In diesem Fall ist

$$\|f\|_1 := \int_{-\infty}^{\infty} |f(x)| dx = \sup_{R>0} \int_{-R}^R |f(x)| dx = \lim_{R \rightarrow \infty} \int_{-R}^R |f(x)| dx < \infty.$$

und es gilt

$$\int_{-\infty}^{\infty} f(x) dx = \lim_{R \rightarrow \infty} \int_{-R}^R f(x) dx \quad \text{und} \quad \left| \int_{-\infty}^{\infty} f(x) dx \right| \leq \|f\|_1.$$

Die Menge der absolut integrierbaren Funktionen bezeichnen wir mit $\mathcal{A}(\mathbb{R}, \mathbb{C})$.

Bemerkung: $\mathcal{A}(\mathbb{R}, \mathbb{C})$ ist ein komplexer Vektorraum, und für $\alpha \in \mathbb{C}$ und $f, g \in \mathcal{A}(\mathbb{R}, \mathbb{C})$ gilt

$$\|\alpha f\|_1 = |\alpha| \|f\|_1, \quad \|f + g\|_1 \leq \|f\|_1 + \|g\|_1.$$

Allerdings ist $\|\cdot\|_1$ keine Norm, denn aus $\|f\|_1 = 0$ folgt i.a. nicht $f = 0$.

Beispiele: (1) Ist $f : \mathbb{R} \rightarrow \mathbb{C}$ stetig und $\{x \in \mathbb{R} : f(x) \neq 0\}$ beschränkt, so ist $f \in \mathcal{A}(\mathbb{R}, \mathbb{C})$.

(2) Die durch $f(x) := (1 + |x|)^{-\alpha}$ definierte Funktion $f : \mathbb{R} \rightarrow \mathbb{C}$ ist genau dann absolut integrierbar, wenn $\alpha > 1$ gilt.

(3) Die durch $g(x) := e^{-\alpha|x|}$ definierte Funktion $g : \mathbb{R} \rightarrow \mathbb{C}$ ist für jedes $\alpha \in \mathbb{C}$ mit $\operatorname{Re} \alpha > 0$ absolut integrierbar.

(4) Eine absolut integrierbare Funktion muss nicht beschränkt sein. Die Funktion $f : \mathbb{R} \rightarrow \mathbb{C}$, die gegeben ist durch $f(x) := \sqrt{n}$ für $x \in [n, n + 1/n^2)$ und $n \in \mathbb{N}$ und $f(x) := 0$ sonst, ist absolut integrierbar.

25.3. Eigenschaften: Seien $f, g : \mathbb{R} \rightarrow \mathbb{C}$ für jedes $R > 0$ über $[-R, R]$ integrierbar.

(a) Gilt $|f| \leq |g|$ auf $\mathbb{R}$ und ist $g \in \mathcal{A}(\mathbb{R}, \mathbb{C})$, so ist auch $f \in \mathcal{A}(\mathbb{R}, \mathbb{C})$.

(b) Ist $f \in \mathcal{A}(\mathbb{R}, \mathbb{C})$ und g beschränkt, so ist $fg \in \mathcal{A}(\mathbb{R}, \mathbb{C})$ und

$$\|fg\|_1 \leq \|f\|_1 \|g\|_{\infty}.$$

25.4. Fouriertransformation: Sei $f \in \mathcal{A}(\mathbb{R}, \mathbb{C})$. Die *Fouriertransformierte* von f definiert man durch

$$\mathcal{F}f(\xi) := \widehat{f}(\xi) := \int_{-\infty}^{\infty} f(x) e^{-i\xi x} dx \quad (\xi \in \mathbb{R}).$$

Das Integral konvergiert dabei für jedes $\xi \in \mathbb{R}$ absolut, da $|f(x)e^{-i\xi x}| = |f(x)|$ ($x \in \mathbb{R}$) gilt und f absolut integrierbar ist (vgl. 25.3). Wir erhalten

$$|\widehat{f}(\xi)| \leq \|f\|_1 \quad (\xi \in \mathbb{R}),$$

so dass $\widehat{f} : \mathbb{R} \rightarrow \mathbb{C}$ beschränkt ist mit $\|\widehat{f}\|_{\infty} \leq \|f\|_1$.

Beispiele: (1) $f : \mathbb{R} \rightarrow \mathbb{C}$, $f(x) = e^{-a|x|}$, ist für $a > 0$ absolut integrierbar. Es gilt

$$\widehat{f}(\xi) = \frac{2a}{a^2 + \xi^2} \quad (\xi \in \mathbb{R}).$$

Die Funktion $\widehat{f} : \mathbb{R} \rightarrow \mathbb{C}$ ist absolut integrierbar.

(2) Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ definiert durch $f(x) = 1$ für $x \in [-1, 1]$ und $f(x) = 0$ sonst. Dann ist f absolut integrierbar und

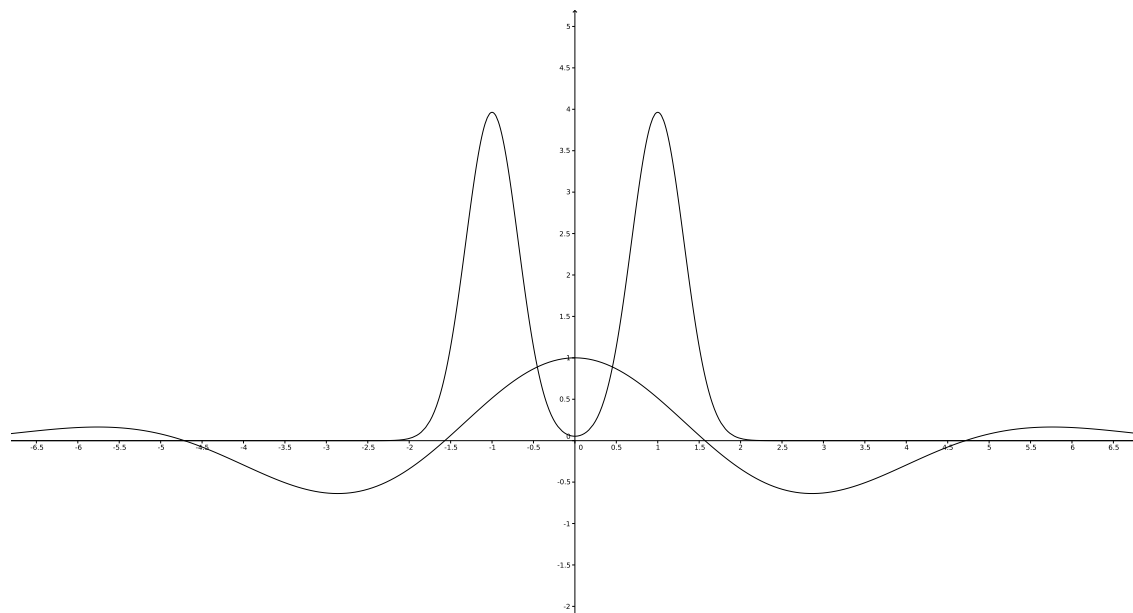
$$\widehat{f}(\xi) = \begin{cases} 2 \frac{\sin \xi}{\xi}, & \xi \in \mathbb{R} \setminus \{0\} \\ 2, & \xi = 0 \end{cases}.$$

Die Funktion $\widehat{f} : \mathbb{R} \rightarrow \mathbb{C}$ ist nicht absolut integrierbar.

(3) $f : \mathbb{R} \rightarrow \mathbb{C}$, $f(x) = e^{-x^2/20} \cos(x) = e^{-x^2/20} (e^{ix} + e^{-ix})/2$ ist absolut integrierbar. Es gilt

$$\widehat{f}(\xi) = \sqrt{5\pi} e^{-5(\xi+1)^2} (e^{20\xi} + 1) \quad (\xi \in \mathbb{R}).$$

Allgemein gilt: Ist f (wie in diesem Beispiel und auch in den Beispielen (1),(2)) reellwertig und gerade, so ist auch $\widehat{f}$ reellwertig und gerade (Übung). Folgendes Bild zeigt f und $\widehat{f}$ in diesem Beispiel.



25.5. Satz: (Riemann-Lebesgue). Es sei $f \in \mathcal{A}(\mathbb{R}, \mathbb{C})$. Dann ist die Funktion $\widehat{f} : \mathbb{R} \rightarrow \mathbb{C}$ gleichmäßig stetig und es gilt $\widehat{f}(\xi) \rightarrow 0$ ($|\xi| \rightarrow \infty$).

Beweis. Sei $\varepsilon > 0$. Dann existiert ein $\delta > 0$ mit $|e^{-ix} - 1| \leq \varepsilon$ ($|x| \leq \delta$). Für alle $\xi \in \mathbb{R}$ und $h \neq 0$ gilt dann

$$\begin{aligned} |\widehat{f}(\xi + h) - \widehat{f}(\xi)| &\leq \int_{-\infty}^{\infty} |f(x)(e^{-i(\xi+h)x} - e^{-i\xi x})| dx = \int_{-\infty}^{\infty} |f(x)| |e^{-ihx} - 1| dx \\ &\leq \underbrace{\varepsilon \int_{|x| \leq \delta/|h|} |f(x)| dx}_{\leq \int_{-\infty}^{\infty} |f(x)| dx} + 2 \underbrace{\int_{|x| \geq \delta/|h|} |f(x)| dx}_{\rightarrow 0 \text{ (} h \rightarrow 0 \text{)}}. \end{aligned}$$

Also ist $\widehat{f}$ gleichmäßig stetig.

Ist f gegeben durch $f(x) = 1$ ($x \in [a, b]$), und $f(x) = 0$ sonst, so gilt für $\xi \neq 0$:

$$\widehat{f}(\xi) = \int_a^b e^{-i\xi x} dx = \frac{e^{-i\xi b} - e^{-i\xi a}}{-i\xi}$$

und somit $\widehat{f}(\xi) \rightarrow 0$ für $|\xi| \rightarrow \infty$. Dies gilt dann auch für alle Linearkombinationen solcher Funktionen. Den allgemeinen Fall erhält man durch ein Approximationsargument. $\square$

25.6. Rechenregeln: Es sei $f \in \mathcal{A}(\mathbb{R}, \mathbb{C})$. Dann gilt:

- (a) Ist $a \in \mathbb{R} \setminus \{0\}$, so gilt $\mathcal{F}\{x \mapsto f(ax)\}(\xi) = \frac{1}{|a|} \mathcal{F}f\left(\frac{\xi}{a}\right)$ ($\xi \in \mathbb{R}$).
- (b) Ist $b \in \mathbb{R}$, so gilt $\mathcal{F}\{x \mapsto f(x - b)\}(\xi) = e^{-i\xi b} \mathcal{F}f(\xi)$ ($\xi \in \mathbb{R}$).
- (c) Ist $b \in \mathbb{R}$, so gilt $\mathcal{F}\{x \mapsto e^{ibx} f(x)\}(\xi) = \mathcal{F}f(\xi - b)$ ($\xi \in \mathbb{R}$).
- (d) Ist $x \mapsto xf(x)$ absolut integrierbar, so ist $\mathcal{F}f$ stetig differenzierbar und es gilt

$$(\mathcal{F}f)'(\xi) = \mathcal{F}\{x \mapsto (-ix)f(x)\}(\xi) \quad (\xi \in \mathbb{R}).$$

- (e) Ist f stetig und stückweise stetig differenzierbar derart, dass f' wieder absolut integrierbar ist, so gilt

$$\mathcal{F}\{f'\}(\xi) = i\xi \mathcal{F}f(\xi) \quad (\xi \in \mathbb{R}).$$

Dabei nennen wir f stückweise stetig differenzierbar, falls eine streng monoton wachsende Folge $(x_j)_{j \in \mathbb{Z}}$ existiert mit $x_j \rightarrow \pm\infty$ für $j \rightarrow \pm\infty$ so, dass f auf jedem Intervall (x_j, x_{j+1}) stetig differenzierbar ist und die einseitigen Grenzwerte $f(x_j \pm)$, $f'(x_j \pm)$ existieren. An den Stellen x_j kann f' beliebig gesetzt werden.

Beweis. Man erhält (a) und (b) durch naheliegende Substitutionen und (c) direkt.

Bei (d) findet man zu $\varepsilon > 0$ ein $\delta > 0$ mit $|1 - \frac{e^{it} - 1}{it}| < \varepsilon$ für $|t| < \delta$. Außerdem ist $\sup_{t \in \mathbb{R}} |1 - \frac{e^{it} - 1}{it}| =: M < \infty$ ($(e^z - 1)/z$ in 0 stetig durch 1 fortgesetzt). Für $h \neq 0$ ist dann

$$\left| \frac{\widehat{f}(\xi + h) - \widehat{f}(\xi)}{h} + \int_{-\infty}^{\infty} ix f(x) e^{-ix\xi} dx \right| = \left| \int_{-\infty}^{\infty} \left(1 - \frac{e^{-ihx} - 1}{-ihx}\right) ix f(x) e^{-ix\xi} dx \right|$$

$$\begin{aligned}
&\leq \varepsilon \int_{|x| \leq \delta/|h|} |xf(x)| dx + M \int_{|x| \geq \delta/|h|} |xf(x)| dx \\
&\leq \varepsilon \int_{-\infty}^{\infty} |xf(x)| dx + M \int_{|x| \geq \delta/|h|} |xf(x)| dx,
\end{aligned}$$

wobei das letzte Integral nach Voraussetzung wieder für $h \rightarrow 0$ gegen 0 geht. Die Stetigkeit von $(\mathcal{F}f)'$ folgt mit 25.5.

Bei (e) beachte man, dass absolute Integrierbarkeit von f' die Existenz der Grenzwerte $f(\pm\infty) := \lim_{x \rightarrow \pm\infty} f(x)$ impliziert, da

$$f(x) - f(0) = \int_0^x f'(t) dt$$

für $x \rightarrow \pm\infty$ konvergiert. Da f außerdem absolut integrierbar ist, muss $f(\pm\infty) = 0$ gelten. Für $R > 0$ ist dann mittels partieller Integration

$$\int_{-R}^R f'(x) e^{-i\xi x} dx = f(x) e^{-i\xi x} \Big|_{-R}^R + i\xi \int_{-R}^R f(x) e^{-i\xi x} dx,$$

und (e) folgt für $R \rightarrow \infty$. □

Bemerkung: Aussage (e) zeigt insbesondere den folgenden Spezialfall von 25.5: Ist $f \in \mathcal{A}(\mathbb{R}, \mathbb{C})$ stetig und stückweise stetig differenzierbar derart, dass f' wieder absolut integrierbar ist, so gilt $\widehat{f}(\xi) \rightarrow 0$ ($|\xi| \rightarrow \infty$), denn für $\xi \neq 0$ ist

$$|\widehat{f}(\xi)| = \left| \frac{1}{i\xi} \widehat{f'}(\xi) \right| \leq \frac{1}{|\xi|} \int_{-\infty}^{\infty} |f'(x)| dx \rightarrow 0 \quad (|\xi| \rightarrow \infty).$$

25.7. Beispiel: Die Funktion $f(x) = e^{-x^2/2}$ ($x \in \mathbb{R}$) ist Lösung der homogenen linearen Differentialgleichung $f'(x) = -xf(x)$. Wir wenden die Fouriertransformation auf diese Gleichung an (beachte $f, f' \in \mathcal{A}(\mathbb{R}, \mathbb{C})$) und erhalten:

$$i\xi \mathcal{F}f(\xi) = \mathcal{F}\{f'\}(\xi) = \mathcal{F}\{x \mapsto -xf(x)\}(\xi) = -i(\mathcal{F}f)'(\xi) \quad (\xi \in \mathbb{R}),$$

also $(\mathcal{F}f)'(\xi) = -\xi(\mathcal{F}f)(\xi)$. Somit löst $\mathcal{F}f$ dieselbe Differentialgleichung wie f und es folgt $\mathcal{F}f = cf$ für ein $c \in \mathbb{R}$. Wegen $f(0) = 1$ und

$$\widehat{f}(0) = \int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{2\pi}$$

folgt $c = \sqrt{2\pi}$, also

$$\widehat{f}(\xi) = \sqrt{2\pi} e^{-\xi^2/2} \quad (\xi \in \mathbb{R}).$$

25.8. Dancing-Hat-Lemma: Seien $f, g \in \mathcal{A}(\mathbb{R}, \mathbb{C})$. Dann gilt:

$$\int_{-\infty}^{\infty} \widehat{f}(\xi) g(\xi) d\xi = \int_{-\infty}^{\infty} f(\eta) \widehat{g}(\eta) d\eta.$$

Beweis. Nur für f, g mit $f(x) = g(x) = 0$ für $x \notin [-a, a]$ (allgemeiner Fall durch Approximation). Wir vertauschen die Integrationsreihenfolge

$$\int_{-a}^a \left(\int_{-a}^a f(\eta) e^{-i\xi\eta} d\eta \right) g(\xi) d\xi = \int_{-a}^a f(\eta) \left(\int_{-a}^a g(\xi) e^{-i\xi\eta} d\xi \right) d\eta.$$

□

25.9. Fourierinversionsformel: Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ stetig, beschränkt, absolut integrierbar und derart, dass $\widehat{f} : \mathbb{R} \rightarrow \mathbb{C}$ auch absolut integrierbar ist. Dann gilt:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{f}(\xi) e^{ix\xi} d\xi \quad (x \in \mathbb{R}).$$

Beweis. Es reicht, die Formel für $x = 0$ zu zeigen (verwende die Verschiebungsregel 25.6 (b)). Wir setzen $h(x) = e^{-x^2/2}$ und $g(x) = h(ax)$, mit $a > 0$. Dann gilt nach 25.8 und 25.6 (a):

$$\int_{-\infty}^{\infty} \widehat{f}(\xi) h(a\xi) d\xi = \int_{-\infty}^{\infty} f(x) \widehat{g}(x) dx = \int_{-\infty}^{\infty} f(x) \frac{1}{a} \widehat{h}(x/a) dx = \int_{-\infty}^{\infty} f(ax) \widehat{h}(x) dx.$$

Wir zeigen:

Für $a \rightarrow 0+$ konvergiert die rechte Seite gegen $2\pi f(0)$ und die linke Seite gegen $\int_{-\infty}^{\infty} \widehat{f}(\xi) d\xi$.

Linke Seite: Es gilt

$$\left| \int_{-\infty}^{\infty} \widehat{f}(\xi) h(a\xi) d\xi - \int_{-\infty}^{\infty} \widehat{f}(\xi) d\xi \right| \leq \int_{-\infty}^{\infty} |\widehat{f}(\xi)| |h(a\xi) - 1| d\xi =: I.$$

Zu gegebenem $\varepsilon > 0$ finden wir ein $\delta > 0$ mit $|h(\xi) - 1| \leq \varepsilon$ ($|\xi| \leq \delta$), und es gilt $|h(\xi) - 1| \leq 2$ ($\xi \in \mathbb{R}$). Wir können dann weiter abschätzen

$$I = \int_{|\xi| \leq \delta/a} \dots + \int_{|\xi| \geq \delta/a} \dots \leq \varepsilon \int_{-\infty}^{\infty} |\widehat{f}(\xi)| d\xi + 2 \underbrace{\int_{|\xi| \geq \delta/a} |\widehat{f}(\xi)| d\xi}_{\rightarrow 0, (a \rightarrow 0+)}.$$

Rechte Seite: Nach 25.7 gilt:

$$\int_{-\infty}^{\infty} \widehat{h}(x) dx = \int_{-\infty}^{\infty} \sqrt{2\pi} e^{-x^2/2} dx = 2\pi.$$

Zu gegebenem $\varepsilon > 0$ finden wir $\delta > 0$ mit $|f(x) - f(0)| \leq \varepsilon$ ($|x| \leq \delta$). Da f beschränkt ist, gibt es ein $M > 0$ mit $|f(x) - f(0)| \leq M$ ($x \in \mathbb{R}$). Wir erhalten

$$\begin{aligned} \left| \int_{-\infty}^{\infty} f(ax) \widehat{h}(x) dx - 2\pi f(0) \right| &\leq \int_{-\infty}^{\infty} |f(ax) - f(0)| \widehat{h}(x) dx = \int_{|x| \leq \delta/a} \dots + \int_{|x| \geq \delta/a} \dots \\ &\leq \varepsilon 2\pi + M \underbrace{\int_{|x| \geq \delta/a} \widehat{h}(x) dx}_{\rightarrow 0, (a \rightarrow 0+)} \end{aligned}$$

□

Bemerkung: Die Voraussetzung “ f beschränkt” kann in 25.9 weggelassen werden, da man zeigen kann, dass dies (bei stetigem f) aus der absoluten Integrierbarkeit von $\widehat{f}$ folgt.

Beispiel: Betrachte $f(x) = e^{-|x|}$ ($x \in \mathbb{R}$). Nach 25.4 (1) gilt

$$\widehat{f}(\xi) = \frac{2}{1 + \xi^2} \quad (\xi \in \mathbb{R}).$$

Mit 25.9 folgt

$$e^{-|x|} = e^{-|-x|} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{2}{1 + \xi^2} e^{-ix\xi} d\xi \quad (x \in \mathbb{R}),$$

also nach vertauschen von x und ξ

$$e^{-|\xi|} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{2}{1 + x^2} e^{-i\xi x} dx \quad (\xi \in \mathbb{R}).$$

Also gilt:

$$\mathcal{F}\{x \mapsto \frac{2}{1 + x^2}\}(\xi) = 2\pi e^{-|\xi|} \quad (\xi \in \mathbb{R}).$$

25.10. Satz von Plancherel: Sind $f, g \in \mathcal{A}(\mathbb{R}, \mathbb{C})$ und gilt

$$\int_{-\infty}^{\infty} |f(x)|^2 dx < \infty, \quad \int_{-\infty}^{\infty} |g(x)|^2 dx < \infty,$$

so ist

$$\int_{-\infty}^{\infty} f(x) \overline{g(x)} dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{f}(\xi) \overline{\widehat{g}(\xi)} d\xi,$$

also insbesondere

$$\int_{-\infty}^{\infty} |f(x)|^2 dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\widehat{f}(\xi)|^2 d\xi.$$

Beweis. Sind zusätzlich $\widehat{f}, \widehat{g}$ absolut integrierbar, so folgt mit 25.9

$$\begin{aligned}
\int_{-\infty}^{\infty} f(x) \overline{g(x)} dx &= \int_{-\infty}^{\infty} f(x) \overline{\left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{g}(\xi) e^{ix\xi} d\xi \right)} dx \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) \left(\int_{-\infty}^{\infty} \overline{\widehat{g}(\xi)} e^{-ix\xi} d\xi \right) dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(x) \overline{\widehat{g}(\xi)} e^{-ix\xi} d\xi \right) dx \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(x) \overline{\widehat{g}(\xi)} e^{-ix\xi} dx \right) d\xi = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(x) e^{-ix\xi} dx \right) \overline{\widehat{g}(\xi)} d\xi \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{f}(\xi) \overline{\widehat{g}(\xi)} d\xi.
\end{aligned}$$

Den allgemeinen Fall erhält man durch ein Approximationsargument. $\square$

25.11. Quadratintegrierbare Funktionen: Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ eine Funktion, die für jedes $R > 0$ über $[-R, R]$ integrierbar ist. Wir nennen die Funktion f *quadratintegrierbar* (qib), falls das uneigentliche Integral

$$\int_{-\infty}^{\infty} |f(x)|^2 dx$$

konvergiert. In diesem Fall setzen wir

$$\|f\|_2 := \left(\int_{-\infty}^{\infty} |f(x)|^2 dx \right)^{1/2}.$$

Die Menge der quadratintegrierbaren Funktionen bezeichnen wir mit $\mathcal{A}^2(\mathbb{R}, \mathbb{C})$.

Beispiele: (1) Ist $f : \mathbb{R} \rightarrow \mathbb{C}$ stetig und $\{x \in \mathbb{R} : f(x) \neq 0\}$ beschränkt, so ist $f \in \mathcal{A}^2(\mathbb{R}, \mathbb{C})$.

(2) Die durch $f(x) := (1 + |x|)^{-\alpha}$ definierte Funktion $f : \mathbb{R} \rightarrow \mathbb{C}$ ist genau dann quadratintegrierbar, wenn $\alpha > 1/2$ gilt. Es gibt also quadratintegrierbare Funktionen, die nicht absolut integrierbar sind.

(3) Die durch $g(x) := e^{-\alpha|x|}$ definierte Funktion $g : \mathbb{R} \rightarrow \mathbb{C}$ ist für jedes $\alpha \in \mathbb{C}$ mit $\operatorname{Re} \alpha > 0$ quadratintegrierbar.

Bemerkungen: (1) Sind $f, g \in \mathcal{A}^2(\mathbb{R}, \mathbb{C})$ quadratintegrierbar, so ist $fg \in \mathcal{A}(\mathbb{R}, \mathbb{C})$ und

$$\|fg\|_1 \leq \|f\|_2 \|g\|_2.$$

(2) $\mathcal{A}^2(\mathbb{R}, \mathbb{C})$ ist ein komplexer Vektorraum. Die durch

$$(f|g) := \int_{-\infty}^{\infty} f(x) \overline{g(x)} dx$$

definierte Abbildung hat die Eigenschaften eines Skalarproduktes (bis auf die Definitheit, die man aber erhalten kann, wenn man Funktionen f, g mit $\|f - g\|_2 = 0$ identifiziert). Es gilt

$$\|\alpha f\|_2 = |\alpha| \|f\|_2 \quad (\alpha \in \mathbb{C}), \quad \|f + g\|_2 \leq \|f\|_2 + \|g\|_2.$$

Die Aussage von 25.10 bedeutet also: Ist $f \in \mathcal{A}^2(\mathbb{R}, \mathbb{C}) \cap \mathcal{A}(\mathbb{R}, \mathbb{C})$, so ist $\widehat{f} \in \mathcal{A}^2(\mathbb{R}, \mathbb{C})$ und

$$\|\widehat{f}\|_2 = \sqrt{2\pi} \|f\|_2,$$

d.h. die Fouriertransformation ändert die “2-Norm” $\|\cdot\|_2$ nur um eine Konstante.

(3) Ist $f : \mathbb{R} \rightarrow \mathbb{C}$ beschränkt und absolut integrierbar, so ist $f \in \mathcal{A}^2(\mathbb{R}, \mathbb{C})$ und

$$\|f\|_2^2 = \int_{-\infty}^{\infty} |f(x)|^2 dx \leq \|f\|_{\infty} \|f\|_1.$$

25.12. Heisenbergsche Unschärferelation (1-dim.): Sei $\psi : \mathbb{R} \rightarrow \mathbb{C}$ absolut integrierbar, beschränkt und stetig differenzierbar. Außerdem seien ψ' und $x \mapsto x\psi(x)$ absolut integrierbar und beschränkt. Dann gilt

$$\sqrt{\frac{\pi}{2}} \|\psi\|_2^2 \leq \|x\psi\|_2 \|\xi \widehat{\psi}\|_2.$$

Beachte dabei, dass nach der letzten Bemerkung in 25.11 die Funktionen ψ , ψ' und $x \mapsto x\psi(x)$ quadratintegrierbar sind. Nach 25.10 (Plancherel) ist dann $\widehat{\psi}'$ quadratintegrierbar und somit ist auch $\xi \mapsto \xi \widehat{\psi}(\xi) = -i\widehat{\psi}'(\xi)$ quadratintegrierbar (verwende 25.6 (e)).

Beweis. Wir zeigen zuerst $x|\psi(x)|^2 \rightarrow 0$ ($x \rightarrow \pm\infty$): Es gilt

$$\frac{d}{dx} x|\psi(x)|^2 = \frac{d}{dx} x\psi(x)\overline{\psi(x)} = |\psi(x)|^2 + \psi'(x)(x\overline{\psi(x)}) + (x\psi(x))\overline{\psi'(x)}.$$

Da ψ , ψ' und $x \mapsto x\psi(x)$ quadratintegrierbar sind, sind

$$x \mapsto x|\psi(x)|^2 \quad \text{und} \quad x \mapsto \frac{d}{dx} x|\psi(x)|^2$$

absolut integrierbar: Nach Bemerkung (1) in 25.11 sind die Funktionen

$$x \mapsto \psi(x)\overline{\psi(x)}, \quad x \mapsto (x\psi(x))\overline{\psi(x)}, \quad x \mapsto \psi'(x)(x\overline{\psi(x)}), \quad x \mapsto (x\psi(x))\overline{\psi'(x)}$$

absolut integrierbar (jeweils als Produkt zweier quadratintegrierbarer Funktionen). Also ist $x \mapsto x|\psi(x)|^2$ stetig differenzierbar und absolut integrierbar mit absolut integrierbarer Ableitung. Wie im Beweis von 25.6 (e) folgt damit $x|\psi(x)|^2 \rightarrow 0$ ($x \rightarrow \pm\infty$).

Nun ist

$$\begin{aligned}
2\operatorname{Re}(x\psi|\psi') &= (x\psi|\psi') + (\psi'|x\psi) \\
&= \int_{-\infty}^{\infty} x\psi(x)\overline{\psi'(x)} + \psi'(x)\overline{x\psi(x)} dx \\
&= \int_{-\infty}^{\infty} x(\psi(x)\overline{\psi'(x)} + \psi'(x)\overline{\psi(x)}) dx \\
&= x|\psi(x)|^2 \Big|_{x=-\infty}^{\infty} - \int_{-\infty}^{\infty} |\psi(x)|^2 dx \\
&= -\|\psi\|_2^2.
\end{aligned}$$

Mit der Cauchy-Schwarzschen Ungleichung folgt

$$(*) \quad \frac{1}{2}\|\psi\|_2^2 = |\operatorname{Re}(x\psi|\psi')| \leq |(x\psi|\psi')| \leq \|x\psi\|_2\|\psi'\|_2.$$

Außerdem ist nach 25.10 (Plancherel) und 25.6 (e)

$$\|\psi'\|_2 = (2\pi)^{-1/2}\|\widehat{\psi'}\|_2 = (2\pi)^{-1/2}\|\xi\widehat{\psi}\|_2.$$

Einsetzen in (*) liefert die Behauptung. □

Korollar: Ist ψ wie oben und sind $x_0, \xi_0 \in \mathbb{R}$, so gilt

$$\sqrt{\frac{\pi}{2}}\|\psi\|_2^2 \leq \|(x - x_0)\psi\|_2\|(\xi - \xi_0)\widehat{\psi}\|_2.$$

Beweis. Betrachte $g(x) := e^{-i\xi_0 x}\psi(x + x_0)$ und beachte dabei $|g(x)| = |\psi(x + x_0)|$ und $|\widehat{g}(\xi)| = |\widehat{\psi}(\xi + \xi_0)|$. □

Interpretation: Ist ψ wie in 25.12 mit $\|\psi\|_2 = 1$, so sind

$$x \mapsto |\psi(x)|^2 \quad \text{und} \quad \xi \mapsto (2\pi)^{-1}|\widehat{\psi}(\xi)|^2$$

Wahrscheinlichkeitsdichten auf $\mathbb{R}$. Sind X, Y Zufallsvariablen, die entsprechend verteilt sind und $x_0 := E(X) = \int_{-\infty}^{\infty} x|\psi(x)|^2 dx$ der Erwartungswert von X , sowie $\xi_0 := E(Y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \xi|\widehat{\psi}(\xi)|^2 d\xi$ der Erwartungswert von Y , so sind $\sigma_X := \|(x - x_0)\psi\|_2$ und $\sigma_Y := (2\pi)^{-1/2}\|(\xi - \xi_0)\widehat{\psi}\|_2$ die Standardabweichungen von X bzw. Y , und das Korollar besagt

$$\frac{1}{2} \leq \sigma_X \sigma_Y,$$

d.h. X und Y können nicht gleichzeitig beliebig gut um x_0 bzw. ξ_0 lokalisiert sein. In der Quantenmechanik entsprechen X und Y Ort bzw. Impuls eines Teilchens im Zustand ψ .

Beispiel: Wir betrachten $\psi(x) = e^{-x^2/2}$ ($x \in \mathbb{R}$). Die Heisenbergsche Unschärferelation lautet hier (beachte $\widehat{\psi}(\xi) = \sqrt{2\pi}e^{-\xi^2/2}$):

$$\sqrt{\frac{\pi}{2}} \int_{-\infty}^{\infty} e^{-x^2} dx \leq \left(\int_{-\infty}^{\infty} x^2 e^{-x^2} dx \right)^{1/2} \left(\int_{-\infty}^{\infty} \xi^2 (2\pi e^{-\xi^2}) d\xi \right)^{1/2}.$$

Tatsächlich gilt hier sogar “=”, denn die Funktionen $x \mapsto x\psi(x)$ und ψ' sind reell und linear abhängig, so dass in (*) überall Gleichheit gilt. Es gilt also

$$\frac{1}{2} \int_{-\infty}^{\infty} e^{-x^2} dx = \int_{-\infty}^{\infty} x^2 e^{-x^2} dx.$$

Wegen

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi} \quad \text{folgt} \quad \int_{-\infty}^{\infty} x^2 e^{-x^2} dx = \frac{\sqrt{\pi}}{2}.$$

25.13. Faltung und Fouriertransformation: Seien $f, g \in \mathcal{A}(\mathbb{R}, \mathbb{C})$ und g (oder f) sei beschränkt. Für jedes $t \in \mathbb{R}$ ist dann die Funktion $x \mapsto f(x)g(t-x)$ absolut integrierbar und die Funktion

$$h : \mathbb{R} \rightarrow \mathbb{C}, \quad h(t) := \int_{-\infty}^{\infty} f(x)g(t-x) dx$$

heißt *Faltung von f und g* , geschrieben $h =: f * g$. Die Funktion $f * g$ ist stetig, beschränkt und absolut integrierbar und es gilt

$$\widehat{f * g}(\xi) = \widehat{f}(\xi) \cdot \widehat{g}(\xi) \quad (\xi \in \mathbb{R}).$$

Bemerkung: Gilt zusätzlich $f(t) = g(t) = 0$ für $t < 0$, so ist $(f * g)(t) = 0$ für $t < 0$ und

$$(f * g)(t) = \int_0^t f(x)g(t-x) dx \quad (t \geq 0).$$

Beweis. Zunächst ist das Integral in der Definition absolut konvergent und eine Substitution gibt

$$h(t) = \int_{-\infty}^{\infty} g(x)f(t-x) dx \quad (t \in \mathbb{R}).$$

Es gilt

$$|h(t)| \leq \int_{-\infty}^{\infty} |f(x)||g(t-x)| dx \leq \|g\|_{\infty} \int_{-\infty}^{\infty} |f(x)| dx \quad (t \in \mathbb{R}).$$

Somit ist h beschränkt mit $\|h\|_{\infty} \leq \|g\|_{\infty} \|f\|_1$. Für $t, \tau \in \mathbb{R}$ gilt

$$|h(t+\tau) - h(t)| \leq \int_{-\infty}^{\infty} |g(x)||f(t+\tau-x) - f(t-x)| dx \leq \|g\|_{\infty} \int_{-\infty}^{\infty} |f(\tau+y) - f(y)| dy,$$

wenn wir $x = t - y$ substituieren. Dabei konvergiert das letzte Integral für $\tau \rightarrow 0$ gegen Null (hier ohne Beweis). Also ist h stetig.

Weiter gilt durch Vertauschung der Integrationsreihenfolge (zunächst wieder für f, g stetig mit $f(x) = g(x) = 0$ für $x \notin [-a, a]$, allgemeine Funktionen muss man geeignet approximieren):

$$\begin{aligned} \int_{-\infty}^{\infty} |h(t)| dt &\leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(x)| |g(t-x)| dx dt \\ &= \int_{-\infty}^{\infty} |f(x)| \int_{-\infty}^{\infty} |g(t-x)| dt dx \\ &= \int_{-\infty}^{\infty} |f(x)| dx \int_{-\infty}^{\infty} |g(t)| dt < \infty. \end{aligned}$$

Somit ist h absolut integrierbar. Der Beweis für die Faltungsregel der Fouriertransformation verwendet ähnliche Argumente:

$$\begin{aligned} \widehat{h}(\xi) &= \int_{-\infty}^{\infty} h(t) e^{-i\xi t} dt \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x) e^{-i\xi x} g(t-x) e^{-i\xi(t-x)} dx dt \\ &= \int_{-\infty}^{\infty} f(x) e^{-i\xi x} \int_{-\infty}^{\infty} g(t-x) e^{-i\xi(t-x)} dt dx \\ &= \int_{-\infty}^{\infty} f(x) e^{-i\xi x} dx \cdot \int_{-\infty}^{\infty} g(t) e^{-i\xi t} dt = \widehat{f}(\xi) \cdot \widehat{g}(\xi). \end{aligned}$$

□

25.14. Anwendung: Gegeben sei eine stetige und absolut integrierbare Funktion $f : \mathbb{R} \rightarrow \mathbb{C}$ und $q > 0$. Gesucht ist eine C^2 -Funktion $u : \mathbb{R} \rightarrow \mathbb{C}$ mit

$$-u''(x) + qu(x) = f(x) \quad (x \in \mathbb{R}).$$

Wenn wir annehmen, dass u, u', u'' absolut integrierbar sind, können wir die Fouriertransformation verwenden und erhalten via 25.6 die Gleichung

$$\xi^2 \widehat{u}(\xi) + q \widehat{u}(\xi) = \widehat{f}(\xi) \quad (\xi \in \mathbb{R}).$$

Wegen $q > 0$ ist $\xi \mapsto (\xi^2 + q)^{-1}$ ($\xi \in \mathbb{R}$) beschränkt, und wir erhalten für die Fouriertransformation der Lösung:

$$\widehat{u}(\xi) = (\xi^2 + q)^{-1} \widehat{f}(\xi) \quad (\xi \in \mathbb{R}).$$

Nun wird zurücktransformiert: Finden wir eine absolut integrierbare und beschränkte Funktion $k_q : \mathbb{R} \rightarrow \mathbb{C}$ mit $\widehat{k}_q(\xi) = (\xi^2 + q)^{-1}$, so ist $u = k_q * f$ nach 25.13 ein Kandidat. Da $\xi \mapsto (\xi^2 + q)^{-1}$ absolut integrierbar ist, kann man nach 25.9 ansetzen

$$k_q(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{ix\xi}}{\xi^2 + q} d\xi.$$

Wir berechnen k_q (mithilfe des Beispiels in 25.9):

$$k_q(x) = \frac{1}{2\sqrt{q}} e^{-\sqrt{q}|x|} \quad (x \in \mathbb{R}).$$

Diese Funktion ist beschränkt und absolut integrierbar. Eine Lösung der ursprünglichen Gleichung sollte also durch

$$u(x) := (k_q * f)(x) = \frac{1}{2\sqrt{q}} \int_{-\infty}^{\infty} e^{-\sqrt{q}|x-y|} f(y) dy \quad (x \in \mathbb{R})$$

gegeben sein. Das ist noch durch eine Probe zu überprüfen.

25.15. Die Fouriertransformation im Raum der schnell fallenden Funktionen:

Definition: Eine Funktion $f \in C^\infty(\mathbb{R}, \mathbb{C})$ heißt *schnell fallend* : $\Longleftrightarrow$

$$\forall n, m \in \mathbb{N}_0 : x \mapsto x^m f^{(n)}(x) \text{ ist beschränkt auf } \mathbb{R}.$$

Die Menge

$$\mathcal{S} := \mathcal{S}(\mathbb{R}, \mathbb{C}) := \{f : \mathbb{R} \rightarrow \mathbb{C} : f \text{ ist schnell fallend}\}$$

heißt *Schwartz-Raum*.

Beispiele: (1) $x \mapsto f(x) = p(x)e^{-\alpha x^2}$ ist für jedes $\alpha > 0$ und jedes Polynom $p \in \mathbb{C}[x]$ eine schnell fallende Funktion.

(2) Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ definiert durch

$$f(x) = \begin{cases} \exp\left(-\frac{1}{1-x^2}\right), & |x| < 1 \\ 0, & |x| \geq 1 \end{cases}$$

und $g \in C^\infty(\mathbb{R}, \mathbb{C})$. Dann ist $fg \in \mathcal{S}$.

Eigenschaften von $\mathcal{S}$ und von $\mathcal{F}$ auf $\mathcal{S}$: Es seien $f, g \in \mathcal{S}$, $p \in \mathbb{C}[x]$, und $k \in C^\infty(\mathbb{R}, \mathbb{C})$ mit $k^{(n)}$ beschränkt auf $\mathbb{R}$ ($n \in \mathbb{N}_0$). Dann gilt:

1. $f \in \mathcal{A}(\mathbb{R}, \mathbb{C}) \cap \mathcal{A}^2(\mathbb{R}, \mathbb{C})$ und $\lim_{x \rightarrow \pm\infty} f(x) = 0$.
2. $\forall \alpha, \beta \in \mathbb{C} : \alpha f + \beta g \in \mathcal{S}$ ($\mathcal{S}$ ist also ein komplexer Vektorraum).

3. $kf, pf, \bar{f}, \operatorname{Re} f, \operatorname{Im} f, x \mapsto f(-x) \in \mathcal{S}$.
4. $\widehat{f} \in \mathcal{S}$ und $f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{f}(\xi) e^{ix\xi} d\xi$ ($x \in \mathbb{R}$).
5. $f^{(n)} \in \mathcal{S}$ ($n \in \mathbb{N}$), und $\widehat{f^{(n)}}(\xi) = (i\xi)^n \widehat{f}(\xi)$ ($\xi \in \mathbb{R}, n \in \mathbb{N}$).
6. $x \mapsto f(x-b) \in \mathcal{S}$ ($b \in \mathbb{R}$) und $\mathcal{F}\{x \mapsto f(x-b)\}(\xi) = e^{-i\xi b} \widehat{f}(\xi)$ ($\xi \in \mathbb{R}$).
7. $f * g \in \mathcal{S}$ und $\widehat{f * g} = \widehat{f} \cdot \widehat{g}$.
8. Für $h(x) := e^{-x^2/2}$ ($x \in \mathbb{R}$) gilt: $h \in \mathcal{S}$ und $\widehat{h} = \sqrt{2\pi}h$ auf $\mathbb{R}$.
9. $\sqrt{\frac{\pi}{2}} \|f\|_2^2 \leq \|xf\|_2 \|\xi \widehat{f}\|_2$ (Heisenbergsche Unschärferelation).

Wir zeigen 1. (2.,3. sind leicht und 4.-9. folgen teilweise aus den bereits bekannten Eigenschaften der Fouriertransformation; Rest ohne Beweis):

Die Funktion $x \mapsto (1+x^2)f(x)$ ist beschränkt, also gilt

$$|f(x)| \leq \frac{M}{1+x^2} \quad (x \in \mathbb{R})$$

für ein $M \geq 0$. Damit folgt $\lim_{x \rightarrow \pm\infty} f(x) = 0$. Weiter konvergieren die Integrale

$$\int_{-\infty}^{\infty} \frac{M}{1+x^2} dx \quad \text{und} \quad \int_{-\infty}^{\infty} \frac{M^2}{(1+x^2)^2} dx.$$

Also gilt $f \in \mathcal{A}(\mathbb{R}, \mathbb{C}) \cap \mathcal{A}^2(\mathbb{R}, \mathbb{C})$.

Satz: Die Fouriertransformation $\mathcal{F} : \mathcal{S} \rightarrow \mathcal{S}$ ist ein Isomorphismus (also linear und bijektiv).

Beweis. Offensichtlich ist $\mathcal{F} : \mathcal{S} \rightarrow \mathcal{S}$ linear. Betrachte $\mathcal{G} : \mathcal{S} \rightarrow \mathcal{S}$ definiert durch

$$\mathcal{G}g(x) := \frac{1}{2\pi} \int_{-\infty}^{\infty} g(\xi) e^{ix\xi} d\xi = \frac{1}{2\pi} \widehat{g}(-x) \quad (x \in \mathbb{R}).$$

Nach 25.9 gilt: $\mathcal{G}(\mathcal{F}f) = f$ ($f \in \mathcal{S}$). Umgekehrt gilt für $g \in \mathcal{S}$:

$$\begin{aligned} \mathcal{F}(\mathcal{G}g)(\xi) &= \int_{-\infty}^{\infty} \frac{1}{2\pi} \widehat{g}(-x) e^{-i\xi x} dx \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{g}(x) e^{i\xi x} dx = g(\xi) \quad (\xi \in \mathbb{R}). \end{aligned}$$

Somit gilt auch $\mathcal{F}(\mathcal{G}g) = g$ ($g \in \mathcal{S}$). Damit ist $\mathcal{G} : \mathcal{S} \rightarrow \mathcal{S}$ sowohl Rechts- als auch Linksinverse von $\mathcal{F} : \mathcal{S} \rightarrow \mathcal{S}$. Also ist $\mathcal{F} : \mathcal{S} \rightarrow \mathcal{S}$ bijektiv und $\mathcal{G} = \mathcal{F}^{-1}$ auf $\mathcal{S}$. $\square$

Beispiel: Es sei $f \in \mathcal{S}$. Behauptung: Es gibt genau eine Funktion $u \in \mathcal{S}$ mit

$$(*) \quad 2u(x+1) + u(x) = f(x) \quad (x \in \mathbb{R}).$$

Beweis. Beachte zuerst: Für $k(\xi) := 1/(2e^{i\xi} + 1)$ ($\xi \in \mathbb{R}$) ist $k^{(n)}$ beschränkt auf $\mathbb{R}$ ($n \in \mathbb{N}_0$) (Übung). Daher gilt mit $u_1(x) := u(x+1)$ folgende Rechnung in $\mathcal{S}$:

$$\begin{aligned} 2u_1 + u = f &\iff 2\widehat{u}_1 + \widehat{u} = \widehat{f} \iff 2e^{i\xi}\widehat{u}(\xi) + \widehat{u}(\xi) = \widehat{f}(\xi) \quad (\xi \in \mathbb{R}) \\ \iff \widehat{u}(\xi) &= \frac{\widehat{f}(\xi)}{2e^{i\xi} + 1} \quad (\xi \in \mathbb{R}) \iff u(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\widehat{f}(\xi)e^{ix\xi}}{2e^{i\xi} + 1} d\xi \quad (x \in \mathbb{R}). \end{aligned}$$

□

Bemerkungen: (1) Es sei $f \in \mathcal{S}$ reellwertig und $u \in \mathcal{S}$ die Lösung von (*). Dann ist auch u reellwertig: Mit u ist auch $\operatorname{Re} u \in \mathcal{S}$ und es gilt

$$2\operatorname{Re} u(x+1) + \operatorname{Re} u(x) = f(x) \quad (x \in \mathbb{R}).$$

Wegen der Eindeutigkeit der Lösung von (*) in $\mathcal{S}$ folgt $\operatorname{Re} u = u$.

(2) **Die Lösungsmenge von Gleichungen hängt im allgemeinen stark vom Lösungsbegriff ab, d.h. von der Menge in der man nach Lösungen sucht.**

Z.B. hat (*) in $C^\infty(\mathbb{R}, \mathbb{C})$ stets unendlich viele Lösungen: Betrachte $v(x) := 2^{-x} \sin(\pi x)$ ($x \in \mathbb{R}$). Dann gilt $v \in C^\infty(\mathbb{R}, \mathbb{R})$ und

$$2v(x+1) + v(x) = 2 \cdot 2^{-(x+1)} \sin(\pi(x+1)) + 2^{-x} \sin(\pi x) = 0 \quad (x \in \mathbb{R}).$$

Also hat (*) in $C^\infty(\mathbb{R}, \mathbb{C})$ neben der Lösung $u \in \mathcal{S}$ auch die Lösungen $u + cv$, $c \in \mathbb{C}$.

Wärmeleitungsgleichung: Gesucht ist $u : [0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}$ mit

$$(*) \quad \begin{cases} u_t(t, x) &= u_{xx}(t, x) \quad ((t, x) \in (0, \infty) \times \mathbb{R}) \\ u(0, x) &= f(x) \quad (x \in \mathbb{R}). \end{cases}$$

Hierbei ist wieder t die Zeitvariable und x die Ortsvariable.

Voraussetzung: Es sei $f : \mathbb{R} \rightarrow \mathbb{R}$ in $\mathcal{S}$. Außerdem sei $u(t, \cdot) : \mathbb{R} \rightarrow \mathbb{R}$ für jedes $t > 0$ in $\mathcal{S}$.

Bemerkung: Diese Annahme für u ist physikalisch sinnvoll. Ist z.B. $f(x) \geq 0$ für alle $x \in \mathbb{R}$, so ist $\int_{-\infty}^{\infty} f(x) dx$ proportional zur Gesamtenergie. Wir erwarten $u(t, x) \geq 0$ und wegen Energieerhaltung

$$\int_{-\infty}^{\infty} u(t, x) dx = \int_{-\infty}^{\infty} f(x) dx \quad (t > 0).$$

Wir verwenden Fouriertransformation in x und setzen

$$\begin{aligned}\widehat{f}(\xi) &= \int_{-\infty}^{\infty} f(x) e^{-i\xi x} dx \quad (\xi \in \mathbb{R}), \\ \widehat{u}(t, \xi) &= \int_{-\infty}^{\infty} u(t, x) e^{-i\xi x} dx \quad (t > 0, \xi \in \mathbb{R}).\end{aligned}$$

Dann wird $(*)$ zu

$$(*) \quad \begin{cases} \widehat{u}_t(t, \xi) &= -\xi^2 \widehat{u}(t, \xi) \quad ((t, \xi) \in (0, \infty) \times \mathbb{R}) \\ \widehat{u}(0, \xi) &= \widehat{f}(\xi) \quad (\xi \in \mathbb{R}). \end{cases}$$

Beachte dabei: Wir gehen davon aus, dass

$$\widehat{u}_t(t, \xi) = \int_{-\infty}^{\infty} u_t(t, x) e^{-i\xi x} dx \quad ((t, \xi) \in (0, \infty) \times \mathbb{R}).$$

Für festes $\xi \in \mathbb{R}$ ist die eindeutige Lösung des linearen Anfangswertproblems $(*)$ gegeben durch

$$\widehat{u}(t, \xi) = e^{-t\xi^2} \widehat{f}(\xi) \quad (t > 0).$$

Wir hoffen also, dass wir durch Fourierinversion eine Lösung von $(*)$ erhalten. Nach der Faltungsregel ist für $g_t(x) := \mathcal{F}^{-1}\{\xi \mapsto e^{-t\xi^2}\}(x)$:

$$\mathcal{F}(g_t * f)(\xi) = e^{-t\xi^2} \widehat{f}(\xi) \quad (\xi \in \mathbb{R}),$$

also setzen wir

$$u(t, x) := g_t * f(x) \quad ((t, x) \in (0, \infty) \times \mathbb{R}).$$

Wir berechnen die Funktion g_t . Es ist

$$e^{-t\xi^2} = e^{-(\sqrt{2t}\xi)^2/2},$$

also

$$\begin{aligned}g_t(x) &= \mathcal{F}^{-1}\{\xi \mapsto e^{-t\xi^2}\}(x) = \frac{1}{2\pi} \mathcal{F}\{\xi \mapsto e^{-t\xi^2}\}(-x) \\ &= \frac{1}{2\pi} \mathcal{F}\{\xi \mapsto e^{-(\sqrt{2t}\xi)^2/2}\}(-x) = \frac{1}{2\pi} \frac{1}{\sqrt{2t}} \sqrt{2\pi} e^{-(x/\sqrt{2t})^2/2} = \frac{1}{\sqrt{4\pi t}} e^{-x^2/(4t)}.\end{aligned}$$

Somit ist

$$u(t, x) = \frac{1}{\sqrt{4\pi t}} \int_{-\infty}^{\infty} \exp\left(-\frac{(x-y)^2}{4t}\right) f(y) dy \quad ((t, x) \in (0, \infty) \times \mathbb{R}).$$

Ob bzw. in welchem Sinn dies nun eine Lösung von $(*)$ ist, muss man noch überprüfen. Im Fall $f \in \mathcal{S}$ ist tatsächlich u differenzierbar auf $(0, \infty) \times \mathbb{R}$ und ist dort Lösung von $u_t = u_{xx}$. Weiter ist u durch f stetig auf $[0, \infty) \times \mathbb{R}$ ergänzbar: Es gilt

$$\forall x \in \mathbb{R} : \lim_{t \rightarrow 0+} u(t, x) = f(x).$$

Ferner gilt $u(t, \cdot) \in \mathcal{S}$ ($t > 0$).

Bemerkung: Setzt man

$$h(t, x) := g_t(x) = \frac{1}{\sqrt{4\pi t}} e^{-x^2/(4t)},$$

so löst h die Wärmeleitungsgleichung. Auf $(0, \infty) \times \mathbb{R}$ gilt

$$h_t(t, x) = -\frac{1}{2}(4\pi t)^{-3/2} \cdot 4\pi e^{-x^2/(4t)} + (4\pi t)^{-1/2} e^{-x^2/(4t)} \cdot \frac{x^2}{4t^2},$$

$$h_x(t, x) = (4\pi t)^{-1/2} e^{-x^2/(4t)} \cdot \frac{-2x}{4t},$$

also

$$h_{xx}(t, x) = (4\pi t)^{-1/2} e^{-x^2/(4t)} \cdot \frac{x^2}{4t^2} + (4\pi t)^{-1/2} e^{-x^2/(4t)} \cdot \frac{-1}{2t} = h_t(t, x).$$

Ferner gilt

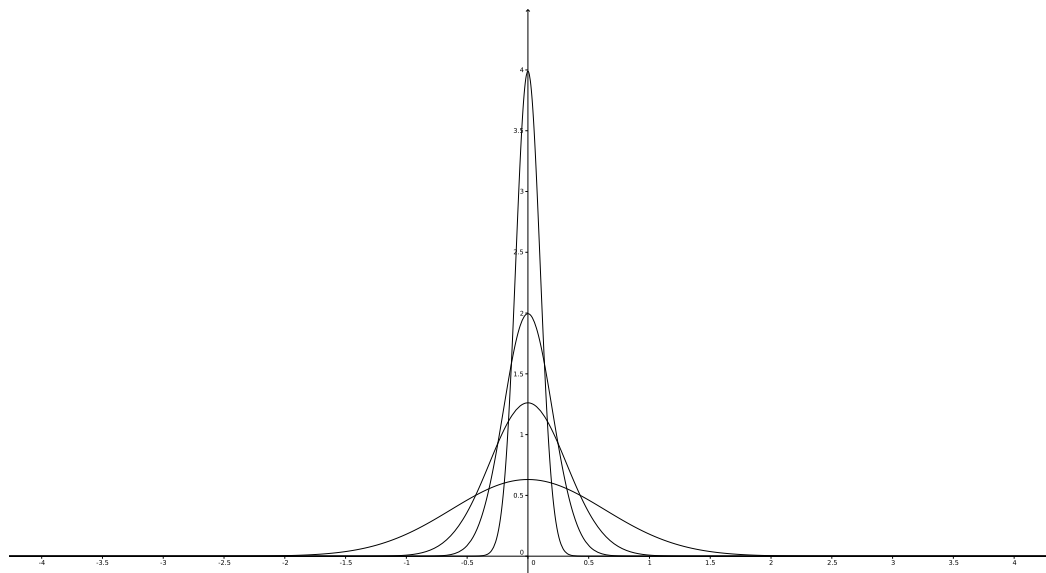
$$\begin{aligned} \forall x \in \mathbb{R} \setminus \{0\} : h(t, x) &\rightarrow 0 \quad (t \rightarrow 0+), \\ h(t, 0) &\rightarrow \infty \quad (t \rightarrow 0+), \end{aligned}$$

und

$$\int_{-\infty}^{\infty} h(t, x) dx = 1 \quad (t > 0)$$

Der zu dieser Lösung gehörige “Anfangswert” wäre die sogenannte δ -Distribution. Dies ist keine Funktion im üblichen Sinn ($\rightarrow$ Distributionentheorie).

Folgendes Bild zeigt $x \mapsto h(t, x)$ für einige $t > 0$.



25.16. Das Abtasttheorem von Shannon: Es sei $f \in \mathcal{A}(\mathbb{R}, \mathbb{C})$ und stetig. Weiter sei f so, dass

$$\widehat{f}(\xi) = 0 \quad (\xi \in \mathbb{R} \setminus (-\pi, \pi)).$$

Nach der Fourierinversionsformel 25.9 (beachte: $\widehat{f} \in \mathcal{A}(\mathbb{R}, \mathbb{C})$) gilt also

$$f(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \widehat{f}(\xi) e^{ix\xi} d\xi \quad (x \in \mathbb{R}).$$

Für festes $x \in \mathbb{R}$ seien $F, G_x : \mathbb{R} \rightarrow \mathbb{C}$ die 2π -periodischen Funktionen mit

$$F(\xi) = \widehat{f}(\xi), \quad G_x(\xi) = e^{-ix\xi} \quad (\xi \in [-\pi, \pi)).$$

Ihre Fourierkoeffizienten sind

$$\widehat{F}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \widehat{f}(\xi) e^{-ik\xi} d\xi = f(-k) \quad (k \in \mathbb{Z})$$

und

$$\widehat{G_x}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ix\xi} e^{-ik\xi} d\xi = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(x+k)\xi} d\xi = \text{si}((x+k)\pi) \quad (k \in \mathbb{Z}),$$

wobei $\text{si} : \mathbb{R} \rightarrow \mathbb{R}$ den *Sinus cardinalis* bezeichnet:

$$\text{si}(x) = \begin{cases} \frac{\sin x}{x}, & x \neq 0 \\ 1, & x = 0 \end{cases}.$$

In $C([-\pi, \pi], \mathbb{C})$ versehen mit dem Skalarprodukt $(\cdot | \cdot)$ und dem vollständigen Orthonormalsystem $(e_k)_{k \in \mathbb{Z}}$ aus 24.6 gilt nun (ebenso wie in $C_{\text{per}}([-\pi, \pi], \mathbb{C})$):

$$\begin{aligned} f(x) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} F(\xi) \overline{G_x(\xi)} d\xi = (F | G_x) = \sum_{k=-\infty}^{\infty} (F | e_k) \overline{(G_x | e_k)} \\ &= \sum_{k=-\infty}^{\infty} \widehat{F}(k) \overline{\widehat{G_x}(k)} = \sum_{k=-\infty}^{\infty} f(-k) \text{si}((x+k)\pi) = \sum_{k=-\infty}^{\infty} f(k) \text{si}((x-k)\pi), \end{aligned}$$

d.h. man kann f aus den Werten $f(k)$ ($k \in \mathbb{Z}$) reproduzieren.

Bemerkung: Wegen $(f(k))_{k \in \mathbb{Z}}, (\text{si}((x-k)\pi))_{k \in \mathbb{Z}} \in l_{\mathbb{C}}^2(\mathbb{Z})$ ist obige Reihe absolut konvergent für jedes $x \in \mathbb{R}$. Weiter ist sie auf jeder beschränkten Teilmenge von $\mathbb{R}$ gleichmäßig konvergent.

Obige Überlegungen kann man auf andere Intervalle transformieren.

Definition: Es sei $f \in \mathcal{A}(\mathbb{R}, \mathbb{C})$ und stetig. Wenn die Fouriertransformierte $\widehat{f} : \mathbb{R} \rightarrow \mathbb{C}$ außerhalb eines beschränkten Intervalls 0 ist, so heißt f *bandbeschränkt* (technisch: Die Frequenzdichte des Signals verschwindet außerhalb eines beschränkten Intervalls).

In diesem Fall ist es möglich f aus den Werten auf einem hinreichend feinen Raster $\{kT : k \in \mathbb{Z}\}$, $T > 0$ zu reproduzieren. Wie schnell man messen muß (d.h. wie klein man T wählen sollte) präzisiert das *Abtasttheorem von Shannon*:

Satz: Es sei $f \in \mathcal{A}(\mathbb{R}, \mathbb{C}) \cap C(\mathbb{R}, \mathbb{C})$ und

$$\exists \omega > 0 \quad \forall \xi \in \mathbb{R} \setminus (-\omega, \omega) : \hat{f}(\xi) = 0.$$

Dann gilt für jedes $T \leq \frac{\pi}{\omega}$:

$$f(x) = \sum_{k=-\infty}^{\infty} f(kT) \operatorname{si} \left((x - kT) \frac{\pi}{T} \right) \quad (x \in \mathbb{R}).$$

Bemerkung: Es gibt Funktionen in $\mathcal{S}$ die bandbeschränkt sind. Da $\mathcal{F} : \mathcal{S} \rightarrow \mathcal{S}$ bijektiv ist, ist z.B. die Funktion aus Beispiel (2) in 25.15 die Fouriertransformierte einer Funktion aus $\mathcal{S}$.

26 Diskrete Fouriertransformation (DFT)

Wird tatsächlich ein Signal *gemessen*, so erhält man in der Regel eine Zahlentabelle die als Vektor im $\mathbb{C}^n$ aufgefasst werden kann. Die *diskrete Fouriertransformation (DFT)* bietet eine Möglichkeit zur Bestimmung und Bearbeitung (digitale Filter) der in einem solchen Signal hauptsächlich vorkommenden Frequenzen.

26.1. Die orthonormale Fourierbasis: Wir versehen $\mathbb{C}^n$ wieder mit dem Standardskalarprodukt, verzichten auf die Pfeile über den Vektoren und schreiben $z = (z_0, z_1, \dots, z_{n-1})$, also

$$(z|w) = \sum_{j=0}^{n-1} z_j \overline{w_j}, \quad \|z\|^2 = (z|z) = \sum_{j=0}^{n-1} |z_j|^2 \quad (z, w \in \mathbb{C}^n).$$

Wir setzen

$$\omega := \exp \left(\frac{2\pi i}{n} \right),$$

und wir definieren die Vektoren $b_0, \dots, b_{n-1} \in \mathbb{C}^n$ durch

$$b_k := \frac{1}{\sqrt{n}} \begin{pmatrix} 1 \\ \omega^k \\ \omega^{2k} \\ \omega^{3k} \\ \vdots \\ \omega^{(n-1)k} \end{pmatrix} = \frac{1}{\sqrt{n}} (\omega^{jk})_{j=0}^{n-1} = \frac{1}{\sqrt{n}} \left(\exp \left(\frac{2\pi i j k}{n} \right) \right)_{j=0}^{n-1}.$$

Diese Vektoren $b_0, \dots, b_{n-1}$ bilden eine Orthonormalbasis von $\mathbb{C}^n$. Diese ONB heißt auch die *orthonormale Fourierbasis* von $\mathbb{C}^n$: Für $k, l \in \{0, \dots, n-1\}$ gilt (beachte $\bar{\omega} = 1/\omega$):

$$(b_k | b_l) = \frac{1}{n} \sum_{j=0}^{n-1} \omega^{jk} \bar{\omega}^{jl} = \frac{1}{n} \sum_{j=0}^{n-1} \frac{\omega^{jk}}{\omega^{jl}} = \frac{1}{n} \sum_{j=0}^{n-1} (\omega^{k-l})^j.$$

Im Fall $k = l$ gilt also $(b_k | b_l) = 1$, und im Fall $k \neq l$ ist $\omega^{k-l} \neq 1$ (beachte $|k-l| \leq n-1$), also

$$(b_k | b_l) = \frac{1}{n} \frac{1 - (\omega^{k-l})^n}{1 - \omega^{k-l}} = \frac{1}{n} \frac{1 - \exp(2\pi i(k-l))}{1 - \omega^{k-l}} = 0.$$

Insbesondere gilt für diese (wie für jede ONB) des $\mathbb{C}^n$:

$$\forall z, w \in \mathbb{C}^n : \quad z = \sum_{k=0}^{n-1} (z | b_k) b_k, \quad (z | w) = \sum_{k=0}^{n-1} (z | b_k) \overline{(w | b_k)}, \quad \|z\|^2 = \sum_{k=0}^{n-1} |(z | b_k)|^2.$$

26.2. Fouriertransformation in $\mathbb{C}^n$:

Definition Für $z \in \mathbb{C}^n$ heißt der Vektor $\widehat{z} \in \mathbb{C}^n$ mit den Komponenten

$$\widehat{z}_k = \sqrt{n}(z | b_k) = \sum_{j=0}^{n-1} z_j \exp\left(-\frac{2\pi i j k}{n}\right) = \sum_{j=0}^{n-1} z_j \bar{\omega}^{jk} \quad (k = 0, \dots, n-1).$$

die *endliche Fouriertransformierte* von z .

Setzt man

$$F = \begin{pmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & \bar{\omega} & \bar{\omega}^2 & \bar{\omega}^3 & \dots & \bar{\omega}^{n-1} \\ 1 & \bar{\omega}^2 & \bar{\omega}^4 & \bar{\omega}^6 & \dots & \bar{\omega}^{2(n-1)} \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & \bar{\omega}^{(n-1)} & \bar{\omega}^{2(n-1)} & \bar{\omega}^{3(n-1)} & \dots & \bar{\omega}^{(n-1)(n-1)} \end{pmatrix},$$

so ist also

$$\forall z \in \mathbb{C}^n : \quad \widehat{z} = Fz.$$

Offensichtlich gilt: F ist symmetrisch, also $F^T = F$. Weiter ist

$$U := \frac{1}{\sqrt{n}} F$$

eine unitäre Matrix: Nach Definition der Orthonormalbasis $b_0, \dots, b_{n-1}$ ist

$$U^* = \frac{1}{\sqrt{n}} F^* = \frac{1}{\sqrt{n}} \bar{F} = (b_0 \ b_1 \ \dots \ b_{n-1}).$$

Somit ist U^* , also auch U unitär. Es gilt also $U^*U = I_n$, d.h. $\frac{1}{n}F^*F = I_n$ und somit

$$F^{-1} = \frac{1}{n}F^* = \frac{1}{n}\overline{F} = \frac{1}{n} \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \omega^3 & \cdots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \omega^6 & \cdots & \omega^{2(n-1)} \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & \omega^{(n-1)} & \omega^{2(n-1)} & \omega^{3(n-1)} & \cdots & \omega^{(n-1)(n-1)} \end{pmatrix}.$$

Die Formel für die *inverse endliche Fouriertransformation* lautet also

$$z_k = \frac{1}{n} \sum_{j=0}^{n-1} \widehat{z}_j \exp\left(\frac{2\pi i j k}{n}\right) = \frac{1}{n} \sum_{j=0}^{n-1} \widehat{z}_j \omega^{jk} \quad (k = 0, \dots, n-1).$$

Kurz zusammengefasst:

endliche Fouriertransformation: $z \mapsto \widehat{z} = Fz$,

endliche Fourierinversionsformel: $z = \frac{1}{n}F^*\widehat{z}$.

Beispiel: Für $n = 4$ ist $\omega = \exp(\pi i/2) = i$, $\overline{\omega} = -i$, also

$$F = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix}, \quad F^{-1} = \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix}.$$

Bemerkung: Der Betrag der Determinante von F lässt sich leicht berechnen:

$$n^n = \det(F^*F) = \det(\overline{F})\det(F) = \overline{\det(F)}\det(F),$$

also $|\det(F)| = n^{n/2}$ und somit $|\det(F^{-1})| = n^{-n/2}$.

In Zusammenhang mit der diskreten Fouriertransformation ist es üblich die Vektoren aus $\mathbb{C}^n$ periodisch zu Folgen auf $\mathbb{Z}$ fortzusetzen. Dann ist $z_{k+n} = z_k$ ($k \in \mathbb{Z}$):

$$z = (\dots, z_0, z_1, \dots, z_{n-1}, z_0, z_1, \dots, z_{n-1}, z_0, z_1, \dots, z_{n-1}, \dots),$$

also z.B. $z_{-1} = z_{n-1}$, $z_{n+1} = z_1$. Wir bezeichnen den zugehörigen Folgenraum mit $\mathbb{C}_{\text{per}}^n$.

Für $z \in \mathbb{C}_{\text{per}}^n$ ist dann $\widehat{z} \in \mathbb{C}_{\text{per}}^n$ ebenfalls periodisch fortgesetzt definiert. Nun ist

$$\widehat{z}_k = \sum_{j=0}^{n-1} z_j \exp\left(-\frac{2\pi i j k}{n}\right) \quad (k \in \mathbb{Z}),$$

und

$$z_k = \frac{1}{n} \sum_{j=0}^{n-1} \widehat{z}_j \exp\left(\frac{2\pi i j k}{n}\right) \quad (k \in \mathbb{Z}).$$

Z.B. gilt in dieser Schreibweise:

$$\widehat{\widehat{z}}_k = \sum_{j=0}^{n-1} \overline{z_j} \exp\left(-\frac{2\pi i j k}{n}\right) = \overline{\sum_{j=0}^{n-1} z_j \exp\left(\frac{2\pi i j k}{n}\right)} = \overline{\widehat{z}_{-k}} \quad (k \in \mathbb{Z}).$$

Ist z reellwertig, so ist also $\widehat{z}_k = \overline{\widehat{z}_{-k}}$ ($k \in \mathbb{Z}$). Ist umgekehrt $z \in \mathbb{C}_{\text{per}}^n$ und gilt $\widehat{z}_k = \overline{\widehat{z}_{-k}}$ ($k \in \mathbb{Z}$), so folgt aus der Inversionsformel

$$\begin{aligned} z_k &= \frac{1}{n} \sum_{j=0}^{n-1} \widehat{z}_j \exp\left(\frac{2\pi i j k}{n}\right) = \frac{1}{n} \sum_{j=0}^{n-1} \overline{\widehat{z}_{-j}} \exp\left(\frac{2\pi i j k}{n}\right) = \overline{\frac{1}{n} \sum_{j=0}^{n-1} \widehat{z}_{-j} \exp\left(-\frac{2\pi i j k}{n}\right)} \\ &= \overline{\frac{1}{n} \sum_{j=0}^{n-1} \widehat{z}_j \exp\left(\frac{2\pi i j k}{n}\right)} = \overline{z_k} \quad (k \in \mathbb{Z}). \end{aligned}$$

Also ist z reellwertig.

Um den Zusammenhang mit Fourierreihen zu verstehen betrachten wir im folgenden T -periodische Funktionen $f: \mathbb{R} \rightarrow \mathbb{C}$ mit $f \in R([-T/2, T/2], \mathbb{C})$. Transformation der Formeln aus 24.3 für 2π -periodische Funktionen führt auf die Fourierkoeffizienten

$$\widehat{f}(k) = \frac{1}{T} \int_{-T/2}^{T/2} f(t) \exp\left(-\frac{2\pi i k t}{T}\right) dt = \frac{1}{T} \int_0^T f(t) \exp\left(-\frac{2\pi i k t}{T}\right) dt,$$

mit der zugehörigen Fourierreihe

$$\sum_{k \in \mathbb{Z}} \widehat{f}(k) \exp\left(\frac{2\pi i k t}{T}\right).$$

Es sei $T_s > 0$ mit $n = T/T_s \in \mathbb{N}$, also $T_s = T/n$. Wir betrachten die Zerlegung

$$0 < T_s < 2T_s < \dots < \left(\frac{T}{T_s} - 1\right) T_s < T$$

von $[0, T]$ und setzen

$$z_j := f(jT_s) = f\left(j\frac{T}{n}\right) \quad (j = 0, \dots, n-1)$$

zu $z \in \mathbb{C}_{\text{per}}^n$ fort.

Dann ist (wg. $1/n = T_s/T$)

$$\frac{1}{n}\hat{z}_k = \frac{1}{n} \sum_{j=0}^{n-1} z_j \exp\left(-\frac{2\pi i j k}{n}\right) = \frac{1}{T} \sum_{j=0}^{n-1} f(jT_s) \exp\left(-\frac{2\pi i k}{T} jT_s\right) T_s \quad (k \in \mathbb{Z}).$$

Die Summe ist eine Riemannsumme für das Integral $\int_0^T f(t) e^{-\frac{2\pi i k t}{T}} dt$. Die Frage ist für welche $k \in \mathbb{Z}$ man $\hat{z}_k/n$ als Approximation von $\hat{f}(k)$ betrachten soll. Wir betrachten dazu den Fall, dass n gerade ist.

Da $\hat{z}$ eine n -periodische Folge ist gilt $\hat{z}_{-k} = \hat{z}_{n-k}$ ($k \in \mathbb{Z}$), also insbesondere

$$\hat{z}_{-1} = \hat{z}_{n-1}, \quad \hat{z}_{-2} = \hat{z}_{n-2}, \quad \dots, \quad \hat{z}_{-n/2} = \hat{z}_{n/2},$$

und speziell für $f(t) = \exp(\frac{2\pi i l t}{T})$ mit $l \in \mathbb{Z}$, $|l| < n/2$ liefert die DFT den richtigen Wert:

$$\frac{\hat{z}_k}{n} = \hat{f}(k) = \delta_{kl} \quad (-n/2 \leq k \leq n/2).$$

Aufgrund dieser Beobachtungen macht man folgende Interpretation:

Die Werte $\hat{z}_k/n$ approximieren die Fourierkoeffizienten $\hat{f}(k)$ mit $|k| \leq n/2$ gut, wenn die $|\hat{f}(k)|$ mit $|k| \geq n/2$ sehr klein sind.

Enthält das abgetastete Signal Frequenzanteile $\hat{f}(k) \neq 0$ mit $|k| \geq n/2$, so werden dadurch die Approximationen von $\hat{f}(k)$ ($|k| \leq n/2$) verfälscht und es kommt zum sogenannten *Alias-Effekt*. Eine technische Lösung ist das Signal dann erst durch einen Tiefpassfilter zu schicken (oder n größer zu wählen).

Wählt man also eine Schrittweite von T_s Sekunden mit $n = T/T_s \in \mathbb{N}$, also eine Abtastfrequenz von $f_s = 1/T_s$ Hertz so approximiert $\hat{z}_k/n$ den Fourierkoeffizienten $\hat{f}(k)$ zur Frequenz

$$\frac{k}{T} = \frac{k}{nT_s} = \frac{k}{n} f_s \quad (|k| \leq n/2).$$

Ist ω die höchste im Signal vorkommende Frequenz, so sollte also $f_s > 2\omega$ sein.

Bemerkung: Man sieht auch, dass mit T die Frequenzauflösung wächst.

In vielen Anwendungen ist man an der relativen Magnitude der vorhandenen Frequenzanteile interessiert und betrachtet daher einfach $|\hat{z}_k|$ ($|k| \leq n/2$). Ist f reellwertig, so ist dann $|\hat{z}_k| = |\overline{\hat{z}_{-k}}| = |\hat{z}_{-k}|$.

Beispiele: (1) Wir betrachten das sinusförmige Signal

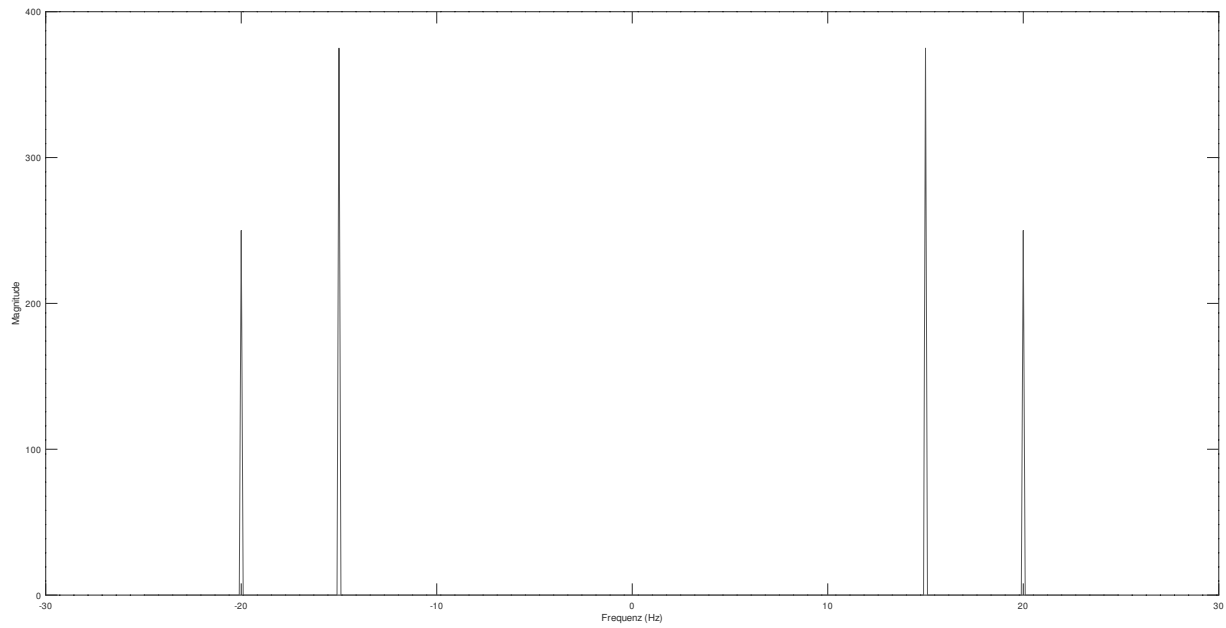
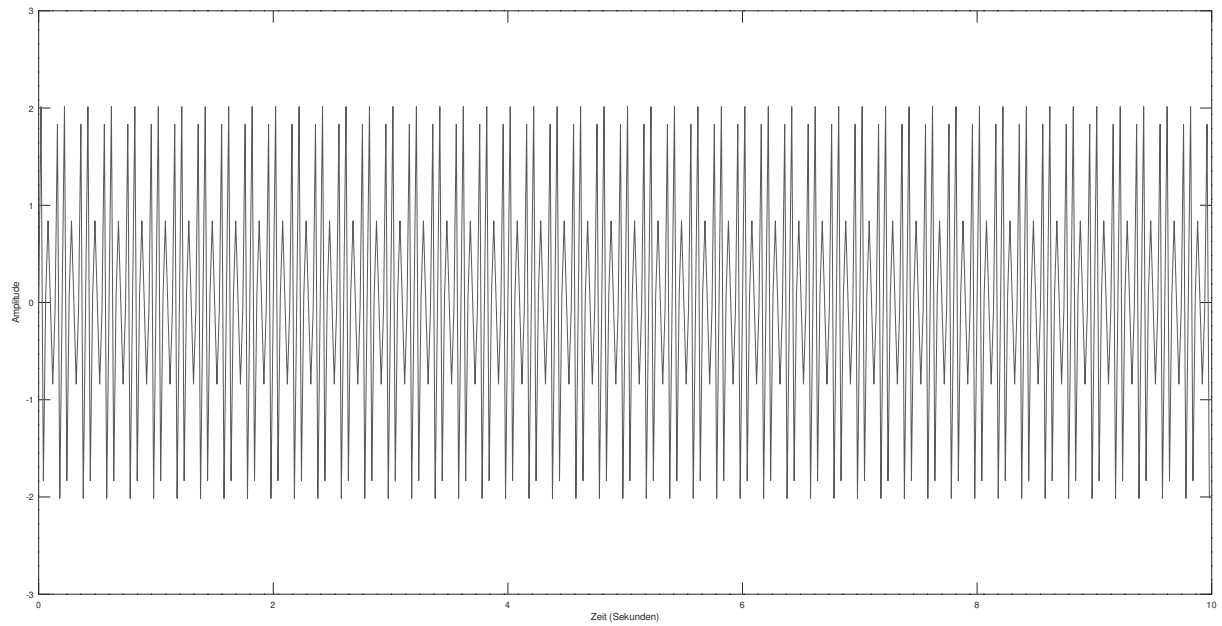
$$f(t) = \frac{3}{2} \sin(15(2\pi t)) + \sin(20(2\pi t)) \quad (t \in \mathbb{R})$$

das eine Funktion der Zeit t mit Frequenzkomponenten von 15Hz und 20Hz ist und verwenden einen Zeitvektor, der in Schritten von $T_s = 1/50$ Sekunden über einen Zeitraum von $T = 10$ Sekunden das Signal abgetastet. Hier ist also $n = 500$, $n/2 = 250$ was Frequenzschritten von $f_s/n = \frac{50}{500} \text{ Hz} = \frac{1}{10} \text{ Hz}$ im Bereich 0Hz - 25Hz entspricht.

GNU Octave bzw. MATLAB code für ein file dft.m:

```
function ans=dft
Ts = 1/50;
fs=1/Ts;
t = 0:Ts:10-Ts;
x = 1.5*sin(2*pi*15*t) + sin(2*pi*20*t);
plot(t,x)
xlabel('Zeit (Sekunden)')
ylabel('Amplitude')
pause(5)
y=fft(x);
n = length(x);
fshift = (-n/2:n/2-1)*(fs/n);
yshift = fftshift(y);
plot(fshift,abs(yshift));
xlabel('Frequenz (Hz)')
ylabel('Magnitude')
end
```

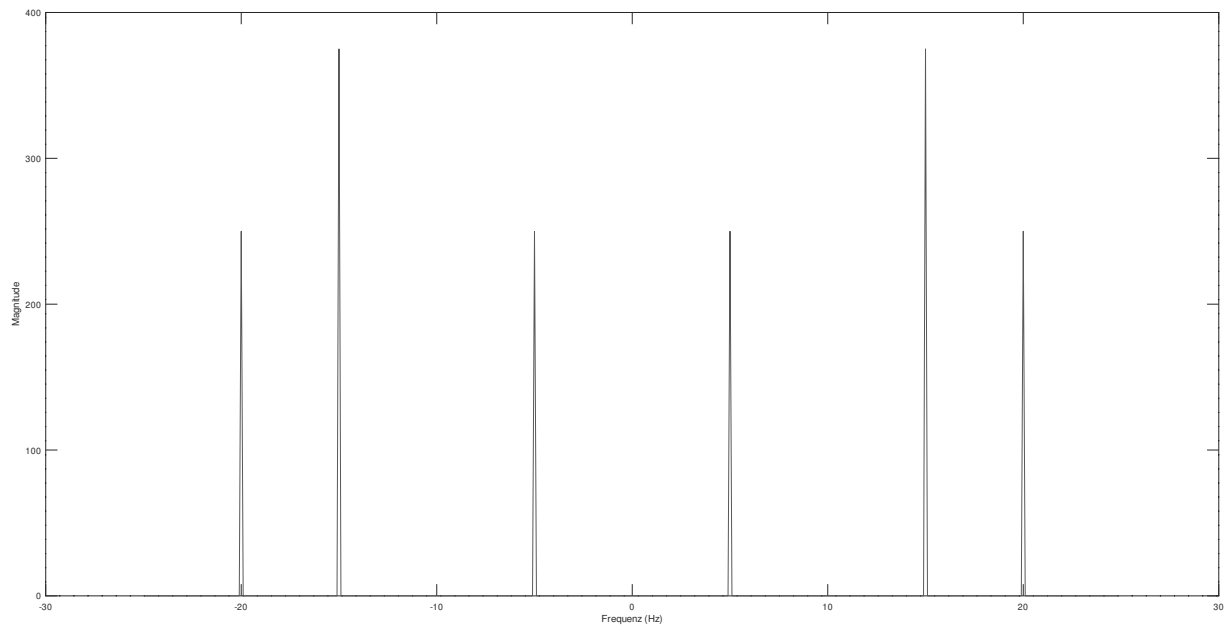
Folgende Bilder zeigen das Signal und das transformierte Signal als Funktion der Frequenz.



(2) Transformiert man hingegen

$$f(t) = \frac{3}{2} \sin(15(2\pi t)) + \sin(20(2\pi t)) + \sin(45(2\pi t)) \quad (t \in \mathbb{R})$$

mit derselben Abtastfrequenz f_s , so sieht man den Alias Effekt:



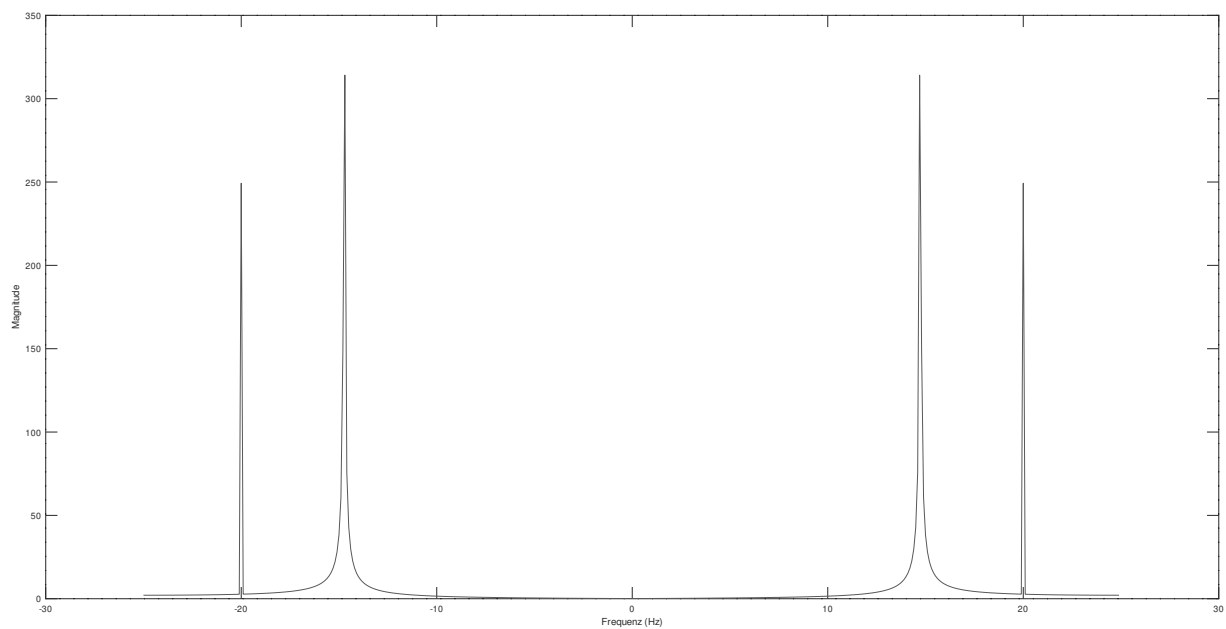
Die Frequenz 5Hz kommt im Signal f in Wirklichkeit nicht vor und die Frequenz 45Hz kann nicht angezeigt werden weil

$$50\text{Hz} = f_s < 2\omega = 90\text{Hz}.$$

(3) Transformiert man mit demselben Programm die Funktion

$$f(t) = \frac{3}{2} \sin((13 + \sqrt{3})(2\pi t)) + \sin(20(2\pi t)) \quad (t \in \mathbb{R})$$

so ist der Abtastbereich keine Periode der Funktion und es entsteht der sogenannte *Leck-Effekt*: Das Signal enthält Frequenzen, die nicht zu den von der DFT berechneten diskreten Frequenzen gehören. Die DFT nähert diese Frequenzen durch die benachbarten Frequenzen an.



Der Leck-Effekt tritt bei Anwendung der DFT auf “natürliche” Signale praktisch immer auf.

(4) Folgende Bilder zeigen eine Unterwasseraufnahme eines Blauwalgesangs und das durch n dividierte transformierte Signal als Funktion der Frequenz. Hier ist $n = 2^{17} = 131072$.

