Höhere Mathematik II für die Fachrichtung Elektrotechnik und Informationstechnik Sommersemester 2015

PD Dr. Peer Christian Kunstmann
Karlsruher Institut für Technologie (KIT)
Institut für Analysis

Das Skript wird von dem Dozenten
der Vorlesung geändert
Dozent: Ioannis Anapolitanos
e-mail: ioannis.anapolitanos@kit.edu

Contents

Innerhalb der Veranstaltung *Höhere Mathematik II für die Fachrichtung elektrotechnik und Informationstechnik inklusive Komplexe Analysis und Integraltransformationen* wird der Teil zu *Höhere Mathematik II* jeweils in den Vorlesungen am Dienstag und Donnerstag behandelt.

Dies ist eine Vorlesungszusammenfassung, gedacht zur Vorlesungsbegleitung und als Gedächtnisstütze, nicht jedoch als etwas, das für sich selbst stehen könnte (wie etwa ein Lehrbuch). Der Besuch der Vorlesung ist durch die Lektüre in keinem Fall zu ersetzen, es gibt dort noch viel mehr an mündlichen Erklärungen, Erläuterungen und veranschaulichenden Skizzen, die für Verständnis und Einordnung des präsentierten Stoffes unabdingbar sind.

14 Mehr zu Matrizen und linearen Abbildungen

14.1. Das Produkt von Matrizen: Sie $\mathbb{K} = \mathbb{C}$ oder $\mathbb{R}$. Seien $n, m, q \in \mathbb{N}$ und $A \in \mathbb{K}^{n\times m}$, $B \in \mathbb{K}^{m\times q}$ Matrizen mit $A = (a_{jk})_{j=1}^{n}{}_{k=1}^{m}$ und $B = (b_{kl})_{k=1}^{m}{}_{l=1}^{q}$. Das <u>Matrixprodukt</u> $AB \in \mathbb{K}^{n\times q}$, ist die Matrix $C = (c_{jl})_{j=1}^{n}{}_{l=1}^{q}$ mit

$$c_{jl} = \sum_{k=1}^{m} a_{jk} b_{kl}, \qquad j = 1, \ldots, n,\ l = 1, \ldots, q.$$

Man beachte, dass hierbei die Anzahl der Spalten von A gleich der Anzahl der Zeilen von B ist (nämlich m), andernfalls ist das Matrixprodukt **nicht definiert**.

Bemerkung: Man erhält die l-te Spalte von AB, indem man die Matrix A mit der l-ten Spalte von B multipliziert (im Sinne des Matrix-Vektor-Produktes).

Beispiel mit $n = m = 2$:

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

Eigenschaften des Produktes von Matrizen: Für alle $A_1, A_2 \in \mathbb{K}^{n\times m}$, $B_1, B_2 \in \mathbb{K}^{m\times q}$, $C \in \mathbb{K}^{q\times r}$ und alle $\alpha \in \mathbb{K}$ gilt:

$$(A_1+\alpha A_2)B_1 = A_1B_1+\alpha A_2B_1, \quad A_1(B_1+\alpha B_2) = A_1B_1+\alpha A_1B_2, \quad (A_1B_1)C = A_1(B_1C).$$

Warnung: Im allgemeinen gilt $AB \neq BA$! Damit beide Produkte existieren, muss zunächst $n = q$ sein. Dann ist AB eine $n \times n$-Matrix und BA eine $m \times m$-Matrix. Gleichheit kann also höchstens für $n = m = q$ gelten. Für $n = m = q = 2$ ist aber z.B.

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

14.2. Invertierbare Matrizen: Für $n \in \mathbb{N}$ heißt $I_n := (\delta_{jk})_{j,k=1}^{n} \in \mathbb{K}^{n\times n}$ die <u>Einheitsmatrix</u> (wir schreiben auch kurz I, wenn n aus dem Zusammenhang klar ist).

Es gilt dann

$$IA = AI = A \quad \text{für alle } A \in \mathbb{K}^{n\times n}$$

Definition: Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt <u>regulär</u>, falls Rang $A = n$ ist, und <u>invertierbar</u>, falls es eine Matrix $B \in \mathbb{K}^{n\times n}$ gibt mit $AB = BA = I$.

Bemerkung und Definition: Ist A invertierbar, so ist die Matrix B in obiger Definition eindeutig bestimmt, denn für eine weitere Matrix $\tilde{B}$ mit diesen Eigenschaften folgt:

$$B = BI = B(A\tilde{B}) = (BA)\tilde{B} = I\tilde{B} = \tilde{B}. \qquad (*)$$

Diese Matrix B heißt Inverse von A (oder zu A inverse Matrix) und wird mit A^{-1} bezeichnet.

Rechenregeln: Seien $A, B \in \mathbb{K}^{n\times n}$ invertierbar. Dann sind A^{-1} und AB invertierbar, und es gilt:

$$(A^{-1})^{-1} = A \qquad \text{und} \qquad (AB)^{-1} = B^{-1}A^{-1}.$$

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

$$A \text{ regulär} \iff \operatorname{Bild} A = \mathbb{K}^n \iff \operatorname{Kern} A = \{\vec{0}\} \iff A \text{ invertierbar}.$$

Für die induzierte lineare Abbildung $\phi_A : \mathbb{K}^n \to \mathbb{K}^n$ mit $\phi_A(\vec{x}) = A\vec{x}$ gilt:

$$\phi_A \text{ surjektiv} \iff \phi_A \text{ injektiv} \iff \phi_A \text{ bijektiv}.$$

Bemerkung: Sind $A, B \in \mathbb{K}^{n\times n}$ mit $AB = I$, so sind A und B invertierbar und es gilt $A^{-1} = B$. So ist etwa $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)^{-1} = \left(\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right)$ (vgl. Beispiel in 14.2).

Beweis. Gilt $AB = I$, so ist $\operatorname{Kern} B = \{\vec{0}\}$ und $\operatorname{Bild} A = \mathbb{K}^n$, und nach dem Satz sind A, B invertierbar. Für $A^{-1} = B$ verwenden wir ein Argument wie in (*). □

Berechnung von A^{-1}: Man bringt die Matrix $(A|I)$ auf Zeilennormalform und bekommt $(I|A^{-1})$.

Beispiel: Sei $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbb{K}^{2\times 2}$ invertierbar und $a \neq 0$. Wir setzen $\delta := ad - bc$ und sehen an der dritten Matrix, dass $\delta \neq 0$ ist.

$$\left(\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right) \quad \left(\begin{array}{cc|cc} 1 & b/a & 1/a & 0 \\ 0 & d - bc/a & -c/a & 1 \end{array}\right) \quad \left(\begin{array}{cc|cc} 1 & b/a & 1/a & 0 \\ 0 & \delta/a & -c/a & 1 \end{array}\right)$$
$$\left(\begin{array}{cc|cc} 1 & b/a & 1/a & 0 \\ 0 & 1 & -c/\delta & a/\delta \end{array}\right) \quad \left(\begin{array}{cc|cc} 1 & 0 & \frac{1}{a} + \frac{bc}{\delta a} & -b/\delta \\ 0 & 1 & -c/\delta & a/\delta \end{array}\right).$$

Wegen $\frac{1}{a} + \frac{bc}{\delta a} = d/\delta$ ist also

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{\delta}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Den Fall $c \neq 0$ behandelt man analog, er führt auf dieselbe Formel. Wir sehen auch, dass Invertierbarkeit äquivalent ist zu $\delta \neq 0$.

14.3. Lineare Abbildungen als Matrizen: Seien V und W $\mathbb{K}$-Vektorräume von Dimension m bzw. n und $\phi : V \to W$ eine lineare Abbildung. In V und W seien geordnete[1]

[1] das bedeutet, dass die Reihenfolge der Elemente der Basis wichtig ist

Basen $B = (\vec{b_1}, \vec{b_2}, \ldots, \vec{b_m})$ bzw. $C = (\vec{c_1}, \vec{c_2}, \ldots, \vec{c_n})$ gegeben. Dann gibt es genau eine Matrix bezeichnet durch ${}_C[\phi]_B \in \mathbb{K}^{n\times m}$ mit der Eigenschaft

$$\phi\Big(\sum_{k=1}^{m} \alpha_k \vec{b_k}\Big) = \sum_{j=1}^{n} \beta_j \vec{c_j} \iff {}_C[\phi]_B \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}.$$

Durch die Matrix ${}_C[\phi]_B$ werden also die Koordinatentupel $(\alpha_1, \alpha_2, \ldots, \alpha_m) \in \mathbb{K}^m$ von Vektoren $v \in V$ bzgl. der Basis B auf die Koordinatentupel $(\beta_1, \beta_2, \ldots, \beta_n) \in \mathbb{K}^n$ von $\phi(v)$ bzgl. der Basis C abgebildet. ${}_C[\phi]_B$ heißt dann Matrix zu ϕ bzgl. der Basen B und C.

Man erhält die Matrix ${}_C[\phi]_B$ wie folgt:

> Zu jedem Basisvektor b_k, $k = 1, \ldots, m$ bestimme man die Koordinaten von $\phi(b_k) \in W$ bzgl. der Basis $c_1, c_2, \ldots, c_n$. Diese bilden die k-te Spalte von A.

Beispiele:(1) Die lineare Abbildung $\phi : \mathbb{C}^3 \to \mathbb{C}^3$ sei gegeben durch

$$\phi(x_1, x_2, x_3) = (2x_1 - x_2 + x_3, -x_1 + 2x_2 + x_3, x_1 - x_2 + 2x_3).$$

Wir bestimmen die Matrix ${}_B[\phi]_B$, wobei $B = (\vec{b}_1, \vec{b}_2, \vec{b}_3)$ mit

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

Es ist

$$\begin{aligned}
\phi(\vec{b}_1) &= \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = 1 \cdot \vec{b}_1 + 0 \cdot \vec{b}_2 + 0 \cdot \vec{b}_3, \\
\phi(\vec{b}_2) &= \begin{pmatrix} 3 \\ 0 \\ 3 \end{pmatrix} = 0 \cdot \vec{b}_1 + 3 \cdot \vec{b}_2 + 0 \cdot \vec{b}_3, \\
\phi(\vec{b}_3) &= \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} = 1 \cdot \vec{b}_1 + (-1) \cdot \vec{b}_2 + 2 \cdot \vec{b}_3,
\end{aligned}$$

also

$${}_B[\phi]_B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & -1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Sei S die Standardbasis von $\mathbb{C}^3$. Dann mit ähnlichen Berechnungen bekommt man

$${}_S[\phi]_B = \begin{pmatrix} 1 & 3 & 0 \\ 1 & 0 & 3 \\ 0 & 3 & 1 \end{pmatrix} \qquad {}_S[\phi]_S = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & -1 & 2 \end{pmatrix}.$$

(2) Sei Id : $\mathbb{C}^3 \to \mathbb{C}^3$, $\vec{x} \mapsto \vec{x}$. Wie betrachten die Basen S, B des vorherigen Beispiels. Dann

$$_B[Id]_B = I, \quad _S[Id]_S = I, \quad _S[Id]_B := \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Da

$$\vec{e}_1 = \frac{1}{2}(\vec{b}_1 + \vec{b}_2 - \vec{b}_3), \quad \vec{e}_2 = \frac{1}{2}(\vec{b}_1 - \vec{b}_2 + \vec{b}_3), \quad \vec{e}_3 = \frac{1}{2}(-\vec{b}_1 + \vec{b}_2 + \vec{b}_3),$$

bekommen wir

$$_B[Id]_S := \frac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix}.$$

Der folgende Satz sagt in einfachen Worten, dass die Komposition von zwei linearen Abbildungen mit Produkt von Matrizen dargestellt werden kann.

Satz: Seien U, V, Z $\mathbb{K}$-Vektorräume und B (bzw. C, bzw. D) eine Basis in U (bzw. V, bzw. C). Seien $\phi : U \to V$, $\psi : V \to Z$ lineare Abbildungen. Dann

$$_D[\psi \circ \phi]_B =_D [\psi]_C \ _C[\phi]_B.$$

Beispiel: Wir betrachten die gleichen Abbildungen und Basen wie in den Beispielen (1), (2). Man kann direkt die Matrix $_S[\phi]_B$ mit Hilfe der Matrizen $_S[Id]_B$,$_B\,[\phi]_B$ berechnen und zwar

$$_S[\phi]_B = {_S[Id \circ \phi]_B} = {_S[Id]_B}{_B[\phi]_B} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & -1 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 0 \\ 1 & 0 & 3 \\ 0 & 3 & 1 \end{pmatrix}.$$

15 Skalarprodukt und Orthogonalität

Wie vorher schreiben wir wieder $\mathbb{K}$ für $\mathbb{R}$ oder $\mathbb{C}$.

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

(S1) $\forall x, y \in V$: $(x|y) = \overline{(y|x)}$,

(S2) $\forall x, y, z \in V$, $\alpha \in \mathbb{K}$: $(\alpha x + y|z) = \alpha(x|z) + (y|z)$,

(S3) $\forall x \in V \setminus \{0\}$: $(x|x) > 0$.

heißt ein <u>Skalarprodukt auf V</u>.

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

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

Zum Beweis der Cauchy-Schwarz-Ungleichung für $y \neq 0$ schreibe man

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

wobei $\alpha = (x|y)/(y|y)$. Wegen (S3) gilt Gleichheit in der Cauchy-Schwarz-Ungleichung nur, wenn $x = \alpha y$, dh 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 $\overline{r} = r$ für alle $r \in \mathbb{R}$.

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

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

Für das Skalarprodukt von $\vec{x}, \vec{y} \in \mathbb{R}^n$ schreibt man häufig auch $\vec{x} \cdot \vec{y}$ (und gelegentlich sogar nur $\vec{x}\vec{y}$).

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

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

ein Skalarprodukt auf $\mathbb{K}^n$.

(3) Ist $V = C([a, b], \mathbb{C})$ (hier ist $\mathbb{K} = \mathbb{C}$), so wird durch

$$(f|g) := \int_a^b f(t)\overline{g(t)}\, dt, \quad f, g \in C[a, b],$$

ein Skalarprodukt auf $C([a, b], \mathbb{C})$ definiert. Die Eigenschaften (S1) und (S2) sind leicht. Zum Nachweis von (S3) stellt man fest, dass für $h \in C[a, b]$ mit $h \geq 0$ und $\int_a^b h(t)\, dt = 0$ zunächst folgt $\int_a^x h(t)\, dt$ für alle $x \in [a, b]$, und dann $h = 0$ nach dem Hauptsatz. Diese Beobachtung wendet man auf $h := |f|^2$ an.

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

(N1) $\forall v \in V$: $\|v\| = 0 \Rightarrow v = 0$, (N2) $\forall v \in V,\ \alpha \in \mathbb{K}$: $\|\alpha v\| = |\alpha|\|v\|$,

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

Beweis der Dreiecksungleichung. Es gilt (unter Verwendung der Cauchy-Schwarz-Ungleichung):

$$\|u+v\|^2 = (u+v|u+v) = \|u\|^2 + 2\mathrm{Re}\,(u|v) + \|v\|^2 \leq \|u\|^2 + 2|(u|v)| + \|v\|^2 \leq (\|u\| + \|v\|)^2,$$

und die Ungleichung folgt durch Wurzelziehen. □

Definition: Sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung $\|\cdot\| : V \to [0, \infty)$ 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, und es gilt $\|0\| = 0$, sowie

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

Beispiel: 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.

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

$$(\vec{x}|\vec{y}) = \|\vec{x}\|\|\vec{y}\| \cos\varphi,$$

wobei φ den von $\vec{x}, \vec{y}$ eingeschlossenen **Winkel** bezeichne. Da das Skalarprodukt in jeder Komoponente 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

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

15.3. Orthogonalität: Sei V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt. Vektoren $v_1, v_2, \ldots, v_m \in V$ heißen orthogonal, falls für alle $j, k \in \{1, 2 \ldots, m\}$ mit $j \neq k$ gilt $(v_j|v_k) = 0$. Statt $(v|w) = 0$ schreibt man auch $v \perp w$.

Die Vektoren $v_1, v_2, \ldots, v_m$ heißen orthonormal oder ein Orthonormalsystem (ONS), falls für alle j, k gilt: $(v_j|v_k) = \delta_{jk}$, dh 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.

Beispiel: Die Standardbasis $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ ist eine Orthonormalbasis von $\mathbb{K}^n$, denn es gilt $(\vec{e}_j|\vec{e}_k) = \delta_{jk}$ für alle $j, k = 1, \ldots, n$.

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

Beweis. Seien $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{K}$ mit

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = 0.$$

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

$$0 = (0|v_k) = (\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n|v_k) = \alpha_k (v_k|v_k)$$

wegen der vorausgesetzten Orthogonalität. Wegen $v_k \neq 0$ ist auch $(v_k|v_k) \neq 0$, und wir erhalten $\alpha_k = 0$. Da k beliebig war, sind $v_1, v_2, \ldots, v_n$ linear unabhängig. $\square$

Bemerkung: Ist $v_1, v_2, \ldots, v_m$ ein Orthonormalsystem in V und $v \in \operatorname{lin}\{v_1, v_2, \ldots, v_m\}$, so lassen sich die Koordinaten von v bzgl. $v_1, v_2, \ldots, v_m$ 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, $k = 1, \ldots, m$:

$$(v|v_k) = \sum_{j=1}^{m} \alpha_j \underbrace{(v_j|v_k)}_{=\delta_{jk}} = \alpha_k.$$

Beispiel: Sei $e_k : [-\pi, \pi] \to \mathbb{C}$, $e_k(t) := e^{ikt}$. In $V = C([-\pi, \pi], \mathbb{C})$ bilden für jedes $n \in \mathbb{N}$ die Funktionen $(e_k)_{|k| \leq n}$ ein Orthonormalsystem bezüglich des durch

$$(f|g) := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \overline{g(t)} \, dt$$

auf V definierten Skalarprodukts, denn für $k, l \in \mathbb{Z}$ gilt

$$(e_k|e_l) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ikt} e^{-ilt} \, dt = \begin{cases} 1 & , k = l \\ 0 & , k \neq l \end{cases}.$$

Für jedes trigonometrische Polynom f vom Grad $\leq n$ gilt

$$f(t) = \sum_{|k| \leq n} (f|e_k) e_k(t) = \sum_{|k| \leq n} \hat{f}(k) e_k(t) = \sum_{|k| \leq n} \hat{f}(k) e^{ikt}, \quad t \in [-\pi, \pi].$$

Die Koordinaten sind hier also gerade die Fourierkoeffizienten.

Wenn wir also im folgenden Fourierreihen betrachten, beziehen wir uns immer auf das hier definierte Skalarprodukt. Die zugehörige Norm ist gegeben durch

$$\|f\| = \sqrt{(f|f)} = \Big(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(t)|^2\Big)^{1/2}.$$

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

$$\mathrm{lin}\{v_1, v_2, \ldots, v_m\} = \mathrm{lin}\{b_1, b_2, \ldots, b_m\}.$$

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

Die Vektoren $b_1, \ldots, b_m$ werden sukzessiv so konstruiert, dass gilt:

$$\mathrm{lin}\{v_1, v_2, \ldots, v_k\} = \mathrm{lin}\{b_1, b_2, \ldots, b_k\} \quad \text{für alle } k = 1, \ldots, m.$$

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

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

oder gleichbedeutend

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

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

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

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

15.5. 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 <u>transponierte Matrix zu A</u> und wird mit A^T bezeichnet. Für alle $k \in \{1, \ldots, n\}$, $j \in \{1, \ldots, m\}$ 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

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

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

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

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

Schreibweisen des Skalarprodukts:

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 erklärt 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

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

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. □

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

$$(A\vec{x}|A\vec{y}) = (\vec{x}|\vec{y}) \quad \text{für alle } \vec{x}, \vec{y} \in \mathbb{K}^n.$$

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

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

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

$$A^T = A \text{ falls } A \in \mathbb{R}^{n \times n} \text{ orthogonal}, A^* = A^{-1} \text{ falls } A \in \mathbb{C}^{n \times n} \text{ unitär}.$$

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 $\theta \in \mathbb{R}$: $A = \left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$.
3) Im $\mathbb{R}^3$ Rotation um die z-Achse bei Spiegelung an der (x, y)-Ebene:

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

15.7. Orthogonalprojektionen: Sei V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt, $b_1, b_2, \dots, b_m$ ein Orthonormalsystem in V und $U := \operatorname{lin}\{b_1, b_2, \dots, b_m\}$. Die lineare Abbildung

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

hat folgende Eigenschaften:

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

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

$$\|v - Pv\| = \min\{\|v - u\| : u \in U\} \quad \text{für jedes } v \in V,$$

dh Pv ist die (eindeutig bestimmte) Bestapproximation von v in U. Die Abbildung P heißt Orthogonalprojektion von V auf U.

Ende Do 17.04.14

Beweis. Klar ist $\operatorname{Bild} P \subseteq U$. Nach der Bemerkung in 15.3 gilt $Pu = u$ für alle $u \in U$, also ist $P \circ P = P$ und $\operatorname{Bild} P = U$. Für jedes $v \in V$ gilt:

$$Pv = 0 \iff (v|b_j) = 0 \text{ für alle } j \iff (v|u) = 0 \text{ für alle } u \in U.$$

Außerdem ist für $k = 1, \dots, m$:

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

und folglich auch $(v - Pv|u) = 0$ für alle $u \in U$. Schließlich gilt für $u \in U$:

$$\begin{aligned}\|v-u\|^2 &= \|v - Pv + Pv - u\|^2 = \|v - Pv\|^2 + \|Pv - u\|^2 + 2\mathrm{Re}\,(v - Pv|Pv - u)\\ &= \|v - Pv\|^2 + \|Pv - u\|^2.\end{aligned}$$

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

Bemerkung: Im Gram-Schmidt-Verfahren in 15.4 hat man im k-ten Schritt (für $k = 2, \ldots, m$) und für $U = \mathrm{lin}\{b_1, b_2, \ldots, b_{k-1}\}$, dass $c_k = v_k - Pv_k$ und somit $c_k \perp u$ für jedes $u \in U$ wie gewünscht.

Beispiel: Sei $n \in \mathbb{N}$. Für eine gegebene 2π-periodische Funkton f ist die Funktion

$$t \mapsto \sum_{|k| \le n} (f|e_k)e_k(t) = \sum_{|k|\le n} \hat{f}(k)e^{ikt}$$

die Bestapproximation von f in $U_n := \mathrm{lin}\{e_k : |k| \le n\}$, also $(Pf)(t) = \sum_{|k|\le n} \hat{f}(k)e^{ikt}$.

15.8. (nicht HM2 Prüfungsrelevant) Besselsche Ungleichung und Parsevalsche Gleichung: Wir setzen das Beispiel von eben fort: wegen $(f - Pf) \perp Pf$ (vgl. 15.7) gilt

$$\|f\|^2 = \|f - Pf + Pf\|^2 = \|f - Pf\|^2 + \|Pf\|^2 = \|f - Pf\|^2 + \sum_{|k|\le n} |(f|e_k)|^2.$$

Daraus folgt zum einen die Besselsche Ungleichung

$$\sum_{k=-\infty}^{\infty} |(f|e_k)|^2 \le \|f\|^2,$$

die analog für jedes durch $\mathbb{N}$ oder $\mathbb{Z}$ indizierte Orthonormalsystem gilt. Bezeichnen wir die Orthogonalprojektion auf U_n mit P_n, so sieht man zum anderen, dass $\lim_{n\to\infty} \|f - P_n f\| = 0$ äquivalent ist zur Parsevalschen Gleichung

$$\|f\|^2 = \sum_{k=-\infty}^{\infty} |(f|e_k)|^2.$$

Ein Orthonormalsystem heißt vollständig in V, wenn dies für alle $f \in V$ gilt.

Bemerkung: Das Orthonormalsystem $(e_k)_{k\in\mathbb{Z}}$ ist vollständig im Raum der stückweise stetigen und normalisierten Funktionen $f : [-\pi, \pi) \to \mathbb{C}$, dh für alle diese Funktionen gilt:

$$\|f\|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(t)|^2\, dt = \sum_{k=-\infty}^{\infty} |\hat{f}(k)|^2$$

und

$$\|f - \sum_{|k|\le n} \hat{f}(k)e_k\| \to 0 \ (n \to \infty).$$

Letzteres bezeichnet man als Konvergenz im quadratischen Mittel und schreibt dann auch $f = \sum_{k=-\infty}^{\infty} \hat{f}(k)e_k$.

16 Determinanten und Kreuzprodukt

16.1. Definierende Eigenschaften der Determinante: Die Determinante ist eine Abbildung $\det : \underbrace{\mathbb{K}^n \times \mathbb{K}^n \times \ldots \times \mathbb{K}^n}_{n} \to \mathbb{K}$ mit den Eigenschaften

(D1) $\det(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n) = 1$,
(D2) für alle $j \in \{1, \ldots, n\}$ und $\vec{a}_1, \ldots, \vec{a}_n, \vec{b}_j \in \mathbb{K}^n$ gilt

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

(D3) Wenn wir zwei Spalten vertauschen, dann ändert sich das Vorzeichen der Determinante.

Bemerkung: Durch die Eigenschaften (D1)–(D3) ist die Determinante $\det$ eindeutig bestimmt.

Bemerkung: Man kann auch die definierenden Eigenschaften mit Hilfe der Zeilen anstatt der Spalten definieren.

(D1) bedeutet eine (naheliegende) Normierung. (D2) bedeutet, dass die Determinante in jeder Spalte linear ist.

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

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

Wir betrachten im folgenden $\det$ meist als Funktion auf $\mathbb{K}^{n\times n}$.

16.2. Folgerungen: (a) Ist eine Spalte $= 0$, 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) Hat die Matrix zwei gleiche Spalten, dann hat sie Determinante Null.

(d) Sind die Spalten von $A \in \mathbb{K}^{n\times n}$ linear abhängig (dh gilt $\operatorname{Rang} A < n$), so ist $\det(A) = 0$.

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

Alle obige Folgerungen gelten wenn Spalten durch Zeilen ersetzt werden.

Erinnerung: Eine Matrix $\in \mathbb{K}^{n\times n}$ heißt regulär, falls sie invertierbar ist, bzw. falls die zugehörige lineare Abbildung $\mathbb{K}^n \to \mathbb{K}^n, x \mapsto Ax$ bijektiv (oder injektiv oder surjektiv) ist (vgl. 14.21 in HM I).
Der Rang einer Matrix ist die Maximalzahl linear unabhängiger Spalten (oder die Maximalzahl linear unabhängigier 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:

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

Wegen (d) muss man bei (e) nur noch zeigen: 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) $\det(A)$ nur mit Zahlen $\neq 0$ multipliziert. Wegen (D1) ist schließlich $\det(I_n) = 1 \neq 0$. □

16.3. Der Fall $n = 2$: Es gilt

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

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

Beispiele: (1) $\begin{vmatrix} 2 & -3 \\ -4 & 6 \end{vmatrix} = 2 \cdot 6 - (-3)(-4) = 0$, 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.

16.4. 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) braucht eine Fallunterscheidung, die wir unten im allgemeinen Fall durchführen.

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

Schema:

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

Beispiel:

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

16.5. Der allgemeine Fall: Gegeben sei die Matrix $A \in \mathbb{K}^{n \times n}$ mit den Einträgen a_{jk}. 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.

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 16.4, für $n = 2$ diejenige aus 16.3.

Beweis. (D1) ist klar, und (D2) ist nicht so schwer. Zum Beweis von (D3) seien zwei Spalten von A gleich, etwa die k_0-te und die k_1-te, wobei $k_0 < k_1$. In der Summe verschwindet dann $\det(A_{1k})$ für alle $k \notin \{k_0, k_1\}$, da in diesen A_{1k} zwei Spalten gleich sind (weder die k_0-te noch die k_1-te sind gestrichen worden). Also ist

$$\det(A) = (-1)^{k_0+1} a_{1k_0} \det(A_{1k_0}) + (-1)^{k_1+1} a_{1k_1} \det(A_{1k_1}).$$

Hierbei ist $a_{1k_0} = a_{1k_1}$ nach Voraussetzung. Wir erhalten A_{1k_1} aus A_{1k_0}, indem wir durch sukzessives Vertauschen benachbarter Spalten die $k_1 - 1$-te Spalte (von A_{1k_0}) an die k_0-te Stelle bringen. Dazu brauchen wir $k_1 - 1 - k_0 = k_1 - k_0 - 1$ Vertauschungen. Also ist

$$\det(A_{1k_1}) = (-1)^{k_1 - k_0 - 1} \det(A_{1k_0}).$$

$\square$

Beispiel: Ist $A \in \mathbb{K}^{n\times n}$ von der Form $\begin{pmatrix} d_1 & 0 & 0 & \cdots & 0 \\ * & d_2 & 0 & \cdots & 0 \\ * & * & \ddots & \ddots & 0 \\ * & * & \cdots & d_{n-1} & 0 \\ * & * & \cdots & * & d_n \end{pmatrix}$, also eine untere Dreiecksmatrix, so gilt $\det(A) = d_1 \cdot d_2 \cdot \ldots \cdot d_n$.

16.6. Determinantenentwicklungssatz: Die Formel in 16.5 nennt man Entwicklung von $\det(A)$ nach der ersten Zeile von $A = (a_{jk})_{jk}$. Mit denselben Argumenten kann man $\det(A)$ nach einer beliebigen Zeile oder Spalte entwickeln:

Für jedes $l \in \{1, \ldots, n\}$ 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:

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

16.7. Das Signum einer Permutation (nicht HM2 Prüfungsrelevant): Eine Permutation ist eine bijektive Abbildung $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$. Die Menge S_n aller Permutationen von $\{1, \ldots, n\}$ hat genau $n!$ Elemente.

Eine Permutation $\sigma \in S_n$ schreibt man zweckmäßigerweise $\left(\begin{smallmatrix} 1 & 2 & & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{smallmatrix}\right)$ oder auch nur $(\sigma(1), \sigma(2), \ldots, \sigma(n))$.

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

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

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

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

(b) Für $\sigma, \tau \in S_n$ gilt $\operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)$. Das folgt aus

$$\frac{\sigma(\tau(j)) - \sigma(\tau(i))}{j-i} = \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \cdot \frac{\tau(j) - \tau(i)}{j-i}.$$

Beispiele: (1) Vertauscht $\tau \in S_n$ gerade zwei bestimmte Zahlen aus $\{1, \dots, n\}$ (eine solche Permutation heißt *Transposition*), so gilt $\operatorname{sgn}\tau = -1$, denn: Vertauscht τ etwa die Zahlen $i_0 < i_1$, so sind die Paare (i,j) mit $i < j$ und $\sigma(i) > \sigma(j)$ gerade (i_0, k) und (k, i_1) mit $i_0 < k < i_1$ und (i_0, i_1). Das m in Bemerkung (a) ist also ungerade.

(2) Lässt sich σ als Hintereinanderausführung von Transpositionen $\tau_1, \tau_2, \dots, \tau_m$ schreiben (dies ist tatsächlich für jede Permutation der Fall), so ist $\operatorname{sgn}\sigma = 1$ für gerades m und $= -1$ für ungerades m. Das m ist dabei nicht eindeutig bestimmt!

(3) $n = 4$, $\sigma = (2, 3, 4, 1)$: Es ist $\operatorname{sgn}\sigma = -1$, zB da man durch drei Transpositionen die 1 nach vorne bekommt. Oder da die Paare (i,j) mit $i < j$ und $\sigma(i) > \sigma(j)$ gerade $(1,4), (2,4), (3,4)$ sind.

16.8. Die Leibnizformel für Determinanten (nicht HM2 Prfungsrelevant): Gegeben sei die Matrix $A \in \mathbb{K}^{n\times n}$ mit den Einträgen a_{jk}. Dann gilt

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

(ohne Beweis). Ebenfalls ohne Beweis:

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

Beispiel: Ist $A \in \mathbb{K}^{n\times n}$ von der Form $\begin{pmatrix} d_1 & * & * & \cdots & * \\ 0 & d_2 & * & \cdots & * \\ 0 & 0 & \ddots & \ddots & * \\ 0 & 0 & \cdots & d_{n-1} & * \\ 0 & 0 & \cdots & 0 & d_n \end{pmatrix}$, also eine <u>obere Dreiecksmatrix</u>, so gilt $\det(A) = d_1 \cdot d_2 \cdot \ldots \cdot d_n$.

16.9. Determinantenmultiplikationssatz: Für beliebige $A, B \in \mathbb{K}^{n\times n}$ gilt

$$\det(AB) = \det(A)\det(B).$$

(ohne Beweis).

Insbesondere gilt für eine reguläre Matrix A: $\det(A^{-1}) = 1/\det(A)$.

Interpretation: Die Matrix B habe die Spalten $\vec{b}_1, \dots, \vec{b}_n$. Der von diesen Spalten aufgespannte Spat hat das Volumen $\det(B)$. Dieser Spat wird von der zur Matrix A gehörenden

linearen Abbildung $\phi_A : \mathbb{R}^n \to \mathbb{R}^n, \vec{x} \mapsto A\vec{x}$ auf den von den Spalten $A\vec{b}_1, \ldots, A\vec{b}_n$ der Matrix AB aufgespannten Spat mit Volumen $\det(AB)$ abgebildet.
Bildet man also mit der zur Matrix A gehörenden Abbildung einen beliebigen Spat ab, **so muss man dessen Volumen mit** $\det(A)$ **multiplizieren**.

16.10. 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 $\vec{a}_1, \ldots, \vec{a}_n \in \mathbb{K}^n$ habe. Wenn die Matrix A regulär ist, so hat das Gleichungssystem eine eindeutige Lösung $\vec{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$, wobei für jedes $j \in \{1, \ldots, n\}$ gilt:

$$x_j = \det(\vec{a}_1, \ldots, \underbrace{\vec{b}}_{j}, \ldots, \vec{a}_n)/\det(A).$$

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. Wir haben $\vec{b} = \sum_{l=1}^{n} x_l \vec{a}_l$, also ist wegen (D2):

$$\det(\vec{a}_1, \ldots, \underbrace{\vec{b}}_{j}, \ldots, \vec{a}_n) = \sum_{l=1}^{n} x_l \det(\vec{a}_1, \ldots, \underbrace{\vec{a}_l}_{j}, \ldots, \vec{a}_n).$$

Wegen (D3) bleibt rechts nur der Summand für $l = j$ stehen, dh $x_j \det(A)$. □

16.11. Eine Formel für die inverse Matrix: Sei $A \in \mathbb{K}^{n\times n}$ regulär mit Spalten $\vec{a}_1, \ldots, \vec{a}_n \in \mathbb{K}^n$. Geht man zur Berechnung von A^{-1} wie in 14.20 (HM I) vor und verwendet die Cramersche Regel 16.10, so erhält man

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

Die Formel für $n = 2$ in 14.21 (HM I) ist ein Spezialfall.

16.12. Orientierung (nicht HM2 Prüfungsrelevant): Die Idee in 16.1 war, dass $\det(\vec{b}_1, \ldots, \vec{b}_n)$ das Volumen des von $\vec{b}_1, \ldots, \vec{b}_n \in \mathbb{R}^n$ aufgespannten Spates beschreibt. Wegen (D2) (und (D1)) nimmt det auch negative Werte an. Das eigentliche Volumen ist $|\det(\vec{b}_1, \ldots, \vec{b}_n)|$. Aber auch das Vorzeichen von $\det(\vec{b}_1, \ldots, \vec{b}_n)$ trägt Information.

Zwischenspiel: Ist V ein n-dimensionaler reeller Vektorraum und $\varphi : V \to V$ eine lineare Abbildung, so kann man φ eine Determinante $\det \varphi$ zuordnen, indem man eine Basis

$\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n$ von V wählt, die Abbildung φ bzgl. dieser Basis durch eine Matrix A darstellt und $\det \varphi := \det(A)$ setzt.

Ist nämlich $\vec{c}_1, \ldots, \vec{c}_n$ eine weitere Basis von V, so erhalten wir die Darstellungsmatrix $\widetilde{A}$ von φ bzgl. dieser Basis als $\widetilde{A} = S^{-1}AS$, wobei S die Darstellungsmatrix der Identität $V \to V$ ist, wenn man "vorne" die Basis $\vec{c}_1, \ldots, \vec{c}_n$ und "hinten" die Basis $\vec{b}_1, \ldots, \vec{b}_n$ nimmt. Wegen 16.9 ist dann

$$\det(\widetilde{A}) = \det(S^{-1})\det(A)\det(S) = \det(S)^{-1}\det(A)\det(S) = \det(A) = \det\varphi,$$

dh die Definition ist unabhängig von der Wahl der Basis.

Definition: Eine bijektive Abbildung $\varphi : V \to V$ heißt orientierungstreu, falls $\det\varphi > 0$ 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 $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$ orientierungstreu, so ist auch $\varphi(\vec{b}_1), \varphi(\vec{b}_2), \varphi(\vec{b}_3)$ ein Rechtssystem.

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.

16.13. Das Kreuzprodukt (Vektorprodukt) im $\mathbb{R}^3$: Für zwei Vektoren $\vec{x} = (x_1, x_2, x_3), \vec{y} = (y_1, y_2, y_3) \in \mathbb{R}^3$ ist das Kreuzprodukt $\vec{x} \times \vec{y} \in \mathbb{R}^3$ derjenige Vektor, der senkrecht auf $\vec{x}$ und $\vec{y}$ steht, dessen Länge der Flächeninhalt des von $\vec{x}$ und $\vec{y}$ aufgespannten Parallelogramms ist und für den $\det(\vec{x}, \vec{y}, \vec{x} \times \vec{y}) \geq 0$ ist.

Hieraus ergeben sich folgende **Rechenregeln**:

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

(1) $\vec{e}_1 \times \vec{e}_2 = \vec{e}_3$, $\vec{e}_2 \times \vec{e}_3 = \vec{e}_1$, $\vec{e}_3 \times \vec{e}_1 = \vec{e}_2$.

(2) $(\alpha\vec{x} + \beta\vec{w}) \times \vec{y} = \alpha(\vec{x} \times \vec{y}) + \beta(\vec{w} \times \vec{y})$ 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 $\vec{x}, \vec{w}, \vec{y}, \vec{z} \in \mathbb{R}^3$, dh das Kreuzprodukt ist linear in jeder Komponente.

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

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

dh man kann zu einer Variablen ein Vielfaches der anderen dazuaddieren.

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

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 (zB $\to$ Elektromotor). 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. Die größte Kraft wirkt somit, wenn $\vec{v}$ und $\vec{B}$ 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

$$\vec{x} \times \vec{y} = \begin{vmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} \vec{e}_1 - \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} \vec{e}_2 + \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \vec{e}_3 = \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix}.$$

Begründung: Bezeichnet man die rechte Seite mit $\vec{z}$, so sieht man leicht $(\vec{z}|\vec{x}) = \det(\vec{x}, \vec{x}, \vec{y}) = 0$ und $(\vec{z}|\vec{y}) = \det(\vec{y}, \vec{x}, \vec{y}) = 0$ ein. Außerdem ist

$$\det(\vec{x}, \vec{y}, \vec{z}) = \det(\vec{z}, \vec{x}, \vec{y}) = \begin{vmatrix} \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} & -\begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} & \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = \|\vec{z}\|^2 \geq 0.$$

Der Flächeninhalt a des von $\vec{x}, \vec{y}$ aufgespannten Parallelogramms ist gegeben durch $a = \|\vec{x}\|\|\vec{y}\||\sin\varphi|$, wobei φ der von $\vec{x}$ und $\vec{y}$ eingeschlossene Winkel ist. Wegen $(\vec{x}|\vec{y}) = \|\vec{x}\|\|\vec{y}\|\cos\varphi$ erhalten wir

$$a^2 = \|\vec{x}\|^2\|\vec{y}\|^2\Big(1 - \cos^2\varphi\Big) = \|\vec{x}\|^2\|\vec{y}\|^2 - (\vec{x}|\vec{y})^2.$$

Nun rechnet man nach, dass

$$\|\vec{z}\|^2 + (\vec{x}|\vec{y})^2 = \|\vec{x}\|^2\|\vec{y}\|^2$$

gilt (zur Übung empfohlen).

Beispiel:

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

Warnung: Das Kreuzprodukt ist nicht assoziativ! So ist zB

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

16.14. 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)–(D2) der Determinante nach, und dazu die Folgerung (c). Es reicht zu beobachten dass (D2) und (c) implizieren (D3). $\square$

Beispiel: $\|\vec{x} \times \vec{y}\|^2 = (\vec{x} \times \vec{y}) \cdot (\vec{x} \times \vec{y}) = \det(\vec{x}, \vec{y}, \vec{x} \times \vec{y})$ (vgl. die Begründung oben). Anschaulich ist das auch klar, da $\det(\vec{x}, \vec{y}, \vec{x} \times \vec{y})$ das Volumen des aufgespannten Spates ist, welches man wegen der Orthogonalität von $\vec{x} \times \vec{y}$ auf dem durch $\vec{x}, \vec{y}$ aufgespannten Parallelogramm als Produkt von $\|\vec{x} \times \vec{y}\|$ mit der Parallelogrammfläche erhält.

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.

Eine Basis $\vec{x}, \vec{y}, \vec{z}$ von $\mathbb{R}^3$ ist genau dann ein Rechtssystem, wenn $\vec{z}$ auf derselben Seite der durch $\vec{x}, \vec{y}$ aufgespannten Ebene liegt wie $\vec{x} \times \vec{y}$.

17 Eigenwerte und Diagonalisierung von Matrizen

17.1. Eigenwerte und Eigenvektoren: Sei $A \in \mathbb{K}^{n\times n}$. 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.

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

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

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

Beachte: $\vec{0}$ ist kein Eigenvektor!

Beispiele: (1) 1 ist der einzige Eigenwert von I_n.

(2) Die Diagonalelemente $d_1, d_2, \ldots, d_n$ einer Diagonalmatrix $D = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & d_n \end{pmatrix}$ sind Eigenwerte von D. Für $j = 1, \ldots, n$ ist der j-te Einheitsvektor $\vec{e}_j$ ein Eigenvektor zum Eigenwert d_j.

(3) Die reelle Matrix $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ hat den Eigenwert i mit Eigenvektor $\binom{i}{1}$.

Wir betrachten von nun an komplexe Matrizen.

17.2. 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 $\operatorname{Kern}(A - \lambda I_n) \neq \{\vec{0}\}$ ist. In diesem Fall ist $E_A(\lambda) = \operatorname{Kern}(A - \lambda I_n)$.

(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 $\lambda_1, \ldots, \lambda_k$ die verschiedenen Eigenwerte von A, so gilt $k \leq n$ und

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

dh **die Summe der geometrischen Vielfachheiten der Eigenwerte von A ist höchstens n.**

Beweis. (d) folgt aus (c). (c) zeigt man durch Induktion nach der Anzahl k verschiedener Eigenwerte. Der Induktionsanfang $k = 1$ ist klar. Für den Induktionsschritt seien $\lambda_1, \ldots, \lambda_{k+1}$ verschiedene Eigenwerte mit zugehörigen Eigenvektoren $\vec{x}_1, \ldots, \vec{x}_{k+1}$. Zum Beweis von deren Unabhängigkeit seien $\alpha_1, \ldots, \alpha_{k+1} \in \mathbb{C}$ 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 $j = 1, \ldots, k$, also $\alpha_j = 0$ für $j = 1, \ldots, k$ wegen $\lambda_j \neq \lambda_{k+1}$. Schließlich folgt auch $\alpha_{k+1} = 0$ wegen $\vec{x}_{k+1} \neq 0$. □

Beispiel: Die Matrix A aus Beispiel 17.1(3) hat genau die Eigenwerte $i, -i$ und es gilt $E_A(i) = \mathrm{lin}\{\binom{i}{1}\}$, $E_A(-i) = \mathrm{lin}\{\binom{i}{-1}\}$. Die geometrische Vielfachheit von i und $-i$ ist also jeweils 1.

17.3. Charakteristisches Polynom und algebraische Vielfachheit: Sei $A \in \mathbb{C}^{n \times n}$ eine Matrix. Das <u>charakteristische Polynom χ_A von A</u> ist gegeben durch

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

Das charakteristische Polynom ist ein Polynom vom Grad n (wg. 16.8).

Ein $\lambda \in \mathbb{C}$ ist Eigenwert von A genau dann, wenn λ Nullstelle von p_A ist (vgl 16.2(e) und 17.2(a)).

Definition: Die <u>algebraische Vielfachheit eines Eigenwertes λ von A</u> 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 $\lambda_1, \ldots, \lambda_k$ die verschiedenen Eigenwerte und $m_1, m_2, \ldots, m_k$ die jeweiligen algebraischen Vielfachheiten, so gilt $m_1 + m_2 + \ldots + m_k = n$ (vgl. 5.4, HM I).

Beispiele: (1) Der Eigenwert 1 von I_n hat algebraische und geometrische Vielfachheit n: Es ist $E_{I_n}(1) = \mathbb{C}^n$ und $p_{I_n}(\lambda) = \det(I_n - \lambda I_n) = (1 - \lambda)^n$.

(2) In Beispiel 17.2 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 $p_A(\lambda) = (1-\lambda)^2 = (\lambda - 1)^2$ und $E_A(1) = \text{Kern} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \text{lin}\{\binom{1}{0}\}$.

17.4. Ähnliche Matrizen: Zwei Matrizen $A, B \in \mathbb{C}^{n\times n}$ heißen <u>ähnlich</u>, falls es eine reguläre Matrix $S \in \mathbb{C}^{n\times n}$ gibt mit $B = S^{-1}AS$.

Bemerkung: (a) Es gilt dann $A = SBS^{-1}$, denn

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

dh auch B, A sind ähnlich.

(b) Sind A, B ähnlich und B, C ähnlich, so sind auch A, C ähnlich, 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.

(c) Ähnliche Matrizen haben dieselbe Determinante und dieselben Eigenwerte mit denselben algebraischen und geometrischen Vielfachheiten. Ist nämlich $B = S^{-1}AS$, 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_A die lineare Abbildung $\vec{x} \mapsto A\vec{x}$, so wird φ bzgl. der Basis $\binom{1}{1}, \binom{1}{-1}$ durch die Matrix B dargestellt, dh φ streckt in Richtung von $\binom{1}{1}$ um den Faktor 4 und in Richtung $\binom{1}{-1}$ um den Faktor 2.

Anwendung: Gegeben sei das lineare Differentialgleichungssystem

$$\begin{aligned} x'(t) &= 3x(t) + y(t), \\ y'(t) &= x(t) + 3y(t). \end{aligned}$$

Setzt man $\binom{u(t)}{v(t)} := S^{-1}\binom{x(t)}{y(t)}$, so erhält man

$$\begin{pmatrix} u'(t) \\ v'(t) \end{pmatrix} = S^{-1}\begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} = S^{-1}A\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \underbrace{S^{-1}AS}_{=B}\begin{pmatrix} u(t) \\ v(t) \end{pmatrix},$$

also das entkoppelte System $u'(t) = 4u(t)$, $v'(t) = 2v(t)$, welches sich leicht lösen lässt: $u(t) = ae^{4t}$, $v(t) = be^{2t}$. Die Lösungen des ursprünglichen Systems erhält man dann durch

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = S\begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} ae^{4t} + be^{2t} \\ ae^{4t} - be^{2t} \end{pmatrix}, \quad t \in \mathbb{R}.$$

Die eindeutige Lösung zum Anfangswert $\binom{x(0)}{y(0)} = \binom{1}{0}$ ist wegen $\binom{a+b}{a-b} = \binom{1}{0} \iff a = b = 1/2$ also gegeben durch $x(t) = (e^{4t} + e^{2t})/2$, $y(t) = (e^{4t} - e^{2t})/2$ für $t \in \mathbb{R}$.

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

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

dh als Summe der Diagonaleinträge.

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

(b) Ähnliche Matrizen haben dieselbe Spur.

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

Beweis. (a) Gilt $C = (c_{kl})_{kl}$, so ist $AC = (\sum_{k=1}^n a_{jk}c_{kl})_{jl}$ und $CA = (\sum_{j=1}^n c_{lj}a_{jk})_{lk}$, also

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

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

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

(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 Induktoinsbeweis fhren und z.B. $\det(A - \lambda I)$ nach der ersten Zeile entwickeln). □

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

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

Hier sind die Eigenwerte von A sogar durch $\det A$ und $\mathrm{Spur}\, A$ eindeutig bestimmt.

17.6. Diagonalisierung von Matrizen: Eine Matrix $A \in \mathbb{C}^{n\times n}$ heißt <u>diagonalisierbar</u>, falls sie ähnlich zu einer Diagonalmatrix D ist, dh 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 17.4 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 $D = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & d_n \end{pmatrix}$, so ist jedes $d \in \{d_j : j = 1, \ldots, n\}$ ein Eigenwert von D mit algebraischer Vielfachheit = Anzahl der $j \in \{1, \ldots, n\}$ mit $d_j = d$. Der Eigenraum zu d ist

$$E_D(d) = \mathrm{Kern}\,(D - dI) = \mathrm{lin}\{\vec{e}_j : j \in \{1, \ldots, n\}, d_j = d\},$$

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 A n unabhängige Eigenvektoren hat, dh $\mathbb{C}^n$ hat eine Basis aus Eigenvektoren von A.

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

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

und

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

dh 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) Für jeden Eigenwert von A geometrische und algebraische Vielfachheit übereinstimmen.

(b) Die Summe der geometrischen Vielfachheiten der Eigenwerte von A ist n.

Folgerung: Hat die Matrix $A \in \mathbb{C}^{n\times n}$ n verschiedene Eigenwerte, so ist sie diagonalisierbar (dann ist nämlich für jeden Eigenwert algebraische und geometrische Vielfachheit $= 1$).

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

Beispiele: Die Matrix $A \in \mathbb{R}^{2\times 2}$ aus dem Beispiel in 17.4 ist symmetrisch. Die Matrix $\begin{pmatrix} 0 & -2i \\ 2i & 0 \end{pmatrix}$ ist hermitesch.

Satz: (a) Eine hermitesche Matrix A hat nur reelle Eigenwerte. Die Eigenvektoren zu verschiedenen Eigenwerten sind paarweise orthogonal.

(b) Jede hermitesche Matrix A läßt sich diagonalisieren, wobei die Matrix S unitär, dh mit $S^* = S^{-1}$, gewählt werden kann. Für reelle symmetrische Matrizen kann $S \in \mathbb{R}^{n\times n}$ orthogonal, dh mit $S^T = S^{-1}$ gewählt werden (ohne Beweis).

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

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

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 mit Eigenvektor $\vec{y}$, 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 $\lambda \neq \mu$. $\square$

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 17.4 und 17.6).

Die Matrix $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ aus dem Beispiel in 17.6 ist symmetrisch, die dort angegebene Matrix S ist nicht orthogonal. Hier reicht es nicht, die gefundenen Eigenvektoren auf Länge 1 zu bringen, da die beiden Eigenvektoren zum Eigenwert 4 nicht orthogonal sind. Wir können aber das Gram-Schmidt-Verfahren verwenden und setzen

$$\vec{s}_1 := \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad \vec{s}_2 := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix},$$

$$\vec{c} := \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} - \Big(\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \vec{s}_2 \Big) \vec{s}_2 = \frac{1}{2} \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}, \qquad \vec{s}_3 := \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}.$$

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

Bemerkung: Allgemeiner gilt, dass sich eine Matrix $A \in \mathbb{C}^{n\times n}$ genau dann unitär diagonalisieren läßt, wenn sie normal ist, dh genau dann, wenn $AA^* = A^*A$ gilt.

Es sei noch der folgende tiefliegendere Satz angegeben:

17.8. Satz (Jordan-Normalform): 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, dh auf der Hauptdiagonalen steht bei J_j ein Eigenwert, auf der Nebendiagonalen stehen Einsen.

Insgesamt stehen in der Jordannormalform J auf der Diagonalen die Eigenwerte von A, gemäß ihren 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 $\mathrm{Kern}\,((A - \lambda I_n)^m)$, wobei λ die Eigenwerte von A durchläuft und m ist kleiner gleich als die allgebraische Vielfachheit von λ.

Beispiel (zum Hauptraum): Sei $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. Dann gilt $p_A(\lambda) = -\lambda^3$, dh 0 ist einziger Eigenwert mit algebraischer Vielfachheit 3. Es gilt $A^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ und $A^3 = 0$, also $\dim \mathrm{Kern}\,(A) = 1$, $\dim(\mathrm{Kern}\,(A^2)) = 2$ und $\dim(\mathrm{Kern}\,(A^3)) = 3$.

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

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

dh $\det(A)$ ist das Produkt der Eigenwerte und $\mathrm{Spur}\,(A)$ 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

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

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

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

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

17.9. Definitheit reeller symmetrischer Matrizen: Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch und $q : \mathbb{R}^n \to \mathbb{R}$, $\vec{x} \mapsto \vec{x}^T A\vec{x}$, die zugehörige quadratische Form.

A heißt positiv definit, falls $\vec{x}^T A\vec{x} > 0$ für alle $\vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\}$.

A heißt negativ definit, falls $\vec{x}^T A\vec{x} < 0$ für alle $\vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\}$.

A heißt positiv semidefinit, falls für alle $\vec{x} \in \mathbb{R}^n$ gilt: $\vec{x}^T A\vec{x} \geq 0$.

A heißt negativ semidefinit, falls für alle $\vec{x} \in \mathbb{R}^n$ gilt: $\vec{x}^T A\vec{x} \leq 0$.

A heißt indefinit, falls es $\vec{x}, \vec{y} \in \mathbb{R}^n$ gibt mit $\vec{x}^T A\vec{x} > 0$ und $\vec{y}^T A\vec{y} < 0$.

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

$$q(\vec{x}) = \vec{x}^T A\vec{x} = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_jx_k \quad \text{für alle } \vec{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n.$$

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

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

Also ist A positiv definit.

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

A ist positiv definit genau dann, wenn alle Eigenwerte von A positiv sind.

A ist negativ definit genau dann, wenn alle Eigenwerte von A negativ sind.

A ist positiv semidefinit genau dann, wenn $\lambda \geq 0$ für alle Eigenwerte von A gilt.

A ist negativ semidefinit genau dann, wenn $\lambda \leq 0$ für alle Eigenwerte von A gilt.

A ist indefinit genau dann, wenn A sowohl positive als auch negative Eigenwerte hat.

Beweis. Seien $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von A (gemäß algebraischer Vielfachheit wiederholt). Diagonalisiere A mit einer orthogonalen Matrix S: $S^T AS = D$, wobei D die Spalten $\lambda_1\vec{e}_1, \ldots, \lambda_n\vec{e}_n$ hat. Schreibe $\vec{x} = S\vec{y}$, wobei $\vec{y} = (y_1, \ldots, y_n)$. Dann ist $\vec{x} = \vec{0}$ äquivalent zu $\vec{y} = \vec{0}$ und

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

□

Folgerung: Sei $A \in \mathbb{R}^{2\times 2}$ symmetrisch. Dann ist A indefinit genau dann, wenn $\det(A) < 0$ ist.

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

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

Beispiele: Die Matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ ist indefinit. Die Matrix $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ ist positiv definit. Die Matrix $\begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$ ist negativ definit.

Für größere Matrizen kann man das folgende verwenden:

Kriterium von Hurwitz: Eine symmetrische Matrix $A = (a_{jk})_{jk} \in \mathbb{R}^{n\times n}$ ist positiv definit genau dann, wenn alle <u>Hauptunterdeterminanten</u> positiv sind, dh wenn

$$\begin{vmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mm} \end{vmatrix} > 0$$

für alle $m = 1, 2, \ldots, n$ gilt.

18 Differentialgleichungen

18.1. Allgemeines: Eine Differentialgleichung erster Ordnung hat die Form

$$y'(t) = f(t, y(t)).$$

Die Vorstellung ist, dass $y(t)$ den zeitabhängigen Zustand eines Systems beschreibt und dass die Gleichung angibt, wie die zukünftige Entwicklung des Systems (infinitesimal gegeben durch $y'(t)$) vom gegenwärtigen Zustand $y(t)$ und dem Zeitpunkt t abhängt. Wenn man jetzt zum Zeitpunkt t_0 den Zustand y_0 des Systems gegeben hat, so sucht man eine Lösung des Anfangswertproblems

$$y'(t) = f(t, y(t)), \qquad y(t_0) = y_0,$$

dh eine stetig differenzierbare Funktionen $t \mapsto \phi(t)$, die auf einem Intervall I mit $t_0 \in I$ definiert ist, und den Bedingungen $\phi(t_0) = y_0$ und $\phi'(t) = f(t, \phi(t))$ für alle $t \in I$ genügt. Diese beschreibt dann hoffentlich den Systemverlauf.

Beispiel RL-Kreis: In einem Stromkreis mit Widerstand R und einer Spule der Induktivität L gilt, wenn wir die zeitabhängige Spannung $U(t)$ anlegen, für den Strom $J(t)$:

$$\frac{d}{dt} J(t) = -\frac{R}{L} J(t) + \frac{U(t)}{L}.$$

Hier gehen wir etwa davon aus, dass zum Zeitpunkt $t_0 = 0$ kein Strom fließt, dh wir haben als Anfangswert $J(0) = 0$.

Allgemeine Fragen: 1) Existenz: Hat ein gegebenes Anfangswertproblem eine Lösung?

2) Eindeutigkeit: Stimmen je zwei Lösungen auf dem Schnitt ihrer Definitionsbereiche überein? Wenn nicht, so ist der Systemverlauf durch das Anfangswertproblem nicht eindeutig festgelegt!

3) Globale Lösungen: Wenn 1), 2) geklärt sind, wie groß kann man ggf. das Existenzintervall I machen bzw. findet man Lösungen mit $I = [t_0, \infty)$?

4) Asymptotisches Verhalten: Wenn $I = [t_0, \infty)$ ist, wie verhält sich $\phi(t)$ für $t \to \infty$?

Wir betrachten im weiteren nur lineare Differentialgleichungen, für die sich 1), 2), 3) leicht beantworten lassen. Im allgemeinen ist das nicht unbedingt so, wie die folgenden Beispiele zeigen:

Beispiele: (a) $y'(t) = \sqrt{|y(t)|}$, $y(0) = 0$. Eine Lösung ist gegeben durch $\phi_1(t) = 0$ für $t \in \mathbb{R}$, eine weitere durch $\phi_2(t) = t|t|/2$, $t \in \mathbb{R}$. Hier hat man globale Existenz, aber keine Eindeutigkeit der Lösung des Anfangswertproblems (Verkleinern der Existenzintervalle hilft nicht).

(b) $y'(t) = 1 + y(t)^2$, $y(0) = 0$. Hier ist die (tatsächlich eindeutige) Lösung des Anfangswertproblems gegeben durch $\phi(t) = \tan t$, $t \in (-\frac{\pi}{2}, \frac{\pi}{2})$, und das Existenzintervall lässt sich nicht vergrößern. Dass es gerade bei $\frac{\pi}{2}$ nicht weitergeht, sieht man dem Problem aber nicht direkt an.

18.2. Lineare Differentialgleichung erster Ordnung: Sei $I \subseteq \mathbb{R}$ Intervall und seien $a, b : I \to \mathbb{C}$ stetig. Für jedes $t_0 \in I$ und jedes $y_0 \in \mathbb{C}$ hat das Anfangswertproblem

$$\begin{aligned} y'(t) &= a(t)y(t) + b(t) \\ y(t_0) &= y_0, \end{aligned} \tag{1}$$

genau eine Lösung $\phi : I \to \mathbb{C}$. Diese ist gegeben durch

$$\phi(t) = e^{A(t)}y_0 + e^{A(t)} \int_{t_0}^{t} e^{-A(\tau)} b(\tau)\, d\tau, \quad t \in I,$$

wobei $A(t) := \int_{t_0}^{t} a(r)\, dr$ für $t \in I$ gesetzt wurde.

Man kann nachrechnen (beachte $A'(t) = a(t)$), dass ϕ eine Lösung ist.

Beobachtungen dazu: 1) Sind $y(t)$, $z(t)$ Lösungen der inhomogenen Gleichung $y'(t) = a(t)y(t) + b(t)$, so löst $w(t) := y(t) - z(t)$ die homogene Gleichung $w'(t) = a(t)w(t)$. Die Lösung erhält man also als $y(t) = w(t) + z(t)$, wobei $z(t)$ irgendeine Lösung der inhomogenen Gleichung und $w(t)$ eine geeignete Lösung der homogenen Gleichung ist.

2) Alle Lösungen der homogenen Gleichung $w'(t) = a(t)w(t)$ sind durch $w(t) = ce^{A(t)}$ mit $c \in \mathbb{C}$ gegeben (hierbei ist $c = w(0)$). Zum einen sind dies Lösungen der homogenen Gleichung, zum anderen muss man einsehen, dass das Anfangswertproblem $w'(t) = a(t)w(t)$, $w(0) = 0$ nur die Nullfunktion als Lösung hat. [2]

3) Eine Lösung $z(t)$ der inhomogenen Gleichung erhält man durch Variation der Konstanten in $w(t) = ce^{A(t)}$, dh durch den Ansatz $z(t) = c(t)e^{A(t)}$:

$$c'(t)e^{A(t)} + a(t)c(t)e^{A(t)} = z'(t) \overset{!}{=} a(t)z(t) + b(t) = a(t)c(t)e^{A(t)} + b(t),$$

also $c'(t) = e^{-A(t)}b(t)$.
Dann ist etwa $c(t) = \int_{t_0}^{t} e^{-A(\tau)}b(\tau)\, d\tau$ und $z(t) = c(t)e^{A(t)} = e^{A(t)} \int_{t_0}^{t} e^{-A(\tau)}b(\tau), d\tau$ wie oben.

[2] Skizze dazu für $t_0 = 0$ als Bonus: Es gilt

$$w(t) = \int_0^t a(\tau)w(\tau)\, d\tau, \quad \text{also} \quad |w(t)| \le \int_0^t |a(\tau)||w(\tau)|\, d\tau.$$

Ist nun $|a(t)| \le K$ und $|w(t)| \le L$ auf I, so erhält man der Reihe nach

$$|w(t)| \le LKt, \quad |w(t)| \le K^2 L \int_0^t \tau\, d\tau = \frac{LK^2t^2}{2}, \quad |w(t)| \le K^3 L \int_0^t \frac{\tau^2}{2}\, d\tau = \frac{LK^3t^3}{3!}$$

etc, also $|w(t)| \le \frac{L(Kt)^n}{n!} \to 0$ $(n \to \infty)$, und es folgt $w(t) = 0$ auf I.

Beispiele: (1) Die Lösung für den RL-Kreis in 18.1 mit Anfangswert $J(0) = 0$ ist nach der Formel gegeben durch

$$J(t) = e^{-\frac{R}{L}t} \int_0^t e^{\frac{R}{L}\tau} \frac{U(\tau)}{L}\, d\tau, \quad t \geq 0.$$

Ist etwa $U(t) = U_0$ konstant (zum Zeitpunkt $t_0 = 0$ wird die konstante Spannung U_0 angelegt), so erhält man

$$J(t) = \frac{U_0}{R}\Big(1 - e^{-\frac{R}{L}t}\Big), \quad t \geq 0,$$

dh die Induktivität L macht sich beim Anlegen der Spannung bemerkbar, der Effekt klingt exponentiell ab und zwar umso schneller, je kleiner die Induktivität ist.

(2) $y'(t) = -(\sin t)y(t) + \sin^3 t$. Hier ist $A(t) = \cos t$ und die Lösung der homogenen Gleichung ist gegeben durch $y(t) = ce^{\cos t}$. Eine Lösung der inhomogenen Gleichung erhalten wir durch die "Variation-der-Konstanten"-Formel (wobei wir $r = \cos\tau$, $dr = -\sin\tau\, d\tau$ substituieren):

$$\begin{aligned} z(t) &= e^{\cos t} \int_0^t e^{-\cos\tau} \sin^3\tau\, d\tau \\ &= e^{\cos t} \int_1^{\cos t} (r^2 - 1)e^{-r}\, dr \\ &= -e^{\cos t}(r^2 - 1 + 2r + 2)e^{-r}\Big|_1^{\cos t} \\ &= \sin^2 t - 2\cos t - 2 + \underbrace{4e^{\cos t - 1}}_{\text{Lsg. der hom. Gl.}} \end{aligned}$$

Die allgemeine Lösung der Differentialgleichung ist also gegeben durch

$$y(t) = \sin^2 t - 2\cos t - 2 + ce^{\cos t}, \quad t \in \mathbb{R},$$

wobei $c \in \mathbb{C}$ eine Konstante ist.

18.3. Gleichungen höherer Ordnung: Viele Systeme werden durch Differentialgleichungen höherer (insbesondere zweiter) Ordnung beschrieben. Wir beschränken uns auf Gleichungen mit konstanten Koeffizienten.

Beispiel RLC-Kreis: Wir nehmen in den Stromkreis zusätzlich einen Kondensator der Kapazität C auf. Die an ihm entstehende Spannung ist $U_c(t) = \frac{1}{C}q(t)$, wobei q die Ladung bezeichnet. Das führt zur Differentialgleichung

$$LJ'(t) + RJ(t) + \frac{1}{C}q(t) = U(t)$$

und nach Division durch L wegen $J(t) = \frac{d}{dt}q(t)$ zur Differentialgleichung

$$q''(t) + \frac{R}{L}q'(t) + \frac{1}{LC}q(t) = \frac{1}{L}U(t).$$

Hier sucht man zweimal stetig differenzierbare Lösungen $t \mapsto q(t)$.
Wenn wir die Ersetzung nicht machen, haben wir ein gekoppeltes System von Gleichungen erster Ordnung

$$\begin{aligned} q'(t) &= J(t) \\ J'(t) &= -\tfrac{R}{L}J(t) - \tfrac{1}{LC}q(t) + \tfrac{1}{L}U(t) \end{aligned}$$

bzw.

$$\frac{d}{dt}\begin{pmatrix} q(t) \\ J(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\frac{1}{LC} & -\frac{R}{L} \end{pmatrix}\begin{pmatrix} q(t) \\ J(t) \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{U(t)}{L} \end{pmatrix},$$

was die Analogie zu 18.2 unterstreicht. Hier muss man $q(0)$ und $J(0)$ vorschreiben bzw. $q(0)$ und $q'(0)$.

Satz: Sei $I \subseteq \mathbb{R}$ ein Intervall und $b : I \to \mathbb{C}$ stetig. Seien $n \in \mathbb{N}$ und $a_0, a_1, \ldots, a_{n-1} \in \mathbb{C}$. Dann hat für jedes $t_0 \in I$ und jeden Anfangsvektor $(y_0, y_1, \ldots, y_{n-1}) \in \mathbb{C}^n$ das Anfangswertproblem

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \ldots + a_1 y'(t) + a_0 y(t) \;=\; b(t), \quad t \in I$$

$$y(t_0) = y_0, y'(t_0) = y_1, \ldots, y^{(n-1)}(t_0) = y_{n-1},$$

genau eine n-mal stetig differenzierbare Lösung $\phi : I \to \mathbb{C}$ (ohne Beweis).

Beobachtungen dazu: 1) Sind $y(t)$, $z(t)$ Lösungen der inhomogenen Gleichung

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \ldots + a_1 y'(t) + a_0 y(t) = b(t),$$

so löst $w(t) := y(t) - z(t)$ die homogene Gleichung

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \ldots + a_1 y'(t) + a_0 y(t) = 0.$$

Die Lösung erhält man also wieder als $y(t) = w(t) + z(t)$, wobei $z(t)$ irgendeine Lösung der inhomogenen Gleichung und $w(t)$ eine geeignete Lösung der homogenen Gleichung ist.

2) Die Menge $\mathscr{L}_0$ der Lösungen der homogenen Gleichung ist ein komplexer Vektorraum. Die Abbildung $\mathscr{L}_0 \to \mathbb{C}^n$, $\phi \mapsto (\phi(t_0), \phi'(t_0), \ldots, \phi^{(n-1)}(t_0))$ ist linear und nach dem Satz bijektiv. Also hat $\mathscr{L}_0$ die Dimension n. Eine Basis von $\mathscr{L}_0$ heißt Fundamentalsystem der homogenen Gleichung.

3) Der Ansatz $\phi(t) = e^{\lambda t}$ für eine Lösung der homogenen Gleichung führt auf

$$\underbrace{\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0}_{:=p(\lambda)} = 0,$$

dh $\lambda \in \mathbb{C}$ muss Nullstelle des charakteristischen Polynoms p der Gleichung sein. Gilt

$$p(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \cdot \ldots \cdot (\lambda - \lambda_k)^{m_k},$$

wobei $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{C}$ die verschiedenen Nullstellen von p mit Vielfachheiten $m_1, m_2, \ldots, m_k \in \mathbb{N}$ sind, so ist ein Fundamentalsystem gegeben durch die $m_1 + m_2 + \ldots + m_k = n$ Funktionen

$$e^{\lambda_1 t}, te^{\lambda_1 t}, \ldots, t^{m_1-1}e^{\lambda_1 t},\ e^{\lambda_2 t}, te^{\lambda_2 t}, \ldots, t^{m_2-1}e^{\lambda_2 t}, \ldots, e^{\lambda_k t}, te^{\lambda_k t}, \ldots, t^{m_k-1}e^{\lambda_k t}.$$

Wir werden dies im Zusammenhang mit der Laplacetransformation einsehen. [3]

Beispiel $y'''(t) + 3y''(t) + 3y'(t) + y(t) = 0$: Das charakteristische Polynom ist

$$p(\lambda) = \lambda^3 + 3\lambda^2 + 3\lambda + 1 = (\lambda + 1)^3,$$

ein Fundamentalsystem ist also hier gegeben durch $e^{-t}, te^{-t}, t^2e^{-t}$.

4) Lösungen der inhomogenen Gleichung werden wir mithilfe der Laplacetransformation bestimmen.

[3] Will man direkt einsehen, dass dies Lösungen sind, muss man beachten, dass eine m_j-fache Nullstelle von p auch Nullstelle von $p', p'', \ldots, p^{(m_j-1)}$ ist.

19 Mehrdimensionale Differentialrechnung

Wir schreiben Vektoren $\vec{x}$ im $\mathbb{R}^n$ mit den Komponenten $x_1, x_2, \ldots, x_n$ als n-Tupel $(x_1, \ldots, x_n)$ oder als Spaltenvektor $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

Für $n = 2, 3$ schreiben wir häufig $\begin{pmatrix} x \\ y \end{pmatrix}$ bzw. $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

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 alle komponenten der Folge konvergent sind. In diesem Fall ist der Limes der Vektor dessen Komponenten die zugehörigen Grenzwerte sind.

Beispiel: Sei $\vec{x}_k := \begin{pmatrix} e^{-k} \\ 1 - 1/k \end{pmatrix}$ für $k \in \mathbb{N}$. Dann gilt $\vec{x}_k \to \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ für $k \to \infty$.

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

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

offene Kugel um $\vec{x}_0$ mit Radius r.

Eine Teilmenge $Q \subseteq \mathbb{R}^n$ heißt offen, falls es zu jedem $\vec{x}_0 \in Q$ ein (von $\vec{x}_0$ abhängiges) $r > 0$ gibt mit $K(\vec{x}_0, r) \subseteq Q$.

Beispiele: $\mathbb{R}^n$ ist offen. Offene Kugeln sind offen. Der Würfel $(0, 1)^n$ ist offen.

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

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

also $\vec{x} \in K(\vec{y}_0, s)$. $\square$

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

Beispiele: $\mathbb{R}^n$ ist abgeschlossen. Endliche Mengen sind abgeschlossen. Abgeschlossene Kugeln

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

sind abgeschlossen.

Sei $f : I \to \mathbb{R}^n$, $t \mapsto f(t) = \begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix}$ eine Funktion, wobei $I \subseteq \mathbb{R}$ ein Intervall ist. f heißt stetig in I, falls jede komponente f_j von f stetig ist in I.

Sei $D \subset \mathbb{R}^n$ nicht leer. D heißt wegzusammenhängend, falls es für jedes paar $\vec{x}, \vec{y} \in D$ einen weg von $\vec{x}$ nach $\vec{y}$ in D gibt, d.h. eine stetige Funktion $f : [0,1] \to D$ mit $f(0) = \vec{x}$ und $f(1) = \vec{y}$.

Beispiel Die Kugel $K(\vec{x_0}, r)$ ist eine wegzusammenhängende Menge. $D := K((0,0),1) \cup K((3,3),2)$ ist keine wegzusammenhängende Menge in $\mathbb{R}^2$, weil zum Beispiel $(0,0),(3,3) \in D$ es gibt aber keinen Weg von $(0,0)$ nach $(3,3)$ in D.

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

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

Definition: Sei $f : D \to \mathbb{R}^m$, wobei $D \subset \mathbb{R}^n$ wegzusammenhängend ist, mit mehr als ein Elemente. Wir sagen, dass

$$\lim_{\vec{x} \to \vec{x_0}} f(\vec{x}) = \vec{c},$$

falls für alle $\epsilon > 0$ ein $\delta > 0$ existiert so, dass $0 < \|\vec{x} - \vec{x_0}\| < \delta \implies \|f(\vec{x}) - \vec{c}\| < \epsilon$.

Satz: Es gilt

$$\lim_{\vec{x} \to \vec{x_0}} f(\vec{x}) = \vec{c},$$

dann und nur dann, wenn $f(\vec{x_n}) \to \vec{c}$, für alle Folgen $\vec{x_n}$ mit $\vec{x_n} \to \vec{x_0}$.

Definition: Sei $f : D \to \mathbb{R}^m$, wobei $D \subset \mathbb{R}$ wegzusammenhängend ist, mit mehr als ein Elemente. Die Funktion f heißt stetig in $\vec{x}_0 \in D$, falls

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

f heißt stetig in D, falls f in **jedem** $\vec{x}_0 \in D$ stetig ist [4].

Beispiele: (1) Sei $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = \begin{cases} xy/(x^2+y^2) & ,(x,y) \neq (0,0) \\ 0 & ,(x,y) = (0,0) \end{cases}$. f ist in jedem Punkt $(x,y) \neq (0,0)$ stetig. f ist aber in $(0,0)$ nicht stetig, denn für $x_k = y_k = 1/k$ gilt

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

[4] Die obigen Definitionen des Limes und der Stetigkeit sind auch anwendbar wenn D eine endliche Vereinigung von solchen Mengen ist. Eigentlich kann man den Begriff der Stetigkeit, für jede Teilmenge D definieren, aber das machen wir hier nicht.

(2) Sei $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = \begin{cases} x|y|^\beta (x^2+y^2)^{-1} & ,(x,y) \neq (0,0) \\ 0 & ,(x,y) = (0,0) \end{cases}$, wobei $\beta > 1$. Die Funktion ist in jedem Punkt $(x,y) \neq (0,0)$ stetig. f ist auch $(0,0)$ stetig, denn wegen $|xy| \leq (x^2+y^2)/2$ gilt

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

(3) Sei $\phi : \mathbb{R}^n \to \mathbb{R}^m$ 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 $\vec{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ gilt

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

Für $\|\vec{x}\| \to 0$ gilt also $\|\phi(\vec{x})\| \to 0$.

(4) Kompositionen von stetigen Funktionen sind stetig. Die Addition $+ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ und Multiplikation $\mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ sind stetig, des weiteren Skalarprodukt, Matrix-Vektor-Produkt, Multiplikation von Matrizen, Determinante und Kreuzprodukt im $\mathbb{R}^3$. Auch die auf den regulären Matrizen erklärte Matrixinversion ist stetig (Cramersche Regel!).

19.4. Differenzierbarkeit von Funktionen $\mathbb{R} \supseteq I \to \mathbb{R}^n$: Sei $f : I \to \mathbb{R}^n$, $t \mapsto f(t) = \begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix}$ eine Funktion, wobei $I \subseteq \mathbb{R}$ ein Intervall ist.

Definition: Die Funktion f heißt <u>differenzierbar in $t_0 \in I$</u>, [bzw. in I] genau dann, wenn jede Komponentenfunktion f_j, $j = 1, \ldots, n$, differenzierbar in t_0 [bzw. in I] ist. Wir definieren dann

$$\dot{f}(t_0) = \begin{pmatrix} \dot{f}_1(t_0) \\ \vdots \\ \dot{f}_n(t_0) \end{pmatrix}.$$

f heißt <u>stetig differenzierbar in I</u> oder <u>eine C^1-Funktion</u>, falls $\dot{f} : I \to \mathbb{R}^n$ zusätzlich stetig ist (d.h. wenn alle $\dot{f}_j : I \to \mathbb{R}$, $j = 1, \ldots, n$, stetig sind.).

Beispiele: Sei $f : I \to \mathbb{R}^2$, $f(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$ gegeben durch

(1) $x(t) = r \cos t$, $y(t) = r \sin t$ mit $r > 0$ und $I = [0, 2\pi]$ (Kreisrand um $(0,0)$ mit Radius r). Dann gilt $\dot{x}(t) = -r \sin t$, $\dot{y}(t) = r \cos t$, und f ist C^1.

(2) $x(t) = e^{-t} \cos t$, $y(t) = e^{-t} \sin t$ mit $I = [0, \infty)$ (<u>logarithmische Spirale</u>). Dann ist

$$\begin{aligned} \dot{x}(t) &= -e^{-t} \cos t - e^{-t} \sin t \\ \dot{y}(t) &= -e^{-t} \sin t + e^{-t} \cos t, \end{aligned}$$

und f ist C^1.

(3) Wir können das Differentialgleichungssystem $\dot{x}(t) = 3x(t) + y(t)$, $\dot{y}(t) = y(t) + 3x(t)$ mittels $z(t) := \binom{x(t)}{y(t)}$ schreiben als $\dot{z}(t) = \left(\begin{smallmatrix} 3 & 1 \\ 1 & 3 \end{smallmatrix}\right) z(t)$.

19.5. Raumkurven: Eine Raumkurve ist eine stetig differenzierbare Abbildung $\gamma : I \to \mathbb{R}^n$, wobei $I \subseteq \mathbb{R}$ ein Intervall ist.

Meist ist I von der Form $[a, b]$. Die Menge $\gamma(I) \subseteq \mathbb{R}^n$ heißt Spur von γ oder Bild von γ.

19.6. Bogenlänge von Raumkurven: Ist $\gamma : [a, b] \to \mathbb{R}^n$ eine Raumkurve, so ist ihre Länge (oder Bogenlänge) gegeben durch

$$L(\gamma) := \int_a^b \|\dot{\gamma}(t)\| \, dt.$$

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

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

(2) Für die logarithmische Spirale $\gamma : [0, \infty) \to \mathbb{R}^2$, $\gamma(t) = \begin{pmatrix} e^{-t}\cos t \\ e^{-t}\sin t \end{pmatrix}$ berechnen wir unter der Verwendung der Formel für $\dot{\gamma}(t)$ aus Beispiel 19.4(2):

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

19.7. Richtungsableitungen: Sei $D \subseteq \mathbb{R}^n$ offen und $f : D \to \mathbb{R}^m$ 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

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

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 $\frac{\partial f}{\partial \vec{v}}(\vec{x}_0) = \dot{g}(0)$ (vgl 19.4).

19.8. Partielle Ableitungen: In der Situation von 19.7 heißen Richtungsableitungen von f in Richtung der Einheitsvektoren $\vec{e}_1, \ldots, \vec{e}_n$ <u>partielle Ableitungen von f</u>, dh

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

für $k = 1, \ldots, m$. Man schreibt oft auch nur $f_{x_k}(\vec{x}_0)$.

Beispiele: (1) $D = \{(x, y, z) \in \mathbb{R}^3 : z > 0\}$ ist offen. Sei $f : D \to \mathbb{R}$, $f(x, y, z) = e^{-x} \cos y + \log z$. In jedem Punkt existieren alle partiellen Ableitungen, und es gilt

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

Hier wurde jeweils in der Notation das Argument (x, y, z) unterdrückt!

(2) Sei $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & , (x, y) \neq (0, 0) \\ 0 & , (x, y) = (0, 0) \end{cases}$. Wir betrachten 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 $t \to 0$ existiert in $\mathbb{R}$ genau dann, wenn $\xi\eta = 0$ ist, dh 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.

(3) Sei $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \begin{cases} \frac{x|x|^{1/2}|y|^{3/2}}{x^2+y^2} & , (x, y) \neq (0, 0) \\ 0 & , (x, y) = (0, 0) \end{cases}$. Wir betrachten Richtungsableitungen im Punkt $(0, 0)$. Für $(\xi, \eta) \neq (0, 0)$ gilt hier

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

und wir erhalten für $t \to 0$:

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

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 $f : D \to \mathbb{R}^m$ 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 $f'(\vec{x}_0) \in \mathbb{R}^m = \mathbb{R}^{m\times 1}$, dh $h \mapsto f'(\vec{x}_0)h$ ist eine lineare Abbildung $\mathbb{R} \to \mathbb{R}^m$.

Definition: f heißt <u>differenzierbar</u> in $\vec{x}_0 \in D$ (gelegentlich <u>total differenzierbar</u> in $\vec{x}_0$), falls es eine lineare Abbildung $\phi : \mathbb{R}^n \to \mathbb{R}^m$ gibt mit

$$\frac{\|f(\vec{x}) - f(\vec{x}_0) - \phi(\vec{x} - \vec{x}_0)\|}{\|\vec{x} - \vec{x}_0\|} \to 0 \quad (\vec{x} \to \vec{x}_0)$$

bzw.

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

In diesem Fall ist die lineare Abbildung ϕ eindeutig bestimmt und heißt <u>Ableitung von f in $\vec{x}_0$</u>, Bezeichnung: $f'(\vec{x}_0) := \phi$. Andere Bezeichnungen: $Df(\vec{x}_0)$, $J_f(\vec{x}_0)$, <u>Jacobimatrix</u>, <u>Funktionalmatrix</u>. Ist $A \in \mathbb{R}^{m\times n}$ die Matrix zur ϕ bezüglich der Standardbasen, dann schreibt man auch $f'(\vec{x}_0) = A$.

f heißt <u>differenzierbar in D</u>, falls f in **jedem** $\vec{x}_0 \in D$ differenzierbar ist.

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

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

Insbesondere gilt dann

$$\frac{\partial f}{\partial x_k}(\vec{x}_0) = f'(\vec{x}_0)\vec{e}_k \quad k\text{-te Spalte von } f'(\vec{x}_0), k = 1, \dots, n,$$

und

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

Im Fall $m = 1$ ist also

$$f'(\vec{x}_0) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\vec{x}_0) & \frac{\partial f}{\partial x_2}(\vec{x}_0) & \dots & \frac{\partial f}{\partial x_n}(\vec{x}_0) \end{pmatrix} \in \mathbb{R}^{1\times n} \quad \text{(Zeilenvektor).}$$

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 $(0,0)$: 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 $\frac{\partial f}{\partial(1,1)}(0,0) = 0$. Nach Beispiel 19.8(3) ist aber $\frac{\partial f}{\partial(1,1)}(0,0) = 1/2 \neq 0$.

(3) Die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x) = \begin{cases} \frac{x|x|^{1/2}y^2}{x^2+y^2} & ,(x,y) \neq (0,0) \\ 0 & ,(x,y) = (0,0) \end{cases}$ 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 $A := (0\ 0)$. Nun schätzen wir ab:

$$\frac{\|f(x,y) - f(0,0) - A\binom{x}{y}\|}{\|\binom{x}{y} - \binom{0}{0}\|} = \frac{|f(x,y)|}{\sqrt{x^2+y^2}} = \frac{|x|^{3/2}y^2}{(x^2+y^2)^{3/2}} \leq \frac{|x|^{3/2}}{(x^2+y^2)^{1/2}} \leq |x|^{1/2} \to 0$$

für $\binom{x}{y} \to \binom{0}{0}$. Somit ist f in $(0,0)$ differenzierbar, und es gilt $f'(0,0) = (0\ 0)$.

(4) Sei $f : \mathbb{R}^n \to \mathbb{R}^m$, $f(\vec{x}) = B\vec{x}$, wobei $B \in \mathbb{R}^{m\times n}$. Für $\vec{x}, \vec{h} \in \mathbb{R}^n$ gilt

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

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

(5) Sei $f : \mathbb{R}^n \to \mathbb{R}$, $f(\vec{x}) = \vec{x}^T B\vec{x}$, wobei $B \in \mathbb{R}^{n\times n}$ symmetrisch sei. Dann gilt für $\vec{x}, \vec{h} \in \mathbb{R}^n$:

$$f(\vec{x}+\vec{h}) - f(\vec{x}) = (\vec{x}+\vec{h})^T B(\vec{x}+\vec{h}) - \vec{x}^T B\vec{x} = \vec{x}^T B\vec{h} + \underbrace{\vec{h}^T B\vec{x}}_{=\vec{x}^T B\vec{h}} + \vec{h}^T B\vec{h} = 2\vec{x}^T B\vec{h} + \vec{h}^T B\vec{h}.$$

Die Abbildung $\vec{h} \mapsto 2\vec{x}^T B\vec{h}$ ist linear, und es gilt

$$|f(\vec{x}+\vec{h}) - f(\vec{x}) - 2\vec{x}^T B\vec{h}| = |\vec{h}^T B\vec{h}| \leq \max\{|\lambda| : \lambda \text{ EW von } B\} \cdot \|\vec{h}\|^2$$

(vergleiche Beweis in 17.9: Die Abschätzung ist klar, wenn B eine Diagonalmatrix ist. Ist B keine Diagonalmatrix, so diagonalisiere B durch eine orthogonale Matrix S, dh $D = S^T BS$ bzw. $B = SDS^T$. Dann folgt:

$$\begin{aligned} |\vec{h}^T B\vec{h}| = |\vec{h}^T SDS^T\vec{h}| = |(S^T\vec{h})D(S^T\vec{h})| &\leq \max\{|\lambda| : \lambda \text{ EW von } D\} \cdot \|S^T\vec{h}\|^2 \\ &= \max\{|\lambda| : \lambda \text{ EW von } B\} \cdot \|\vec{h}\|^2 \end{aligned}$$

wegen $\|S^T\vec{h}\| = \|\vec{h}\|$.)

19.10. Kriterium für Differenzierbarkeit, stetige Differenzierbarkeit: Sei $D \subseteq \mathbb{R}^n$ offen und $f : D \to \mathbb{R}^m$ mit Komponentenfunktionen $f_1, \ldots, f_m : D \to \mathbb{R}$.

Definition: f heißt in D partiell differenzierbar, falls alle partiellen Ableitungen $\frac{\partial f_j}{\partial x_k}$ auf D existieren, und stetig partiell differenzierbar in D, falls die partiellen Ableitungen zusätzlich auf D stetig sind.

Satz: 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.

Definition: Eine stetig partiell differenzierbare Funktion $f : D \to \mathbb{R}^m$ heißt stetig differenzierbar, geschrieben $f \in C^1(D, \mathbb{R}^m)$.

Beispiel: Für die Funktion f aus Beispiel 19.8(2) gilt $f_x = \frac{y(x^2+y^2)-2x^2y}{(x^2+y^2)^2}$ und $f_y = \frac{x(x^2+y^2)-2xy^2}{(x^2+y^2)^2}$ auf $\mathbb{R}^2 \setminus \{(0,0)\}$. Also ist $f \in C^1(\mathbb{R}^2 \setminus \{(0,0)\}, \mathbb{R})$.

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

$$\frac{\partial}{\partial x_l}\frac{\partial f}{\partial x_k}(\vec{x}_0).$$

Entsprechend werden (falls vorhanden) Ableitungen höherer Ordnung definiert.

Schreibweisen: $\frac{\partial^3 f}{\partial^2 x \partial y} = f_{xxy}$ etc.

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

$f : D \to \mathbb{R}^k$ heißt k-mal stetig differenzierbar, $f \in C^k(D, \mathbb{R}^m)$, falls für alle Komponentenfunktionen f_j, $j = 1, \ldots, m$ gilt: $f_j \in C^k(D, \mathbb{R})$.

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.

Beispiel: Für die Funktion f aus Beispiel 19.8(2) existieren partielle Ableitungen beliebiger Ordnung in $\mathbb{R}^2 \setminus \{(0,0)\}$. Also ist $f \in C^k(\mathbb{R}^2 \setminus \{(0,0)\}, \mathbb{R})$ für jedes $k \in \mathbb{N}$. Man schreibt dafür auch $f \in C^\infty(\mathbb{R}^2 \setminus \{(0,0)\}, \mathbb{R})$.

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

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

Man schreibt statt $\operatorname{grad} f(\vec{x}_0)$ 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 $\nabla f : D \to \mathbb{R}^n$ macht.

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

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

(beachte $f'(\vec{x}_0) \in \mathbb{R}^{1\times n}$).

Satz: Sei $f : D \to \mathbb{R}$ in $\vec{x}_0 \in D$ differenzierbar und $\operatorname{grad} f(\vec{x}_0) \neq \vec{0}$. Dann gilt für jeden Vektor $\vec{v} \in \mathbb{R}^n$:

$$\frac{\partial f}{\partial \vec{v}}(\vec{x}_0) = (\operatorname{grad} f(\vec{x}_0)) \cdot \vec{v}.$$

Insbesondere, wenn $\|\vec{v}\| = 1$ gilt

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

Dabei gilt Gleichheit rechts genau dann, wenn $\vec{v} = \operatorname{grad} f(\vec{x}_0)/\|\operatorname{grad} f(\vec{x}_0)\|$ ist (dh "der Gradient zeigt in Richtung des stärksten Anstiegs von f").

Außerdem steht $\operatorname{grad} f(\vec{x}_0)$ senkrecht auf der Niveaulinie $\{\vec{x} \in \mathbb{R}^n : f(\vec{x}) = f(\vec{x}_0)\}$.

Beweis. Es ist

$$\frac{\partial f}{\partial \vec{v}}(\vec{x}_0) \overset{19.9}{=} f'(\vec{x}_0)\vec{v} = (\operatorname{grad} f(\vec{x}_0))^T \vec{v} = (\operatorname{grad} f(\vec{x}_0)) \cdot \vec{v}.$$

Nun verwendet man die Cauchy-Schwarz-Ungleichung. □

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

Beispiele: (1) Sei $F : \mathbb{R}^n \to \mathbb{R}$ differenzierbar und $\phi : I \to \mathbb{R}^n$ differenzierbar, wobei $I \subseteq \mathbb{R}$ ein Intervall ist. Dann ist $F \circ \phi : I \to \mathbb{R}$, $t \mapsto F(\phi(t))$ differenzierbar, und es gilt

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

wobei $\phi_1, \phi_2, \ldots, \phi_n : I \to \mathbb{R}$ die Komponentenfunktionen von ϕ sind. Ist $n = 3$ und schreibt man $F(x, y, z)$, $\phi(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}$ und $\tilde{F}(t) = F(x(t), y(t), z(t))$, so gilt

$$\frac{d\tilde{F}}{dt} = \frac{\partial F}{\partial x}\frac{dx}{dt} + \frac{\partial F}{\partial y}\frac{dy}{dt} + \frac{\partial F}{\partial z}\frac{dz}{dt}.$$

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

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

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 $f : U \to V$ bijektiv ist und die Umkehrabbildung $f^{-1} : V \to U$ stetig differenzierbar ist mit

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

Bemerkung: Die Eigenschaft <u>regulär</u> 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

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

für jedes $\vec{x} \in U$.

ACHTUNG: Der Satz besagt nur, dass es **lokal** eine Umkehrfunktion zu f gibt. Auf ganz D muss dies nicht gelten!

Beispiel: Sei $D := (0, \infty) \times \mathbb{R} \subseteq \mathbb{R}^2$ und $f(r, \varphi) := \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix}$. Dann ist D offen und $f \in C^1(D, \mathbb{R}^2)$, sowie

$$f'(r, \varphi) = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}, \quad r > 0, \varphi \in \mathbb{R}.$$

Wegen $\det f'(r, \varphi) = r > 0$ ist $f'(r, \varphi)$ für alle $(r, \varphi) \in D$ regulär, und nach dem Umkehrsatz ist f lokal bijektiv.
Andererseits ist $f(r, \varphi) = f(r, \varphi + 2\pi)$ für alle $(r, \varphi) \in D$, dh $f : D \to \mathbb{R}^2$ ist nicht injektiv.
Es ist $f(D) = \mathbb{R}^2 \setminus \{(0, 0)\}$, aber $f : D \to \mathbb{R}^2 \setminus \{(0, 0)\}$ ist **nicht** bijektiv.

19.15. Der Satz über implizit definierte Funktionen:

Motivation: Die Gleichung $y - x^2 = 0$ lässt sich eindeutig nach y auflösen durch $y = x^2$. Die Auflösung von $x - y^2 = 0$ nach y ist hingegen global nicht mehr möglich.

Für gegebene (x_0, y_0) mit $x_0 - y_0^2 = 0$ und $y_0 > 0$ bzw. $y_0 < 0$ ist diese Auflösung lokal noch möglich (durch $y = \sqrt{x}$ bzw. $y = -\sqrt{x}$ für (x, y) nahe (x_0, y_0)). Jedoch ist in keiner Umgebung von $(x_0, y_0) = (0, 0)$ eine Auflösung durch eine **Funktion** möglich.

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

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

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** (dh $\det \frac{\partial f}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0) \neq 0$), so gibt es offene Umgebungen U von $\vec{x}_0$ und V von $\vec{y}_0$, sowie eine stetig differenzierbare Funktion $g : U \to V$ derart, dass für alle $\vec{x} \in U$ und $\vec{y} \in V$ gilt: $\det \frac{\partial f}{\partial \vec{y}}(\vec{x}, \vec{y}) \neq 0$ und

$$f(\vec{x}, \vec{y}) = \vec{0} \Longleftrightarrow \vec{y} = g(\vec{x}),$$

dh durch die Funktion $g : U \to V$ 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 (bzw. $\{(\vec{x}, \vec{y}) \in U \times V : f(\vec{x}, \vec{y}) = \vec{0}\}$ ist Graph der Funktion $g : U \to V$).

Bemerkung: Ableitungen von g kann man nach der Kettenregel aus

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

berechnen: es ist

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

also

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

Beispiele: (1) $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = y - x^2$. Hier ist $\frac{\partial f}{\partial y}(x_0, y_0) = 1$ für alle x_0, y_0 mit $f(x_0, y_0) = 0$.

(2) $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x - y^2$. Hier ist $\frac{\partial f}{\partial y}(x_0, y_0) = -2y_0$.

(3) Sei $f : \mathbb{R}^3 \to \mathbb{R}^2$, $f(x,y,z) = \begin{pmatrix} \sinh(yz) + (z-x)^2 - 1 \\ \cos^2(\pi y) + z - x^2/2 \end{pmatrix}$ und $(x_0, y_0, z_0) = (2, 0, 1)$. Hier ist $n = 3$, $p = 1$, $m = 2$, und es gilt

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

Da diese Matrix regulär ist und $f(2,0,1) = \binom{0}{0}$ gilt, ist in einer Umgebung von $(2,0,1)$ eine Auflösung des Gleichungssystems $f(x,y,z) = \binom{0}{0}$ nach y und z (jeweils als Funktion von x) möglich. Dh es gibt Funktionen $g_1, g_2 : U \to \mathbb{R}$, 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 $g = \binom{g_1}{g_2}$ ist dabei C^1 mit

$$g'(x) = \begin{pmatrix} g_1'(x) \\ g_2'(x) \end{pmatrix} = -\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}^{-1} \frac{\partial f}{\partial x}(2,0,1) = -\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

(4) Man kann den Satz verwenden, um zu Parameterdarstellungen von implizit gegebenen Flächen im $\mathbb{R}^3$ zu kommen. Wir betrachten als Beispiel die Gleichung $x^2 + y^2 + z^2 = 1$ der Oberfläche der Einheitskugel (dh der Einheitssphäre) S^2. Hier ist $f : \mathbb{R}^3 \to \mathbb{R}^3$, $f(x,y,z) = x^2 + y2 + z^2 - 1$, und $n = 3$, $p = 2$, $m = 1$. Es gilt

$$f'(x,y,z) = (2x\ 2y\ 2z) \quad \text{für alle } (x,y,z) \in \mathbb{R}^3.$$

Haben wir also einen Punkt $(x_0, y_0, z_0) \in S^2$ mit $z_0 \neq 0$, so finden wir offene Umgebungen U von (x_0, y_0), V von z_0 und auf U eine auflösende Funktion $g(x,y)$ derart, dass für $(x,y,z) \in U \times V$ gilt

$$(x,y,z) \in S^2 \iff z = g(x,y).$$

Hier kann man U, V und g konkret angeben: $g(x,y) = \operatorname{sgn}(z_0)\sqrt{1 - x^2 - y^2}$ auf $U := \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$ und z.B. $V = \operatorname{sgn}(z_0) \cdot (0, \infty)$. In Punkten $(x_0, y_0, z_0) \in S^2$ mit $z_0 = 0$ ist so eine lokale Auflösung nach z nicht möglich. Man kann dann aber zumindest nach einer der Variablen x oder y lokal auflösen.

Als Bonus:

Beweisidee 19.15 $\Rightarrow$ 19.14: Setze $F(\vec{x}, \vec{y}) = \vec{y} - f(\vec{x})$. Es ist $\frac{\partial F}{\partial \vec{x}}(\vec{x}, \vec{y}) = -f'(\vec{x})$ regulär. Also kann man $\vec{y} - f(\vec{x}) = \vec{0}$ lokal nach $\vec{x}$ auflösen. Die Auflösung ist f^{-1}.

Beweisidee 19.14 $\Rightarrow$ 19.15: Setze $F(\vec{x}, \vec{y}) = \binom{\vec{x}}{f(\vec{x}, \vec{y})}$. Dann ist $F'(\vec{x}_0, \vec{y}_0) = \begin{pmatrix} I_p & 0 \\ \frac{\partial f}{\partial \vec{x}} & \frac{\partial f}{\partial \vec{y}} \end{pmatrix}(\vec{x}_0, \vec{y}_0)$. Wegen $\det F'(\vec{x}_0, \vec{y}_0) = \det \frac{\partial f}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0) \neq 0$ ist $F'(\vec{x}_0, \vec{y}_0)$ regulär. Setze $g(\vec{x}) :=$ zweite Komponente von $F^{-1}(\vec{x}, \vec{0})$. Dann gilt:

$$\vec{y} = g(\vec{x}) \Leftrightarrow \begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix} = \begin{pmatrix} \vec{x} \\ g(\vec{x}) \end{pmatrix} = F^{-1}(\vec{x}, \vec{0}) \Leftrightarrow F(\vec{x}, \vec{y}) = \begin{pmatrix} \vec{x} \\ \vec{0} \end{pmatrix} \Leftrightarrow f(\vec{x}, \vec{y}) = \vec{0}.$$

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

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

und

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

Entsprechend definiert man für $k = 1, 2, \ldots, l+1$:

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

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

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

Zur allgemeinen Betrachtung von $(\vec{h} \cdot \nabla)^2 f(\vec{x}_0)$ definieren wir die Hesse-Matrix von f in $\vec{x}_0$:

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

Nach dem Satz von Schwarz ist $H_f(\vec{x}_0)$ symmetrisch, wenn $f \in C^2(D, \mathbb{R})$. Damit ist

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

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

$$f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + \frac{1}{1!}(\vec{h} \cdot \nabla) f(\vec{x}_0) + \frac{1}{2!}(\vec{h} \cdot \nabla)^2 f(\vec{x}_0) + \ldots + \frac{1}{l!}(\vec{h} \cdot \nabla)^l f(\vec{x}_0) + \frac{1}{(l+1)!}(\vec{h} \cdot \nabla)^{l+1} f(\vec{\xi}).$$

Bemerkung: (a) Für $l = 0$ erhält man einen mehrdimensionalen Mittelwertsatz.

(b) Der Ausdruck

$$T_{l,\vec{x}_0}(\vec{h}) := f(\vec{x}_0) + \sum_{k=1}^{l} \frac{(\vec{h} \cdot \nabla)^k f(\vec{x}_0)}{k!}$$

heißt <u>l-tes Taylorpolynom von f in $\vec{x}_0$</u>. Statt $\vec{h}$ schreibt man auch $\vec{x} - \vec{x}_0$.

(c) Für $l = 1$, $f \in C^2(D, \mathbb{R})$ erhalten wir, wenn $\vec{h} \in \mathbb{R}^n$ mit $S[\vec{x}_0, \vec{x}_0 + \vec{h}] \subseteq D$:

$$f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + \operatorname{grad} f(\vec{x}_0) \cdot \vec{h} + \frac{1}{2}\vec{h}^T H_f(\vec{\xi})\vec{h}$$

für ein $\vec{\xi} \in S[\vec{x}_0, \vec{x}_0 + \vec{h}] \subseteq D$. Schreibt man $\vec{x} - \vec{x}_0$ statt $\vec{h}$, so ist $\vec{x}_0 + \vec{h} = \vec{x}$.

Beispiel: Sei $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = e^{xe^y - 2}$ und $(x_0, y_0) = (2, 0)$. Dann ist $f \in C^2(\mathbb{R}^2, \mathbb{R})$ und $f_x = e^y f$, $f_y = xe^y f$, sowie $f_{xx} = e^{2y} f$, $f_{xy} = e^y(f + f_y) = (e^y + xe^{2y})f$, $f_{yy} = xe^y(f + f_y) = (xe^y + x^2e^{2y})f$. Wir erhalten

$$f(2, 0) = 1, f_x(2, 0) = 1, f_y(2, 0) = 2, f_{xx}(2, 0) = 1, f_{xy}(2, 0) = 3, f_{yy}(2, 0) = 6,$$

und damit ist das zweite Taylorpolynom von f in $(2, 0)$ gegeben durch

$$T_{2,(2,0)}(h_1, h_2) = 1 + 1 \cdot h_1 + 2h_2 + \frac{1}{2}h_1^2 + 3h_1h_2 + 3h_2^2$$

bzw., wenn man $h_1 = x - 2$ und $h_2 = y$ berücksichtigt, durch

$$1 + (x - 2) + 2y + \frac{1}{2}(x - 2)^2 + 3(x - 2)y + 3y^2.$$

19.17. Lokale Extremstellen: Sei $D \subseteq \mathbb{R}^n$ offen und $f : D \to \mathbb{R}$.

Definition: f hat in $\vec{x}_0 \in D$ ein <u>lokales Maximum</u> [bzw. <u>lokales Minimum</u>], falls es ein $\delta > 0$ so gibt, dass $K(\vec{x}_0, \delta) \subseteq D$ und $f(\vec{x}) \leq f(\vec{x}_0)$ [bzw. $f(\vec{x}) \geq f(\vec{x}_0)$] für alle $\vec{x} \in K(\vec{x}_0, \delta)$ gilt.

Ein <u>lokales Extremum</u> ist ein lokales Maximum oder ein lokales Minimum.

Satz über lokale Extremstellen: Sei $\vec{x}_0 \in D$.

(a) Hat f in $\vec{x}_0$ ein lokales Extremum und ist f in $\vec{x}_0$ partiell differenzierbar, so ist $\operatorname{grad} f(\vec{x}_0) = \vec{0}$.

(b) Sei $f \in C^2(D, \mathbb{R})$ und $\operatorname{grad} f(\vec{x}_0) = \vec{0}$. Dann gilt:

(i) Ist $H_f(\vec{x}_0)$ positiv definit, so hat f in $\vec{x}_0$ ein lokales Minimum.

(ii) Ist $H_f(\vec{x}_0)$ negativ definit, so hat f in $\vec{x}_0$ ein lokales Maximum.

(iii) Ist $H_f(\vec{x}_0)$ indefinit, so hat f in $\vec{x}_0$ **kein** lokales Extremum, sondern einen Sattelpunkt.

Beispiel für einen Sattelpunkt: Sei $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = xy$. Dann ist $\operatorname{grad} f(x,y) = \binom{y}{x}$ und $\operatorname{grad} f(x_0,y_0) = \binom{0}{0}$ gilt nur für $(x_0,y_0) = (0,0)$. Für die Hessematrix gilt $H_f(x,y) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Diese Matrix hat eine negative Determinante und ist daher indefinit (siehe 17.9). 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 $f(x,y) = xy < 0$.

Bemerkung: Trifft in (b) keiner der Fälle (i), (ii) oder (iii) zu, so ist $H_f(\vec{x}_0)$ (positiv oder negativ) **semidefinit** (vgl. 17.9), und es ist keine allgemeine Aussage möglich.

Nullstellen des Gradienten heißen auch kritische Punkte.

Im Beweis von (b) verwendet man Bemerkung 19.16(c), dh

$$f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + \vec{h}^T H_f(\vec{\xi})\vec{h},$$

und die Tatsache, dass man wegen $f \in C^2$ ein $\delta > 0$ so findet, dass $H_f(\vec{\xi})$ für $\vec{\xi} \in K(\vec{x}_0, \delta)$ dieselben Definitheitseigenschaften wie $H_f(\vec{x}_0)$ hat.

Alternativ kann man für $\vec{h} \in \mathbb{R}^n \setminus \{\vec{0}\}$ und kleine $t \in \mathbb{R}$ definieren $g(t) := f(\vec{x}_0 + t\vec{h})$. Dann ist $g \in C^2$ mit $g'(t) = \operatorname{grad} f(\vec{x}_0 + t\vec{h}) \cdot \vec{h}$, $g'(0) = 0$ und

$$g''(0) = \sum_{j=1}^{n} \Big(\frac{\partial f}{\partial x_j}\Big)'(\vec{x}_0)\vec{h}h_j = \sum_{j,k=1}^{n} \frac{\partial^2 f}{\partial x_k \partial x_j}(\vec{x}_0)h_j h_k = \vec{h}^T H_f(\vec{x}_0)\vec{h}.$$

Verwende nun HM I.

Beispiel: $D = \mathbb{R}^2$, $f(x,y) = x^3 - 12xy + 8y^3$. Hier gilt

$$f_x = 3x^2 - 12y, \quad f_y = -12x + 24y^2.$$

Wir formen äquivalent um:

$$\operatorname{grad} f(x,y) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff \begin{matrix} 3x^2 - 12y = 0 \\ -12x + 24y^2 = 0 \end{matrix} \iff \begin{matrix} x^2 = 4y \\ 2y^2 = x \end{matrix} \iff \begin{matrix} y^4 = y \\ 2y^2 = x \end{matrix} \iff$$
$$\iff \begin{matrix} y \in \{0,1\} \\ 2y^2 = x \end{matrix} \iff \begin{pmatrix} x \\ y \end{pmatrix} \in \Big\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \Big\}.$$

Wir berechnen die Hessematrix

$$H_f(x,y) = \begin{pmatrix} 6x & -12 \\ -12 & 48y \end{pmatrix}.$$

Es ist

$$\det H_f(0,0) = \det \begin{pmatrix} 0 & -12 \\ -12 & 0 \end{pmatrix} = -24 < 0,$$

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 17.9). Also hat f in $(2,1)$ ein lokales Minimum.

19.18. Extremwertaufgaben mit Nebenbedingungen: Sei $\emptyset \neq D \subseteq \mathbb{R}^n$ offen, $f \in C^1(D,\mathbb{R})$, $p \in \mathbb{N}$ mit $p < n$ und $h \in C^1(D,\mathbb{R}^p)$ mit Komponentenfunktionen $h_1, h_2, \ldots, h_p : D \to \mathbb{R}$, sowie $S := \{\vec{x} \in D : h(\vec{x}) = \vec{0}\}$.

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 $f(\vec{x}) \leq f(\vec{x}_0)$ [bzw. $f(\vec{x}) \geq f(\vec{x}_0)$] für alle $\vec{x} \in K(\vec{x}_0,\delta) \cap S$.

Zur Existenz: Eine Menge $K \subseteq \mathbb{R}^n$ heißt <u>kompakt</u>, wenn K abgeschlossen und beschränkt ist. Dabei heißt K <u>beschränkt</u>, falls es ein $M \in \mathbb{R}$ gibt mit $\|\vec{x}\| \leq M$ für alle $\vec{x} \in K$.

Satz: Ist $\emptyset \neq K \subseteq \mathbb{R}^n$ kompakt und $f : K \to \mathbb{R}$ stetig, so ist $f(K)$ kompakt und es gibt $\vec{a}, \vec{b} \in K$ mit

$$f(\vec{a}) \leq f(\vec{x}) \leq f(\vec{b}) \quad \text{für alle } \vec{x} \in K.$$

Beispiel: Seien $n = 3$, $p = 2$ und $D = \mathbb{R}^3$, sowie

$$f(x,y,z) = x + y + z, \qquad h(x,y,z) = \begin{pmatrix} x^2 + y^2 - 2 \\ x + z - 1 \end{pmatrix}.$$

Dann sind f, h stetig differenzierbar auf D 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 $(x_k, y_k, z_k) \to (x_0, y_0, z_0) \in \mathbb{R}^3$ gilt $(x_0, y_0, z_0) \in S$.

Die Menge S ist beschränkt: Sind $(x,y,z) \in S$, so gilt $x^2 + y^2 = 2$ und somit $|x| \leq \sqrt{2}$. Also ist $|z| = |1 - x| \leq 1 + |x| \leq 1 + \sqrt{2}$ und $\|(x,y,z)\| \leq \sqrt{2 + (1+\sqrt{2})^2} =: M$.

Also ist S kompakt, und nach dem Satz gibt es $\vec{a}, \vec{b} \in S$ mit $f(\vec{a}) \leq f(\vec{v}) \leq f(\vec{b})$ für alle $\vec{v} \in S$.

19.19. Multiplikatorenregel von Lagrange: Seien n, D, n, f, p, h und S wie in 19.18. Definiere die Funktion $F : D \times \mathbb{R}^p \to \mathbb{R}$ durch

$$F(\vec{x}, \lambda_1, \lambda_2, \ldots, \lambda_p) := f(\vec{x}) + \lambda_1 h_1(\vec{x}) + \lambda_2 h_2(\vec{x}) + \ldots + \lambda_p h_p(\vec{x})$$

für $\vec{x} \in D$ und $\lambda_1, \lambda_2, \ldots, \lambda_p \in \mathbb{R}$.

Satz: Hat f in $\vec{x}_0 \in D$ ein lokales Extremum unter der Nebenbedingung $h = \vec{0}$ und gilt

$$\text{Rang}\, \underbrace{h'(\vec{x}_0)}_{\in \mathbb{R}^{p\times n}} = p \quad [h'(\vec{x}_0) \text{ hat } \underline{\text{vollen}}, \text{ dh maximalen, Rang}],$$

so gibt es $\lambda_1^0, \lambda_2^0, \ldots, \lambda_p^0 \in \mathbb{R}$ (<u>Lagrangemultiplikatoren</u>) mit

$$\operatorname{grad} F(\vec{x}_0, \lambda_1^0, \lambda_2^0, \ldots, \lambda_p^0) = \vec{0}.$$

Bemerkung: (a) Beachte, dass die Zeilen von $h'(\vec{x}_0) \in \mathbb{R}^{p\times n}$ gerade $\operatorname{grad} h_1(\vec{x}_0)^T, \operatorname{grad} h_2(\vec{x}_0)^T, \ldots, \operatorname{grad} h_p(\vec{x}_0)^T$ sind. Die Voraussetzung an den Rang von $h'(\vec{x}_0)$ bedeutet also, dass $\operatorname{grad} h_1(\vec{x}_0), \operatorname{grad} h_2(\vec{x}_0), \ldots, \operatorname{grad} h_p(\vec{x}_0)$ linear unabhängig sind.

(b) Schreibt man $\vec{x}_0 = (x_1^0, x_2^0, \ldots, x_n^0)$, so ist die Bedingung $\operatorname{grad} F = \vec{0}$ ein Gleichungssystem mit $n+p$ Gleichungen für die $n+p$ Unbekannten $x_1^0, x_2^0, \ldots, x_n^0, \lambda_1^0, \lambda_2^0, \ldots, \lambda_p^0$, nämlich

$$0 = \frac{\partial f}{\partial x_k}(\vec{x}_0) + \sum_{j=1}^{p} \lambda_j^0 \frac{\partial h_j}{\partial x_k}(\vec{x}_0), \quad k = 1, 2, \ldots, n$$

(Gleichungen aus $\frac{\partial F}{\partial x_k} = 0$, $k = 1, 2, \ldots, n$) und

$$h_j(\vec{x}_0) = 0, \quad j = 1, 2, \ldots, p$$

(Gleichungen aus $\frac{\partial F}{\partial \lambda_j} = 0$, diese sind gleichbedeutend mit $\vec{x}_0 \in S$).

(c) Zur Bestimmung der Extrema versucht man, dieses Gleichungssystem zu lösen. Kann man den Satz anwenden (dh 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**!

Beispiel (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 $\operatorname{Rang} h'(x,y,z) < 2$ ist äquivalent zu $x = y = 0$, was jedoch für $(x,y,z) \in S$ wegen $x^2 + y^2 = 2$ nicht vorkommt. Also ist

$$\operatorname{Rang} h'(x,y,z) = 2 \quad \text{für alle } (x,y,z) \in S,$$

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

$$F_x = 1 + 2\lambda_1 x + \lambda_2,\ F_y = 1 + 2\lambda_1 y,\ F_z = 1 + \lambda_2,\ F_{\lambda_1} = x^2 + y^2 - 2,\ F_{\lambda_2} = x + z - 1.$$

Aus $\operatorname{grad} F = \vec{0}$ 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).

Wegen $f(0, \pm\sqrt{2}, 1) = 1 \pm \sqrt{2}$ ist $\vec{a} = (0, -\sqrt{2}, 1)$ die Minimal- und $\vec{b} = (0, \sqrt{2}, 1)$ 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 $f(\vec{a}) = 1 - \sqrt{2}$.

Erläuterung zum Satz (als Bonus): Wir kehren zur allgemeinen Situation zurück, dh n, D, f, p und h wie in 19.18, $S = \{\vec{x} \in D : h(\vec{x}) = \vec{0}\}$. Weiter sei $\vec{x}_0 \in S$ mit

$$\operatorname{Rang} h'(\vec{x}_0) = p.$$

Wir zeigen, wie man S lokal in der Nähe von $\vec{x}_0$ parametrisieren kann. Wir setzen $q := n - p$ und finden p linear unabhängige Spalten von $h'(\vec{x}_0) \in \mathbb{R}^{p\times n}$. Dies seien o.B.d.A. die letzten p Spalten. Wir schreiben $\vec{x} = (\vec{y}, \vec{z})$ mit $\vec{y} = (x_1, \ldots, x_q) \in \mathbb{R}^q$ und $\vec{z} = (x_{q+1}, \ldots, x_{q+p}) \in \mathbb{R}^p$, sowie entsprechend $\vec{x}_0 = (\vec{y}_0, \vec{z}_0)$. Dann ist $\frac{\partial h}{\partial \vec{z}}(\vec{x}_0) \in \mathbb{R}^{p\times p}$ regulär (da die Matrix Rang p hat), und nach 19.15 gibt es offene Umgebungen $U \subseteq \mathbb{R}^q$ von $\vec{y}_0$, $V \subseteq \mathbb{R}^p$ von $\vec{z}_0$ und eine C^1-Funktion $g : U \to V$ mit $g(\vec{y}_0) = \vec{z}_0$ so, dass für alle $(\vec{y}, \vec{z}) \in U \times V$ gilt:

$$h(\vec{y}, \vec{z}) = \vec{0} \iff \vec{z} = g(\vec{y}).$$

Die Funktion $U \to \mathbb{R}^n$, $\vec{y} \mapsto \binom{\vec{y}}{g(\vec{y})}$ ist also eine Parametrisierung von $S \cap (U \times V)$. Außerdem haben wir

$$g'(\vec{y}_0) = -\Big(\frac{\partial h}{\partial \vec{z}}(\vec{x}_0)\Big)^{-1} \frac{\partial h}{\partial \vec{y}}(\vec{x}_0).$$

Hat nun f zusätzlich in $\vec{x}_0 = (\vec{y}_0, g(\vec{y}_0))$ eine lokale Extremstelle unter der Nebenbedingung $h = \vec{0}$, so hat die Funktion $U \to \mathbb{R}$, $\vec{y} \mapsto f(\vec{y}, g(\vec{y}))$, in $\vec{y}_0$ eine lokale Extremstelle. Nach Satz 19.17(a) und der Kettenregel ist also

$$\vec{0} = \frac{\partial f}{\partial \vec{y}}(\vec{x}_0) + \frac{\partial f}{\partial \vec{z}}(\vec{x}_0) g'(\vec{y}_0) = \frac{\partial f}{\partial \vec{y}}(\vec{x}_0) - \frac{\partial f}{\partial \vec{z}}(\vec{x}_0)\Big(\frac{\partial h}{\partial \vec{z}}(\vec{x}_0)\Big)^{-1} \frac{\partial h}{\partial \vec{y}}(\vec{x}_0).$$

Wir setzen $\vec{\lambda}_0^T := \frac{\partial f}{\partial \vec{z}}(\vec{x}_0)\Big(\frac{\partial h}{\partial \vec{z}}(\vec{x}_0)\Big)^{-1} \in \mathbb{R}^{1\times p}$ und haben

$$\frac{\partial f}{\partial \vec{y}}(\vec{x}_0) = \vec{\lambda}_0^T \frac{\partial h}{\partial \vec{y}}(\vec{x}_0), \qquad \frac{\partial f}{\partial \vec{z}}(\vec{x}_0) = \frac{\partial f}{\partial \vec{z}}(\vec{x}_0)\Big(\frac{\partial h}{\partial \vec{z}}(\vec{x}_0)\Big)^{-1} \frac{\partial h}{\partial \vec{z}}(\vec{x}_0) = \vec{\lambda}_0^T \frac{\partial h}{\partial \vec{z}}(\vec{x}_0),$$

also

$$f'(\vec{x}_0) = \vec{\lambda}_0^T h'(\vec{x}_0), \quad \text{bzw.} \quad \operatorname{grad} f(\vec{x}_0) = h'(\vec{x}_0)^T \vec{\lambda}_0.$$

Das ist die Aussage des Satzes.

19.20. Rotation, Divergenz, Laplace: Sei $\emptyset \neq D \subseteq \mathbb{R}^n$. Eine Funktion $f : D \to \mathbb{R}$ heißt Skalarfeld (auf D) und eine Funktion $\vec{v} : D \to \mathbb{R}^n$ heißt Vektorfeld (auf D).

Bemerkung: Ist $f \in C^1$ ein Skalarfeld auf D, so ist $\operatorname{grad} f = \nabla f$ ein Vektorfeld auf D.

Partielle Ableitungen schreiben wir im folgenden als

$$\partial_j := \frac{\partial}{\partial x_j}, \quad j = 1, 2, \ldots, n.$$

Definition: Sei $\vec{v} : D \to \mathbb{R}^n$ ein C^1-Vektorfeld auf D mit Komponentenfunktionen $v_1, v_2, \ldots, v_n : D \to \mathbb{R}$. Dann definiert man die Divergenz von $\vec{v}$ durch

$$\operatorname{div} \vec{v} := \nabla \cdot \vec{v} := \partial_1 v_1 + \partial_2 v_2 + \ldots + \partial_n v_n = \sum_{j=1}^{n} \partial_j v_j$$

und im Fall $n = 3$ die Rotation von $\vec{v}$ durch

$$\operatorname{rot} \vec{v} := \nabla \times \vec{v} := \begin{pmatrix} \partial_2 v_3 - \partial_3 v_2 \\ \partial_3 v_1 - \partial_1 v_3 \\ \partial_1 v_2 - \partial_2 v_1 \end{pmatrix}.$$

Bemerkung: (a) $\operatorname{div} \vec{v}$ ist ein Skalarfeld auf D und $\operatorname{rot} \vec{v}$ ist ein Vektorfeld auf D.

(b) Vektorfelder mit $\operatorname{div} \vec{v} = 0$ heißen quellenfrei und Vektorfelder mit $\operatorname{rot} \vec{v} = \vec{0}$ heißen wirbelfrei.

(c) Für C^2-Skalarfelder $f : D \to \mathbb{R}$ definiert man den Laplaceoperator durch

$$\Delta f := \operatorname{div} \operatorname{grad} f := \nabla \cdot \nabla f := \sum_{j=1}^{n} \partial_j^2 f,$$

und für C^2-Vektorfelder $\vec{v} : D \to \mathbb{R}^n$ durch

$$\Delta \vec{v} = \begin{pmatrix} \Delta v_1 \\ \Delta v_2 \\ \vdots \\ \Delta v_n \end{pmatrix}.$$

Beispiele: (1) Sei $\vec{v} : \mathbb{R}^n \to \mathbb{R}^n$, $\vec{v}(\vec{x}) := \vec{x}$. Dann gilt $\operatorname{div} \vec{v}(\vec{x}) = n$ für jedes $\vec{x} \in \mathbb{R}^n$. Im Falle $n = 3$ gilt $\operatorname{rot} \vec{v} = \vec{0}$ auf $\mathbb{R}^3$.

(2) Sei $\vec{v} : \mathbb{R}^3 \to \mathbb{R}^3$, $\vec{v}(x,y,z) := \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}$. Dann ist $\operatorname{rot} \vec{v}(x,y,z) = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$ für alle $(x,y,z) \in \mathbb{R}^3$.

(3) Sei $f : \mathbb{R}^n \to \mathbb{R}$ eine C^2-Funktion und $A \in \mathbb{R}^{n\times n}$ orthogonal, sowie $g := f \circ A$, dh $g(\vec{x}) = f(A\vec{x})$. Dann gilt $\Delta g = (\Delta f) \circ A$:

Wir schreiben hier $D\vec{v}$ statt $\vec{v}\,'$ für die Ableitung eines Vektorfelds $\vec{v}$. Es gilt dann

$$\operatorname{div} \vec{v} = \operatorname{Spur}(D\vec{v}).$$

Nun haben wir nach der Kettenregel (siehe 19.13):

$$\begin{aligned} \Delta g &= \operatorname{div}(\nabla g) = \operatorname{div}(A^T(\nabla f) \circ A) \\ &= \operatorname{Spur}(D(A^T(\nabla f)\circ A)) = \operatorname{Spur}(A^T D((\nabla f)\circ A)) \\ &= \operatorname{Spur}(A^T((D(\nabla f))\circ A)A). \end{aligned}$$

Nach Satz 17.5(a) ist dies

$$= \operatorname{Spur}(((D(\nabla f))\circ A)\underbrace{AA^T}_{=I_n}) = \operatorname{Spur}((D(\nabla f))\circ A) = (\operatorname{div}(\nabla f))\circ A = (\Delta f)\circ A.$$

Bedeutung: Ist $Jf := f \circ A$, so gilt $\Delta f = J^{-1}\Delta J f$ für jede C^2-Funktion $f : \mathbb{R}^n \to \mathbb{R}$, dh Δ ist invariant unter orthogonalen Koordinatentransformationen.

(4) Der Laplace-Operator in Polarkoordinaten: Sei $u = u(x,y)$ eine C^2-Funktion. Wir verwenden Polarkoordinaten $x = r\cos\varphi$, $y = r\sin\varphi$ und schreiben $v(r,\varphi) := u(r\cos\varphi, r\sin\varphi)$. Nun wenden wir $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ auf $u(x,y) = v(r(x,y),\varphi(x,y))$ an, wollen dies aber mittels v und den Variablen r,φ ausdrücken. Nach Ketten- und Produktregel haben wir:

$$\begin{aligned} \partial_x v(r,\varphi) &= v_r r_x + v_\varphi \varphi_x \\ \partial_x^2 v(r,\varphi) &= \partial_x(v_r) r_x + v_r r_{xx} + \partial_x(v_\varphi)\varphi_x + v_\varphi\varphi_{xx} \\ &= v_{rr}(r_x)^2 + v_{r\varphi} r_x\varphi_x + v_r r_{xx} + v_{\varphi r} r_x \varphi_x + v_{\varphi\varphi}(\varphi_x)^2 + v_\varphi \varphi_{xx} \\ \Delta v(r,\varphi) &= v_r(r_{xx}+r_{yy}) + v_{rr}((r_x)^2+(r_y)^2) + 2v_{r\varphi}(r_x\varphi_x + r_y\varphi_y) \\ &\quad + v_\varphi(\varphi_{xx}+\varphi_{yy}) + v_{\varphi\varphi}((\varphi_x)^2+(\varphi_y)^2). \end{aligned}$$

Ableitungen von r und φ nach x und y berechnen wir aus $r^2 = x^2+y^2$ und $\tan\varphi = y/x$:

$$2rr_x = 2x, \quad \text{also} \quad r_x = x/r = \cos\varphi,\ r_y = y/r = \sin\varphi,\ (r_x)^2 + (r_y)^2 = 1.$$

Weiter ist $rr_{xx} + (r_x)^2 = 1$, also

$$r_{xx} = \frac{1-(r_x)^2}{r},\ r_{yy} = \frac{1-(r_y)^2}{r}, r_{xx}+r_{yy} = \frac{2-(r_x)^2-(r_y)^2}{r} = \frac{1}{r}.$$

Für die Ableitungen von φ erhalten wir

$$(1+\tan^2\varphi)\varphi_x = -y/x^2, \quad (1+\tan^2\varphi)\varphi_y = 1/x$$

und wegen $1+\tan^2\varphi = (\cos\varphi)^{-2}$:

$$\varphi_x = -\cos^2\varphi\,\frac{y}{x^2} = -\frac{\sin\varphi}{r}, \quad \varphi_y = \cos^2\varphi\,\frac{1}{x} = \frac{\cos\varphi}{r}.$$

Wir lesen ab:

$$(\varphi_x)^2 + (\varphi_y)^2 = \frac{1}{r^2}, \quad r_x\varphi_x + r_y\varphi_y = 0.$$

Schließlich ist

$$\varphi_{xx} + \varphi_{yy} = -\cos\varphi\cdot\varphi_x\cdot\frac{1}{r} + \frac{\sin\varphi}{r^2}\cdot r_x - \sin\varphi\cdot\varphi_y\cdot\frac{1}{r} - \frac{\cos\varphi}{r^2}\cdot r_y = 0.$$

Zusammengefasst haben wir also

$$\Delta v(r,\varphi) = v_{rr} + \frac{1}{r}v_r + \frac{1}{r^2}v_{\varphi\varphi}$$

als Formel für den Laplace-Operator in Polarkoordinaten.

19.21. Rechenregeln:

Produktregeln: Seien $f, g : D \to \mathbb{R}$ und $\vec{v} : D \to \mathbb{R}^n$ C^1-Funktionen. Dann gilt:

$$\begin{aligned}
\nabla(fg) &= g\nabla f + f\nabla g \\
\nabla\cdot(f\vec{v}) &= f(\nabla\cdot\vec{v}) + (\nabla f)\cdot\vec{v} \\
\nabla\times(f\vec{v}) &= f(\nabla\times\vec{v}) + (\nabla f)\times\vec{v} \quad (n=3) \\
\Delta(fg) &= (\Delta f)g + 2\nabla f\cdot\nabla g + f(\Delta g) \quad (f,g\in C^2).
\end{aligned}$$

Zum Nachrechnen wende man die eindimensionale Produktregel aus HM I auf die einzelnen partiellen Ableitungen an:

$$\partial_j(\phi\psi) = (\partial_j\phi)\psi + \phi(\partial_j\psi).$$

Hintereinanderausführung: Sei $D \subseteq \mathbb{R}^3$. Sind $f : D \to \mathbb{R}$ und $\vec{v} : D \to \mathbb{R}^3$ C^2-Funktionen, so gilt:

$$\begin{aligned}
\operatorname{rot}(\operatorname{grad} f) &= \vec{0}, \qquad \text{dh} \quad \nabla\times(\nabla f) = \vec{0} \\
\operatorname{div}(\operatorname{rot}\vec{v}) &= 0 \qquad \text{dh} \quad \nabla\cdot(\nabla\times\vec{v}) = 0
\end{aligned}$$

Die beiden Regeln folgen aus dem Satz von Schwarz.

20 Kurvenintegrale und Integralsätze im $\mathbb{R}^2$

20.1. Kurvenintegrale von Skalarfeldern: Sei $D \subseteq \mathbb{R}^n$ offen und $f : D \to \mathbb{R}$ stetig. Für eine reguläre Kurve $\gamma : [a,b] \to D$ ((regulär bedeutet, dass $\gamma'(t) \neq \vec{0}$, für alle $t \in [a,b]$)) setzt man

$$\int_\gamma f\, ds := \int_a^b f(\gamma(t)) \|\dot{\gamma}(t)\|\, dt$$

(beachte, dass $\gamma(t)$ und $\dot{\gamma}(t)$ Vektoren aus $\mathbb{R}^n$ sind).

Bemerkung: (a) Mit diesem Kurvenintegral berechnet man Flächeninhalt von Flächen die nicht Flach sind.

(b) Für $f = 1$ ist $\int_\gamma ds$ die Länge von γ (vgl. 19.6).

(c) Ist γ wie oben und setzt man $\rho : [a,b] \to D$, $\rho(t) := \gamma(a+b-t)$, so durchläuft ρ die Spur von γ in umgekehrter Richtung. Diese Kurve wird auch mit $-\gamma$ bezeichnet. Es gilt dann

$$\int_{-\gamma} f\, ds = \int_\gamma f\, ds.$$

(d) Sind $\gamma_j : [a_{j-1}, a_j] \to D$, $j = 1, 2, \ldots, m$, reguläre Kurven mit $\gamma_j(a_j) = \gamma_{j+1}(a_j)$, so bezeichnet man auch $\gamma : [a_0, a_m] \to D$, $\gamma(t) := \gamma_j(t)$, falls $t \in [a_{j-1}, a_j]$, als Kurve. In den Punkten $a_1, a_2, \ldots, a_{m-1}$ muss γ nicht differenzierbar sein, dh die rechts- und linksseitigen Ableitungen in diesen Punkten müssen nicht übereinstimmen. Man schreibt $\gamma = \gamma_1 + \gamma_2 + \ldots + \gamma_m$ und definiert

$$\int_\gamma f\, ds := \sum_{j=1}^m \int_{\gamma_j} f\, ds.$$

(e) Ist γ geschlossen (d.h. $\gamma(a) = \gamma(b)$ in einfachen Worten sie aufhört wo sie anfängt), so schreibt man auch

$$\int_\gamma f\, ds = \oint_\gamma f\, ds.$$

20.2. Kurvenintegrale von Vektorfeldern: Sei $D \subseteq \mathbb{R}^n$ offen und $\vec{v} : D \to \mathbb{R}^n$ stetig. Für eine reguläre Kurve $\gamma : [a,b] \to D$ (regulär bedeutet, dass $\gamma'(t) \neq \vec{0}$, für alle $t \in [a,b]$) setzt man

$$\int_\gamma \vec{v} \cdot d\vec{s} := \int_a^b \vec{v}(\gamma(t)) \cdot \dot{\gamma}(t)\, dt$$

(beachte, dass $\gamma(t)$ und $\dot{\gamma}(t)$ Vektoren aus $\mathbb{R}^n$ sind und dass $\cdot$ das Skalarprodukt im $\mathbb{R}^n$ bezeichnet).

Wichtige Bemerkung: Interpretation dieses Kuvrenintegrals: Ist $\gamma(t)$ der Ort eines sich bewegenden Punktobjektes in D und $\vec{v}$ eine Kraft an das Objekt, dann ist dieses Kurvenintegral die Arbeit der Kraft $\vec{v}$ entlang der Kurve γ.

Beispiele: (1) Es gilt
$$\int_{-\gamma} \vec{v} \cdot d\vec{s} = -\int_{\gamma} \vec{v} \cdot d\vec{s},$$
dh das Kurvenintegral eines Vektorfeldes ändert bei Orientierungsumkehr der Kurve das Vorzeichen (vgl. aber mit 20.1 Bemerkung (c)).

(2) Ist $f : D \to \mathbb{R}$ ein C^1-Skalarfeld und $\gamma : [a,b] \to D$ eine Kurve, so gilt
$$\int_{\gamma} \nabla f \cdot d\vec{s} = f(\gamma(b)) - f(\gamma(a)).$$
Das liegt an
$$\frac{d}{dt} f(\gamma(t)) = \underbrace{f'(\gamma(t))}_{(\nabla f)(\gamma(t))^T} \dot{\gamma}(t) = (\nabla f)(\gamma(t)) \cdot \dot{\gamma}(t)$$
und dem Hauptsatz (aus HM I):
$$f(\gamma(b)) - f(\gamma(a)) = \int_a^b \frac{d}{dt} f(\gamma(t))\, dt.$$

(3) Sei $\vec{v} : \mathbb{R}^2 \to \mathbb{R}^2$, $\vec{v}(x,y) = \binom{-y}{x}$ und $\gamma(t) := \binom{\cos t}{\sin t}$, $t \in [0, 2\pi]$. Dann ist $\dot{\gamma}(t) = \binom{-\sin t}{\cos t}$ und
$$\int_{\gamma} \vec{v} \cdot d\vec{s} = \int_0^{2\pi} \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \cdot \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} dt = 2\pi.$$
Sei $\vec{w} : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}^2$, $\vec{w}(x,y) = \binom{-y/(x^2+y^2)}{x/(x^2+y^2)}$. Auch hier gilt
$$\int_{\gamma} \vec{w} \cdot d\vec{s} = \int_0^{2\pi} \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \cdot \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} dt = 2\pi.$$

20.3. Gebiete und einfach zusammenhängende Mengen: Eine Teilmenge $K \subseteq \mathbb{R}^n$ heißt <u>konvex</u>, falls für je zwei Punkte $\vec{x}, \vec{y} \in K$ gilt: $S[\vec{x}, \vec{y}] \subseteq K$ ($S[\vec{x}, \vec{y}]$ ist 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: Eine offene und wegzusammenhängende Teilmenge $G \subseteq \mathbb{R}^n$ heißt ein <u>Gebiet</u>.

Beispiele: Ist G offen und konvex, so ist G ein Gebiet. $\mathbb{R}^n \setminus \{\vec{0}\}$ ist ein Gebiet. $\{(x,y) \in \mathbb{R}^2 : x \neq 0\}$ ist kein Gebiet.

Definition: Ein Gebiet G heißt <u>einfach zusammenhängend</u>, wenn es für jede einfach geschlossene Kurve $\gamma : [0,1] \to G$ einen Punkt gibt $\vec{c} \in G$ gibt, und eine stetige Funktion $g : [0,1] \times [0,1] \to G$ mit $g(0,t) = \gamma(t)$, für alle $t \in [0,1]$ und $g(1,t) = \vec{c}$ für alle $t \in [0,1]$.

Dh: Man kann jede einfach geschlossene Kurve in G zu einem Punkt zusammenziehen. Für Teilmengen $G \subseteq \mathbb{R}^2$ bedeutet dies anschaulich: "G hat keine Löcher."

Beispiele: $\mathbb{R}^2 \setminus \{\vec{0}\}$ ist nicht einfach zusammenhängend. Konvexe Mengen sind einfach zusammenhängend.

Gibt es in G einen Punkt $\vec{c} \in G$ mit $S[\vec{x}, \vec{c}] \subseteq G$ für jedes $\vec{x} \in G$ (solche Gebiete heißen sternförmig), so ist G einfach zusammenhängend.

$\mathbb{R}^3 \setminus \{\vec{0}\}$ ist einfach zusammenhängend, aber $\mathbb{R}^3 \setminus \{(0,0,z) : z \in \mathbb{R}\}$ ist **nicht** einfach zusammenhängend.

20.4. Kurvenintegrale und Potentialfelder:

Potentialfelder Sei $\emptyset \neq D \subseteq \mathbb{R}^n$ offen. Ein stetiges Vektorfeld $\vec{v} : D \to \mathbb{R}^n$ heißt <u>Potentialfeld</u> (oder <u>Gradientenfeld</u>, <u>konservatives Feld</u>), falls ein C^1-Skalarfeld (ein <u>Potential</u>) $f : D \to \mathbb{R}$ existiert mit $\vec{v} = \nabla f$ auf D.

Bemerkung: (a) Wegen 19.21 ist ein C^1-Potentialfeld $\vec{v} : D \to \mathbb{R}^3$ wirbelfrei in D.

(b) Nach dem Satz von Schwarz gilt für ein C^1-Potentialfeld $\vec{v} : D \to \mathbb{R}^n$ mit Komponenten $v_1, v_2, \ldots, v_n : D \to \mathbb{R}$:

$$\underbrace{\vec{v}\,'(\vec{x})}_{\in \mathbb{R}^{n \times n}} = \vec{v}\,'(\vec{x})^T, \quad \vec{x} \in D,$$

dh für alle $j, k = 1, \ldots, n$ ist

$$\partial_j v_k = \partial_k v_j \quad \text{auf } D.$$

Satz 1: Sei $G \subseteq \mathbb{R}^n$ ein Gebiet und $\vec{v} : G \to \mathbb{R}^n$ ein stetiges Vektorfeld. Dann sind äquivalent:

(i) $\vec{v}$ ist Potentialfeld in G.

(ii) Für je zwei Punkte $\vec{x}, \vec{y} \in G$ ist $\int_\gamma \vec{v} \cdot d\vec{s}$ unabhängig von der Kurve $\gamma : [a,b] \to G$ mit $\gamma(a) = \vec{x}$, $\gamma(b) = \vec{y}$.

(iii) Für jede geschlossene Kurve $\gamma : [a,b] \to G$ gilt

$$\oint_\gamma \vec{v} \cdot d\vec{s} = 0.$$

Satz 2: Ist G einfach zusammenhängend und $\vec{v} : G \to \mathbb{R}^n$ ein C^1-Vektorfeld, so ist außerdem äquivalent:

(iv) Für alle $j, k = 1, \ldots, n$ gilt: $\partial_j v_k = \partial_k v_j$ auf G (Verträglichkeitsbedingung).

Für $G \subseteq \mathbb{R}^3$ gilt also, ist G einfach zusammenhängend und ist $\operatorname{rot} \vec{v} = \vec{0}$ in G, so ist $\vec{v}$ ein Potentialfeld in G.

Beispiele: (1) Sei $\vec{v}(x,y) = \binom{2xy}{x^2+1}$ auf $G := \mathbb{R}^2$. G ist konvex, also einfach zusammenhängend. Hier gilt

$$\partial_y v_1 = \partial_y(2xy) = 2x = \partial_x(x^2+1) = \partial_x v_2 \quad \text{auf } \mathbb{R}^2.$$

Nach Satz 2 ist $\vec{v}$ ein Potentialfeld auf $\mathbb{R}^2$. Berechnung eines Potentials f etwa durch den Ansatz:

$$f(x,y) = \int v_1(x,y)\,dx + \phi(y) = \int 2xy\,dx + \phi(y) = x^2y + \phi(y).$$

Nun muss $\partial_y(x^2y + \phi(y)) = v_2(x,y) = x^2 + 1$ sein, also $x^2 + \phi'(y) = x^2 + 1$, dh $\phi'(y) = 1$ und etwa $\phi(y) = y$. Ein Potential f auf $\mathbb{R}^2$ ist also gegeben durch $f(x,y) = x^2y + y$.

(2) Sei $\vec{w}(x,y) = \begin{pmatrix} \frac{-y}{x^2+y^2} \\ \frac{x}{x^2+y^2} \end{pmatrix}$ für $(x,y) \in G := \mathbb{R}^2 \setminus \{(0,0)\}$. Dann ist $\vec{w}$ ein C^1-Vektorfeld, das die Bedingung (iv) erfüllt: Es ist

$$\partial_y w_1 = \frac{y^2 - x^2}{(x^2+y^2)^2} = \partial_x w_2.$$

Für die geschlossene Kurve γ aus Beispiel 20.2(3) gilt aber

$$\int_\gamma \vec{w} \cdot d\vec{s} = 2\pi.$$

Also ist $\vec{w}$ nach Satz 1 kein Potentialfeld auf $\mathbb{R}^2 \setminus \{(0,0)\}$. $\mathbb{R}^2 \setminus \{(0,0)\}$ ist **nicht** einfach zusammenhängend.

Auf z.B. $\{(x,y) \in \mathbb{R}^2 : x > 0\}$ (was konvex, also einfach zusammenhängend ist) ist $\vec{w}$ aber nach Satz 2 ein Potentialfeld und $f(x,y) := \arctan(y/x)$ definiert ein Potential.

20.5. Integration über Teilmengen im $\mathbb{R}^2$: Sei $R := [a,b] \times [c,d] \subseteq \mathbb{R}^2$ ein Rechteck. Man erklärt Integrierbarkeit und Integral für beschränkte Funktionen $f : R \to \mathbb{R}$ ähnlich wie in 10.1 und 10.2 (HM I), indem man Zerlegungen $a = x_0 < x_1 < \ldots < x_n = b$ und $c = y_0 < y_1 < \ldots < y_m = d$ und die daraus resultierenden Zerlegungen $[x_{j-1}, x_j] \times [y_{k-1}, y_k]$, $j = 1, \ldots, n$, $k = 1, \ldots, m$ von R betrachtet. Ist f integrierbar, schreiben wir

$$\iint_R f(x,y)\,d(x,y)$$

für das Integral.

Satz 1: Sei $f : R \to \mathbb{R}$ stetig. Dann ist f integrierbar und es gilt

$$\iint\limits_R f(x,y)\,d(x,y) = \int_a^b \int_c^d f(x,y)\,dy\,dx = \int_c^d \int_a^b f(x,y)\,dx\,dy.$$

Beispiel:

$$\iint\limits_{[0,1]\times[0,1]} xy\,d(x,y) = \int_0^1 \Big(\int_0^1 xy\,dy\Big)\,dx = \int_0^1 \Big[x\frac{y^2}{2}\Big]_{y=0}^{y=1}\,dx = \int_0^1 \frac{x}{2}\,dx = \Big[\frac{x^2}{4}\Big] = \frac{1}{4}.$$

Bemerkung: Ist $B \subseteq \mathbb{R}^2$ beschränkt und $f : B \to \mathbb{R}$ beschränkt. Dann heißt f über B integrierbar, falls es ein Rechteck R gibt mit $B \subseteq R$ und die Funktion $f_0 : R \to \mathbb{R}$, $f_0(x,y) := \begin{cases} f(x,y) & ,(x,y) \in B \\ 0 & ,(x,y) \notin B \end{cases}$, integrierbar ist. Man setzt

$$\iint\limits_B f(x,y)\,d(x,y) = \iint\limits_R f_0(x,y)\,d(x,y).$$

Satz 2: Ist $f : B \to \mathbb{R}$ stetig, wobei

$$B := \{(x,y) : x \in [a,b], y \in [c(x), d(x)]\}$$

und $c, d : [a,b] \to \mathbb{R}$ stetige Funktionen sind, so ist f über B integrierbar und

$$\iint\limits_B f(x,y)\,d(x,y) = \int_a^b \Big(\int_{c(x)}^{d(x)} f(x,y)\,dy\Big)\,dx.$$

Entsprechendes gilt, wenn die Rollen von x und y vertauscht werden, dh für Mengen

$$C = \{(x,y) : y \in [c,d], x \in [a(y), b(y)]\},$$

wobei $a, b : [c,d] \to \mathbb{R}$ stetig sind. Es ist dann

$$\iint\limits_C f(x,y)\,d(x,y) = \int_c^d \Big(\int_{a(y)}^{b(y)} f(x,y)\,dx\Big)\,dy.$$

Beispiele: (1) Sei $B := \{(x,y) : x, y \geq 0, x^2 + y^2 \leq 1\}$. Dann ist

$$\iint\limits_B x\,d(x,y) = \int_0^1 \int_0^{\sqrt{1-x^2}} x\,dy\,dx = \int_0^1 x\sqrt{1-x^2}\,dx = \Big[-\frac{1}{3}(1-x^2)^{3/2}\Big]_0^1 = \frac{1}{3}.$$

(2) Sei $B := \{(x,y) : y \geq 0, x^2 + y^2 \leq 1\}$. Dann ist

$$\iint\limits_B d(x,y) = \int_{-1}^{1} \Big(\int_0^{\sqrt{1-x^2}} dy \Big)\, dx = \int_{-1}^{1} \sqrt{1-x^2}\, dx = \frac{\pi}{2}$$

der Flächeninhalt von B.

Bemerkung: (a) In Satz 2 ist B **abgeschlossen**. Integriert man nur über die **offene** Menge

$$\operatorname{inn}(B) := \{(x,y) : a < x < b, c(x) < y < d(x)\},$$

(das Innere von B), so ändert sich das Integral nicht (f ist aber als stetig auf B vorausgesetzt!).

(b) Lässt sich eine abgeschlossene Menge A schreiben als $A = \bigcup_{j=1}^{n} A_j$, wobei jedes A_j von der Form der obigen Mengen B oder C ist und $\operatorname{inn}(A_1), \operatorname{inn}(A_2), \ldots, \operatorname{inn}(A_n)$ paarweise disjunkt sind, so gilt für stetiges $f : A \to \mathbb{R}$:

$$\iint\limits_A f(x,y)\, d(x,y) = \sum_{j=1}^{n} \iint_{A_j} f(x,y)\, d(x,y).$$

(c) Wir werden im folgenden über Gebiete $G \subseteq \mathbb{R}^n$ integrieren. Die Funktionen f werden stetig sein auf

$$\overline{G} := G \cup \partial G,$$

wobei ∂G (der Rand von G) die Menge aller Randpunkte von G ist. Dabei heißt ein $\vec{x} \in \mathbb{R}^n$ Randpunkt von G, falls $K(\vec{x}, \varepsilon) \cap G \neq \emptyset$ und $K(\vec{x}, \varepsilon) \cap (\mathbb{R}^n \setminus G) \neq \emptyset$ für jedes $\varepsilon > 0$ (“in jeder Umgebung von $\vec{x}$ liegen Punkte aus G und Punkte aus $\mathbb{R}^n \setminus G$”). Die Menge $\overline{G}$ ist abgeschlossen.

Beispiele: Es gilt etwa $\partial K(\vec{x}_0, r) = \{\vec{x} \in \mathbb{R}^n : \|\vec{x} - \vec{x}_0\| = r\}$ und $\partial(\mathbb{R}^n \setminus \{\vec{0}\}) = \{\vec{0}\}$. Für $G = \operatorname{inn}(B)$ von oben gilt

$$\partial G = \{(x,y) : x \in [a,b], y \in \{c(x), d(x)\}\} \cup \{(a,y) : y \in [c(a), d(a)]\} \cup \{(b,y) : y \in [c(b), d(b)]\}.$$

20.6. Gaußscher Integralsatz im $\mathbb{R}^2$: Sei $G \subseteq \mathbb{R}^2$ einfach zusammenhängend. Wie nehmen an, dass der Rand ∂G von G die Spur einer einfach geschlossenen regulären Kurve γ ist. Die Orientierung von γ sei so, dass G “links von γ liegt” (positive orientierte Kurve). In dieser Situation schreibt man statt $\oint_\gamma$ suggestiv auch $\oint_{\partial G}$.

Satz: Sei $D \subseteq \mathbb{R}^2$ offen mit $\overline{G} \subseteq D$ und $\vec{v} = \binom{v_1}{v_2} \in C^1(D, \mathbb{R}^2)$. Dann gilt:

$$\oint_{\partial G} \vec{v} \cdot d\vec{s} = \iint\limits_G (\partial_1 v_2(x,y) - \partial_2 v_1(x,y))\, d(x,y).$$

Bemerkung: γ muss nicht überall differenzierbar sein, Stückweise differenzierbarkeit von γ reicht. Das wird ausfürlicher, in der Vorlesung erklärt werden.

Bemerkung: Manchmal wird auch $\oint_{\partial G} v_1(x,y)\,dx + v_2(x,y)\,dy$ statt $\oint_{\partial G} \vec{v}\cdot d\vec{s}$ geschrieben.

Wir betrachten den Fall eines offenen Rechtecks $G = (0,b)\times(0,d)$ und parametrisieren die vier Seiten jeweils durch die Bogenlänge. Dann ist

$$\begin{aligned}
\oint_{\partial G} \vec{v}\cdot d\vec{s} &= \int_0^b v_1(x,0)\,dx + \int_0^d v_2(b,y)\,dy - \int_0^b v_1(x,d)\,dx - \int_0^d v_2(0,y)\,dy \\
&= \int_0^d \underbrace{v_2(b,y)-v_2(0,y)}_{=\int_0^b \partial_1 v_2(x,y)\,dx}\,dy - \int_0^b \underbrace{v_1(x,d)-v_1(x,0)}_{=\int_0^d \partial_2 v_1(x,y)\,dy}\,dx \\
&= \iint\limits_G \Big(\partial_1 v_2(x,y) - \partial_2 v_1(x,y)\Big)\,d(x,y).
\end{aligned}$$

Bemerkung: Der Satz gilt auch, in Fällen wo ∂G von mehreren Kurven besteht (in diesem Fall ist G nicht einfach zusammenhängend). Das wird ausfürlicher in der Vorlesung erklärt werden.

20.7. Ausfluss in $\mathbb{R}^2$: Seien G, ∂G, D wie in 20.6 und sei $\vec{w} = \binom{w_1}{w_2} \in C^1(D,\mathbb{R}^2)$. Der Rand ∂G sei parametrisiert durch eine reguläre Kurve $\gamma(t) = \binom{x(t)}{y(t)}$, $a \le t \le b$. Dann ist $\gamma'(t) = \binom{x'(t)}{y'(t)}$. Sei $\vec{N}(t) := \frac{1}{\|\gamma'(t)\|}\binom{y'(t)}{-x'(t)}$. Da G "links von γ" liegt, ist $\vec{N}(t)$ senkrecht auf ∂G und nach außen gerichtet (äußere Einheitsnormale oder äußerer Normaleneinheitsvektor). Das Kurven Integral

$$\oint_{\partial G} \vec{w}\cdot\vec{N}\,ds$$

heißt der Ausfluss des Vektorfeldes durch ∂G. In diesem Integral wird der Anteil des Vektorfeldes $\vec{w}$ in äußerer Normalenrichtung aufintegriert.

20.8. Divergenzsatz in $\mathbb{R}^2$: Seien G, ∂G, D, N, wie in 20.7. Dann

$$\oint_{\partial G} \vec{w}\cdot\vec{N}\,ds = \iint\limits_G \operatorname{div}\vec{w}\; d(x,y).$$

Bemerkung: γ muss nicht überall stetig differenzierbar sein stückweise stetig differenzierbar reicht.

Beweis. Setzt man $\vec{v} := \binom{-w_2}{w_1}$, so erhält man aus 20.6: Wegen $v_1 = -w_2$ und $v_2 = w_1$ ist nämlich

$$\vec{v}\cdot d\vec{s} = \vec{w}\cdot\vec{N}\,ds$$

und $\partial_1 v_2 - \partial_2 v_1 = \partial_1 w_1 + \partial_2 w_2$. □

Wichtige Bemerkung: Aus diesem Satz folgt: Die Divergenz $\nabla \cdot \vec{w}$ ist ein Maß für die Quelldichte des Vektorfeldes.

Insbesondere gilt: Ist $\vec{w}$ quellenfrei in D (dh $\nabla \cdot \vec{w} = 0$ in D), so ist $\oint_{\partial G} \vec{w} \cdot \vec{N}\, ds = 0$ für alle G wie in 20.6 mit $\overline{G} \subseteq D$ (und umgekehrt).

Beispiel: Sei $r > 0$ und $G := \{(x,y) : x^2 + y^2 \leq r^2\}$. Wir parametrisieren $\partial G = \{(x,y) : x^2 + y^2 = r^2\}$ durch $\gamma(t) := \binom{r\cos t}{r \sin t}$, $t \in [0, 2\pi]$. Anhand einer Skizze ist klar, dass $\vec{N}(\gamma(t)) = \binom{\cos t}{\sin t}$ bzw. $\vec{N}(x,y) = \binom{x/\sqrt{x^2+y^2}}{y/\sqrt{x^2+y^2}}$ für $(x,y) \in \partial G$.

$\Big[$ Parametrisieren wir nach der Bogenlänge, so ist $\vec{\gamma}(s) = \binom{r\cos(s/r)}{r\sin(s/r)}$, $s \in [0, 2\pi r]$, und $\vec{\gamma}\,'(s) = \vec{T}(s) = \binom{-\sin(s/r)}{\cos(s/r)}$, sowie $\vec{N}(s) = \binom{\cos(s/r)}{\sin(s/r)}$ $\Big]$.

Das Vektorfeld sei gegeben durch $\vec{w}(x,y) = \binom{xy}{2y}$. Dann gilt $\vec{w} \in C^1(\mathbb{R}^2, \mathbb{R})$ und

$$\begin{aligned}
\oint_{\partial G} \vec{w} \cdot \vec{N}\, ds &= \iint_G \nabla \cdot \vec{w}\, d(x,y) = \iint_G (y+2)\, d(x,y) \\
&= \int_{-r}^{r} \int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} (y+2)\, dy\, dx = 4 \int_{-r}^{r} \sqrt{r^2 - x^2}\, dx.
\end{aligned}$$

Wir substituieren $x = r\xi$, $dx = r\, d\xi$ und erhalten

$$\oint_{\partial G} \vec{w} \cdot \vec{N}\, ds = 4r^2 \int_{-1}^{1} \sqrt{1 - \xi^2}\, d\xi = 2\pi r^2.$$

21 Oberflächenintegrale und Integralsätze im $\mathbb{R}^3$

Wir beschäftigen uns zunächst mit Volumenintegralen im $\mathbb{R}^3$

21.1. Integration über projizierbare Teilmengen von $\mathbb{R}^3$: Sei $B \subseteq \mathbb{R}^3$ von der folgenden Form:

$$B = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in B_0, g(x, y) \leq z \leq h(x, y)\},$$

wobei $g, h : B_0 \to \mathbb{R}$ stetig mit $g \leq h$ und

$$B_0 = \{(x, y) \in \mathbb{R}^2 : x \in [a, b], u(x) \leq y \leq v(x)\}$$

mit $u, v : [a, b] \to \mathbb{R}$ stetig.

Ist dann $f : B \to \mathbb{R}$ stetig, so ist f über B integrierbar und es gilt

$$\iiint\limits_B f(x, y, z)\, d(x, y, z) = \int_a^b \Big(\int_{u(x)}^{v(x)} \Big(\int_{g(x,y)}^{h(x,y)} f(x, y, z)\, dz \Big)\, dy \Big)\, dx.$$

Bemerkung: (a) Die Rollen von x, y, z können vertauscht werden (vergleiche Satz 2 in 20.5).

(b) Für $f = 1$ erhält man das <u>Volumen</u> $\mathrm{vol}(B)$ <u>von B</u>.

(c) Ähnlicher Satz gilt wenn man integriert über Bereiche von $\mathbb{R}^n$.

Beispiele: (1) Sei $r > 0$ und

$$B = \{(x, y, z) : x^2 + y^2 + z^2 \leq r^2\}$$

(Kugel um $\vec{0}$ mit Radius r). Mit $a = -r$, $b = r$, $u(x) = -\sqrt{r^2 - x^2}$, $v(x) = \sqrt{r^2 - x^2}$, $g(x, y) = -\sqrt{r^2 - (x^2 + y^2)}$, $h(x, y) = \sqrt{r^2 - (x^2 + y^2)}$ erhalten wir

$$\iiint\limits_B d(x, y, z) = \int_{-r}^{r} \int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} \int_{-\sqrt{r^2-(x^2+y^2)}}^{\sqrt{r^2-(x^2+y^2)}} dz\, dy\, dx = \int_{-r}^{r} \int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} 2\sqrt{r^2 - (x^2 + y^2)}\, dy\, dx.$$

Wir substituieren im inneren Integral $y = \sqrt{r^2 - x^2}\eta$, $dy = \sqrt{r^2 - x^2}\, d\eta$ und erhalten

$$\int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} 2\sqrt{r^2 - (x^2 + y^2)}\, dy = 2(r^2 - x^2) \int_{-1}^{1} \sqrt{1 - \eta^2}\, d\eta = \pi(r^2 - x^2).$$

Somit ist

$$\mathrm{vol}(B) = \pi \int_{-r}^{r} (r^2 - x^2)\, dx = \pi \Big[r^2 x - \frac{x^3}{3} \Big]_{-r}^{r} = \frac{4\pi}{3} r^3.$$

(2) Sei

$$B := \{(x,y,z) : 0 \le x^2 + y^2 \le z \le 1\}.$$

Wir berechnen

$$\iiint\limits_B 1\, d(x,y,z) = \int_0^1 \Big(\int_{-\sqrt{z}}^{\sqrt{z}} \Big(\int_{-\sqrt{z-x^2}}^{\sqrt{z-x^2}} 1\, dy \Big)\, dx \Big)\, dz = \int_0^1 \Big(\iint\limits_{\{(x,y): x^2+y^2 \le z\}} 1\, d(x,y) \Big)\, dz.$$

Das innere Integral ist der Flächeninhalt eines Kreises mit Radius $\sqrt{z}$. Dieser Kreis ist der Schnitt durch B mit festgehaltener z-Komponente. Wir erhalten somit

$$\iiint\limits_B 1\, d(x,y,z) = \int_0^1 \pi z\, dz = \frac{\pi}{2}.$$

Bemerkung (Prinzip von Cavalieri): Für $B \subseteq \mathbb{R}^3$ der Form

$$B = \{(x,y,z) : z \in [a,b], (x,y) \in B(z)\}$$

gilt allgemein

$$\iiint\limits_B d(x,y,z) = \int_a^b \int_{B(z)} d(x,y)\, dz.$$

Hierbei ist $B(z)$ der Schnitt durch B mit festgehaltener z-Komponente.

Bemerkung: Wir werden im folgenden über beschränkte Mengen im $\mathbb{R}^3$ integrieren, die sich in endlich viele Gebiete der Form B (mit eventuell vertauschten Rollen der Koordinaten) zerlegen lassen. Wir nennen solche Mengen <u>Integrationsbereiche</u>.

21.2. Transformationsformel: Sei $B \subseteq \mathbb{R}^n$ beschränkt und abgeschlossen (meist ist $n = 2$ oder $n = 3$ und B ein Integrationsbereich) und $U \supseteq B$ ein Gebiet. Sei $\Phi : U \to \mathbb{R}^n$ stetig differenzierbar und injektiv mit $\det \Phi' \neq 0$ auf U, sowie $A := \Phi(B)$ und $f : A \to \mathbb{R}$ beschränkt.

Bemerkung: Nach 19.14 ist dann auch $V := \Phi(U)$ offen und $\Phi : U \to V$ bijektiv und in beiden Richtungen stetig differenzierbar. Da U ein Gebiet und $\det \Phi' : U \to \mathbb{R}$ stetig ist, gilt außerdem $\det \Phi' > 0$ auf U oder $\det \Phi' < 0$ auf U.

Satz: Es ist f integrierbar über A genau dann, wenn $f \circ \Phi |\det \Phi'(\cdot)|$ über B integrierbar ist. In diesem Falle gilt

$$\int_A f(x)\, dx = \int_B f(\Phi(y)) |\det (\Phi'(y))|\, dy.$$

Im einfachsten Fall ist $\Phi(x) = Cx + b$, wobei $C \in \mathbb{R}^{n\times n}$ regulär ist und $b \in \mathbb{R}^n$. Nach der "Interpretation" in 17.9 muss das Volumen eines Quaders $Q \subseteq B$ beim Abbilden mit Φ gerade mit $|\det(C)|$ multipliziert werden. Das ist die Kernidee hinter der Transformationsformel.

21.3. Polarkoordinaten im $\mathbb{R}^2$: Für $(x, y) \in \mathbb{R}^2$ setze $r := \|(x, y)\| = \sqrt{x^2 + y^2}$. Dann findet man Winkel φ mit $x = r\cos\varphi$, $y = r\sin\varphi$.

Für $\Phi(r, \varphi) := \binom{r\cos\varphi}{r\sin\varphi}$ gilt

$$\Phi'(r, \varphi) = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix},$$

also $\det\Phi'(r, \varphi) = r$.

Damit Φ injektiv ist, nehme man etwa $U = (0, \infty) \times (\tilde{\varphi}_1, \tilde{\varphi}_2)$ mit $0 \le \tilde{\varphi}_1 < \tilde{\varphi}_2 \le 2\pi$ und $\tilde{\varphi}_2 - \tilde{\varphi}_1 < 2\pi$. Sind $\tilde{\varphi}_1 < \varphi_1 < \varphi_2 < \tilde{\varphi}_2$, $B := [R_1, R_2] \times [\varphi_1, \varphi_2]$ und $A := \Phi(B)$, so gilt für stetiges $f : A \to \mathbb{R}$:

$$\iint\limits_A f(x, y)\, d(x, y) = \iint\limits_B f(r\cos\varphi, r\sin\varphi) r\, d(r, \varphi) = \int_{\varphi_1}^{\varphi_2} \int_{R_1}^{R_2} f(r\cos\varphi, r\sin\varphi) r\, dr\, d\varphi.$$

Zusatz: Diese Formel gilt auch für $\varphi_1 = 0$ und $\varphi_2 = 2\pi$, sowie für $R_1 = 0$.

Beispiele: (1) $A = \{(x, y) \in \mathbb{R}^2 : y \ge 0, 1 \le x^2 + y^2 \le 4\}$. Hier ist $R_1 = 1$, $R_2 = 2$, $\varphi_1 = 0$, $\varphi_2 = \pi$, also $B = [1, 2] \times [0, \pi]$. Es gilt also für $f(x, y) = y\sqrt{x^2 + y^2}$:

$$\begin{aligned} \iint\limits_A y\sqrt{x^2 + y^2}\, d(x, y) &= \iint\limits_B r\sin\varphi\, rr\, d(r, \varphi) = \int_0^\pi \int_1^2 r^3 \sin\varphi\, dr\, d\varphi \\ &= \int_0^\pi \sin\varphi\, d\varphi \cdot \int_1^2 r^3\, dr = 2 \cdot \left[\frac{r^4}{4}\right]_1^2 = \frac{15}{2}. \end{aligned}$$

(2) Sei $M := \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 \le 4 - z, z \in [0, 4]\}$ und $A := \{(x, y) : x^2 + y^2 \le 4\}$. Dann gilt

$$\begin{aligned} \operatorname{vol}(M) &= \iiint\limits_M 1\, d(x, y, z) = \iint\limits_A \Big(\int_0^{4-(x^2+y^2)} 1\, dz\Big)\, d(x, y) \\ &= \iint\limits_A 4 - (x^2 + y^2)\, d(x, y) = \int_0^{2\pi} \int_0^2 (4 - r^2) r\, dr\, d\varphi = 8\pi. \end{aligned}$$

(3) Wir berechnen das Integral $\int_0^\infty e^{-x^2}\, dx$ (das war in HM I nicht möglich). Dazu sei $R > 0$ und

$$K_R := \{(x, y) : x, y \ge 0, x^2 + y^2 \le R^2\}, \quad Q_R := [0, R] \times [0, R]$$

und $\rho := \sqrt{2}R$, sowie $f(x,y) := e^{-(x^2+y^2)}$ auf $\mathbb{R}^2$.
Dann gilt

$$\iint\limits_{K_R} f(x,y)\,d(x,y) \le \iint\limits_{Q_R} f(x,y)\,d(x,y) \le \iint\limits_{K_\rho} f(x,y)\,d(x,y),$$

$$\iint\limits_{Q_R} f(x,y)\,d(x,y) = \iint\limits_{Q_R} (e^{-x^2}e^{-y^2})\,d(x,y) = \Big(\int_0^R e^{-x^2}\,dx\Big)^2,$$

$$\iint\limits_{K_R} f(x,y)\,d(x,y) = \int_0^{\pi/2}\int_0^R e^{-r^2} r\,dr\,d\varphi = \frac{\pi}{4}(1-e^{-R^2}),$$

und genauso

$$\iint\limits_{K_\rho} f(x,y)\,d(x,y) = \frac{\pi}{4}(1-e^{-\rho^2}) = \frac{\pi}{4}(1-e^{-2R^2}).$$

Wir lassen nun $R \to \infty$ und erhalten

$$\int_0^\infty e^{-x^2}\,dx = \frac{\sqrt{\pi}}{2}, \qquad \int_{-\infty}^\infty e^{-x^2}\,dx = \sqrt{\pi}.$$

21.4. Zylinderkoordinaten im $\mathbb{R}^3$: Hier ist $\Phi(r,\varphi,z) = \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \\ z \end{pmatrix}$, also $x = r\cos\varphi$, $y = r\sin\varphi$ und $z = z$. Es gilt

$$\Phi'(r,\varphi,z) = \begin{pmatrix} \cos\varphi & -r\sin\varphi & 0 \\ \sin\varphi & r\cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

also $\det\Phi'(r,\varphi,z) = r$. Für $A, B \subseteq \mathbb{R}^3$ wie in 21.3 und stetiges $f: A \to \mathbb{R}$ gilt somit:

$$\iiint\limits_A f(x,y,z)\,d(x,y,z) = \iiint\limits_B f(r\cos\varphi, r\sin\varphi, z) r\,d(r,\varphi,z).$$

Der Zusatz aus 21.3 gilt sinngemäß auch hier.

Beispiel: $A = \{(x,y,z) : x^2+y^2 \le 1, 0 \le y \le x, z \in [0,1]\}$, also $B = [0,1]\times[0,\pi/4]\times[0,1]$.

Dann ist für $f(x,y,z) = x^2 + y^2 + yz$:

$$\begin{aligned}
\iiint\limits_A (x^2+y^2+yz)\,d(x,y,z) &= \iiint\limits_B (r^2 + zr\sin\varphi)r\,d(r,\varphi,z) \\
&= \int_0^1 \int_0^{\pi/4} \int_0^1 (r^2 + zr\sin\varphi)r\,dz\,d\varphi\,dr \\
&= \int_0^1 \int_0^{\pi/4} \Big[zr^3 + \frac{z^2}{2}r^2\sin\varphi\Big]_{z=0}^{z=1} d\varphi\,dr \\
&= \int_0^1 \int_0^{\pi/4} r^3 + \frac{r^2}{2}\sin\varphi\,d\varphi\,dr \\
&= \int_0^1 \frac{\pi}{4}r^3 + \frac{r^2}{2}[-\cos\varphi]_0^{\pi/4}\,dr = \frac{\pi}{16} + \frac{1}{6}\Big(1 - \frac{\sqrt{2}}{2}\Big).
\end{aligned}$$

21.5. Kugelkoordinaten im $\mathbb{R}^3$: Man schreibt $r = \|(x,y,z)\| = (x^2+y^2+z^2)^{1/2}$ und

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \Phi(r,\varphi,\vartheta) = \begin{pmatrix} r\cos\varphi\sin\vartheta \\ r\sin\varphi\sin\vartheta \\ r\cos\vartheta \end{pmatrix},$$

wobei $r \geq 0$, $\varphi \in [0, 2\pi]$ und $\vartheta \in [0,\pi]$. Es ist

$$\Phi'(r,\varphi,\vartheta) = \begin{pmatrix} \cos\varphi\sin\vartheta & -r\sin\varphi\sin\vartheta & r\cos\varphi\cos\vartheta \\ \sin\varphi\sin\vartheta & r\cos\varphi\sin\vartheta & r\sin\varphi\cos\vartheta \\ \cos\vartheta & 0 & -r\sin\vartheta \end{pmatrix},$$

also $\det\Phi'(r,\varphi,\vartheta) = -r^2\sin\vartheta$.

Sind A und B wie in 21.2, also $A = \Phi(B)$, und ist $f : A \to \mathbb{R}$ stetig, so gilt:

$$\iiint\limits_A f(x,y,z)\,d(x,y,z) = \iiint\limits_B f(r\cos\varphi\sin\vartheta, r\sin\varphi\sin\vartheta, r\cos\vartheta)r^2\sin\vartheta\,d(r,\varphi,\vartheta).$$

Der Zusatz in 21.3 gilt entsprechend.

Beispiel: Sei $A = \{(x,y,z) : x^2+y^2+z^2 \leq 1, x,y,z \geq 0\}$. Dann ist $B = [0,1] \times [0,\pi/2] \times [0,\pi/2]$, und für $f(x,y,z) = x\sqrt{x^2+y^2+z^2}$ gilt:

$$\begin{aligned}
\iiint\limits_A x\sqrt{x^2+y^2+z^2}\,d(x,y,z) &= \iiint\limits_B r\cos\varphi\sin\vartheta\cdot r\cdot r^2\sin\vartheta\,d(r,\varphi,\vartheta) \\
&= \int_0^{\pi/2}\int_0^{\pi/2}\int_0^1 r^4\cos\varphi\sin^2\vartheta\,dr\,d\varphi\,d\vartheta \\
&= \int_0^1 r^4\,dr \cdot \underbrace{\int_0^{\pi/2}\cos\varphi\,d\varphi}_{=1} \cdot \underbrace{\int_0^{\pi/2}\sin^2\vartheta\,d\vartheta}_{=\pi/4} = \frac{\pi}{20}.
\end{aligned}$$

21.6. Flächendarstellungen im $\mathbb{R}^3$: Es gibt verschiedene Möglichkeiten, Flächen im $\mathbb{R}^3$ darzustellen.

Explizite Darstellung: $z = f(x,y)$, z.B. $z = \pm\sqrt{1-x^2-y^2}$, genauer

$$\mathcal{F} = \left\{ \begin{pmatrix} x \\ y \\ f(x,y) \end{pmatrix} : (x,y) \in U \right\},$$

wobei $U \subseteq \mathbb{R}^2$ Gebiet und $f \in C^1(U,\mathbb{R})$.

Implizite Darstellung: $F(x,y,z) = 0$, z.B. $x^2+y^2+z^2-1=0$, genauer

$$\mathcal{F} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in D : F(x,y,z) = 0 \right\},$$

wobei $D \subseteq \mathbb{R}^3$ Gebiet und $F \in C^1(D,\mathbb{R})$. Hierbei sei $\nabla F(x,y,z) \neq \vec{0}$ für $(x,y,z) \in \mathcal{F}$. Diese Bedingung sorgt dafür, dass man lokal immer nach einer der Variablen x, y oder z auflösen kann (vgl. 19.15).

Parameterdarstellung: Beispiel

$$\mathcal{F} = \left\{ \begin{pmatrix} \cos\varphi \sin\vartheta \\ \sin\varphi \sin\vartheta \\ \cos\vartheta \end{pmatrix} : \varphi \in (0,2\pi), \vartheta \in (0,\pi) \right\},$$

allgemein

$$\mathcal{F} = \{\vec{g}(u,v) : (u,v) \in U\},$$

wobei $U \subseteq \mathbb{R}^2$ Gebiet und $\vec{g} \in C^1(U,\mathbb{R}^3)$ **injektiv** ist mit $\operatorname{Rang} \vec{g}\,'(u,v) = 2$ für alle $(u,v) \in U$ (beachte $\vec{g}\,'(u,v) \in \mathbb{R}^{3\times 2}$). Ein solches $\mathcal{F}$ heißt <u>reguläres Flächenstück im $\mathbb{R}^3$</u> und $\vec{g}$ heißt <u>reguläre Parametrisierung</u> von $\mathcal{F}$.

Bemerkung: Durch $F(x,y,z) = z - f(x,y)$ kommt man von einer expliziten zu einer impliziten Darstellung. Die explizite Darstellung ist ein Spezialfall der Parameterdarstellung via

$$\vec{g}(x,y) := \begin{pmatrix} x \\ y \\ f(x,y) \end{pmatrix}, \quad (x,y) \in U.$$

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 <u>Normaleneinheitsvektor</u>. Die Entscheidung für einen der beiden legt die <u>Orientierung des Flächenstücks</u> 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 $\mathcal{F} \to \mathbb{R}^3$.

In Parameterdarstellung ist

$$\vec{N}(\vec{g}(u,v)) = \pm\frac{\partial_u\vec{g}(u,v)\times\partial_v\vec{g}(u,v)}{\|\partial_u\vec{g}(u,v)\times\partial_v\vec{g}(u,v)\|},\quad (u,v)\in U.$$

Wir verwenden, wenn nichts anderes gesagt wird, hier das **positive** Vorzeichen.

In impliziter Darstellung $F(x,y,z)=0$ ist der Normaleneinheitsvektor

$$\vec{N}(x,y,z) = \pm\frac{\nabla F(x,y,z)}{\|\nabla F(x,y,z)\|},\quad (x,y,z)\in\mathcal{F}$$

(beachte, dass man statt F ebenso $-F$ verwenden kann). Das liegt daran, dass der Gradient von F senkrecht auf der Niveaufläche steht (vgl. 19.12).

Für die explizite Darstellung $z=f(x,y)$ erhalten wir

$$\vec{N}(x,y,f(x,y)) = \pm\frac{1}{\sqrt{1+(\partial_x f(x,y))^2+(\partial_y f(x,y))^2}}\begin{pmatrix}-\partial_x f(x,y)\\ -\partial_y f(x,y)\\ 1\end{pmatrix},\quad (x,y)\in U$$

(beachte, dass $\partial_x\vec{g} = \begin{pmatrix}1\\0\\\partial_x f\end{pmatrix}$, $\partial_y\vec{g} = \begin{pmatrix}0\\1\\\partial_y f\end{pmatrix}$ und $\partial_x\vec{g}\times\partial_y\vec{g} = \begin{pmatrix}-\partial_x f\\-\partial_y f\\1\end{pmatrix}$ ist).

21.7. Oberflächenintegral: Sei $U\subseteq\mathbb{R}^2$ ein Gebiet und $\vec{g}:U\to\mathbb{R}^3$ eine reguläre (insbesondere also injektive) Parametrisierung eines Flächenstücks. Dann heißt

$$do := \|\partial_u\vec{g}(u,v)\times\partial_v\vec{g}(u,v)\|\,d(u,v)$$

skalares Oberflächenelement, und

$$d\vec{o} := \partial_u\vec{g}(u,v)\times\partial_v\vec{g}(u,v)\,d(u,v)$$

heißt vektorielles Oberflächenelement des durch $\vec{g}$ parametrisierten regulären Flächenstücks.

Definition: Sei $B\subseteq U$ ein Integrationsbereich und $\mathcal{F}:=\vec{g}(B)$. Für ein stetiges Skalarfeld $f:\mathcal{F}\to\mathbb{R}$ definiert man das (Oberflächen-)Integral von f über $\mathcal{F}$ durch

$$\iint\limits_{\mathcal{F}} f\,do := \iint\limits_{B} f(\vec{g}(u,v))\,\|\partial_u\vec{g}(u,v)\times\partial_v\vec{g}(u,v)\|\,d(u,v),$$

und für ein stetiges Vektorfeld $\vec{w}:\mathcal{F}\to\mathbb{R}^3$ setzt man

$$\iint\limits_{\mathcal{F}} \vec{w}\cdot d\vec{o} := \iint\limits_{B} \vec{w}(\vec{g}(u,v))\cdot\Big(\partial_u\vec{g}(u,v)\times\partial_v\vec{g}(u,v)\Big)\,d(u,v).$$

Für $f = 1$ erhält man den Flächeninhalt von $\mathcal{F}$:

$$A(\mathcal{F}) := \iint\limits_{\mathcal{F}} do = \iint\limits_{B} \|\partial_u \vec{g}(u,v) \times \partial_v \vec{g}(u,v)\| \, d(u,v).$$

Bemerkung: Die Definitionen sind invariant unter orientierungserhaltenden Parametertransformationen.

Bemerkung: Wenn wir das vektorielle Oberflächenelement mit dem skalaren Oberflächenelement vergleichen und – wie in 21.6 gesagt – für $\vec{N}$ das positive Vorzeichen nehmen, dh

$$\vec{N}(\vec{g}(u,v)) = \frac{\partial_u \vec{g}(u,v) \times \partial_v \vec{g}(u,v)}{\|\partial_u \vec{g}(u,v) \times \partial_v \vec{g}(u,v)\|},$$

so ist

$$\iint\limits_{\mathcal{F}} \vec{w} \cdot d\vec{o} = \iint\limits_{\mathcal{F}} \vec{w} \cdot \vec{N} \, do,$$

das ist der **Fluss** des Vektorfelds $\vec{v}$ durch die mittels $\vec{N}$ orientierte Fläche $\mathcal{F}$.

Beispiel: Wir berechnen den Flächeninhalt einer oberen Halbkugel

$$\mathcal{F} = \{ \begin{pmatrix} x \\ y \\ \sqrt{R^2 - x^2 - y^2} \end{pmatrix} : x^2 + y^2 \leq R^2 \}$$

mit Radius $R > 0$. Wir haben eine explizite Darstellung mit $f(x,y) = \sqrt{R^2 - x^2 - y^2}$ und $\vec{g}(x,y) = \begin{pmatrix} x \\ y \\ f(x,y) \end{pmatrix}$. Es ist $\partial_x \vec{g} \times \partial_y \vec{g} = \begin{pmatrix} -\partial_x f \\ -\partial_y f \\ 1 \end{pmatrix}$, wobei $-\partial_x f = x/\sqrt{R^2 - x^2 - y^2}$ und $-\partial_y f = y/\sqrt{R^2 - x^2 - y^2}$. Wir setzen $B := \{(x,y) : x^2 + y^2 \leq R^2\}$. Somit ist

$$\begin{aligned} A(\mathcal{F}) &= \iint\limits_{B} \sqrt{\frac{x^2}{R^2 - x^2 - y^2} + \frac{y^2}{R^2 - x^2 - y^2} + 1} \, d(x,y) \\ &= \int_0^{2\pi} \int_0^R \sqrt{\frac{R^2}{R^2 - r^2}} \, r \, dr \, d\varphi \\ &= 2\pi R^2 \int_0^1 \rho \frac{1}{\sqrt{1 - \rho^2}} \, d\rho \\ &= 2\pi R^2 \Big[-\sqrt{1 - \rho^2} \Big]_0^1 = 2\pi R^2. \end{aligned}$$

21.8. Der Integralsatz von Stokes im $\mathbb{R}^3$: Sei $U \subseteq \mathbb{R}^2$ ein Gebiet und $\vec{g} : U \to \mathbb{R}^3$ eine reguläre Parametrisierung eines Flächenstücks $\mathcal{F}^* := \vec{g}(U)$.

Sei $G \subseteq U$ ein Gebiet so, dass $\overline{G}$ ein Integrationsbereich ist. Der Rand ∂G von G bestehe aus endlich vielen regulären Kurven $\gamma_1, \ldots, \gamma_m$, dh $\gamma := \gamma_1 + \ldots + \gamma_m$ (im Sinne von Bemerkung 20.1(d)) ist doppelpunktfrei und hat ∂G als Spur. Die Orientierung von γ sei so, dass G "links von γ liegt" (dh ∂G ist positiv orientiert).

Sei $\mathcal{F} := \vec{g}(G)$. Dann ist $\partial\mathcal{F} := \vec{g}(\partial G)$ parametrisiert durch $\vec{g} \circ \gamma$.

Satz: Ist nun $V \subseteq \mathbb{R}^3$ offen mit $\mathcal{F} \subseteq V$ und $\vec{v} : V \to \mathbb{R}^3$ ein C^1-Vektorfeld, so gilt

$$\oint_{\partial\mathcal{F}} \vec{v} \cdot d\vec{s} = \iint_{\mathcal{F}} (\nabla \times \vec{v}) \cdot d\vec{o}.$$

Nach 20.2 können wir die linke Seite auch schreiben als

$$\oint_{\partial\mathcal{F}} \vec{v} \cdot \vec{T}\, ds,$$

wobei $\vec{T} : \mathcal{F} \to \mathbb{R}^3$ der Tangenteneinheitsvektor ist, hier gegeben durch

$$\vec{T}(\vec{g}(\gamma(t))) = \frac{\vec{g}\,'(\gamma(t))\dot{\gamma}(t)}{\|\vec{g}\,'(\gamma(t))\dot{\gamma}(t)\|},$$

und die rechte Seite können wir schreiben als

$$\iint_{\mathcal{F}} (\nabla \times \vec{v}) \cdot \vec{N}\, do.$$

Alternativ: Ist $\vec{T}$ auf $\partial\mathcal{F}$ gegeben (und damit die Orientierung von $\partial\mathcal{F}$ festgelegt), so sei im Punkt $P \in \partial\mathcal{F}$ der Vektor $\vec{n}$ der Vektor der Länge 1, der in der Tangentialebene an $\mathcal{F}$ in P senkrecht auf $\vec{T}$ steht und ins Äußere von $\mathcal{F}$ weist. Die Richtung von $\vec{N}$ ist dann diejenige von $\vec{n} \times \vec{T}$.

Anders ausgedrückt: Die Orientierung von $\vec{N}$ auf $\mathcal{F}$ ergibt sich aus der Orientierung von $\partial\mathcal{F}$ im Sinne der "Rechtsschraubenregel". Man vergleiche hierzu auch die ebene Version des Stokesschen Integralsatzes in 20.7!

Beispiele: (1) Sei $\mathcal{F} := \{(x, y, z) : x^2 + y^2 + z^2 = 1, z \geq 0\}$. Dann ist $\partial\mathcal{F} = \{(x, y, 0) : x^2 + y^2 = 1\}$. Orientieren wir $\partial\mathcal{F}$ durch $\vec{T}(x, y, 0) = \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}$, so erhalten wir $\vec{N}(x, y, z) = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ für $(x, y, z) \in \mathcal{F}$.

Wir wollen

$$J := \iint\limits_{\mathcal{F}} (\nabla \times \vec{v}) \cdot \, d\vec{o}$$

berechnen, wobei $\vec{v} \in C^1(\mathbb{R}^3, \mathbb{R}^3)$ sei.

Nach dem Stokesschen Satz ist

$$J = \oint_{\partial \mathcal{F}} \vec{v} \cdot \, d\vec{s} = \int_0^{2\pi} \vec{v}(\cos t, \sin t, 0) \cdot \begin{pmatrix} -\sin t \\ \cos t \\ 0 \end{pmatrix} dt.$$

Für $\vec{v}(x,y,z) = \begin{pmatrix} -y+z \\ x+z \\ z-x \end{pmatrix}$ erhalten wir

$$J = \int_0^{2\pi} \begin{pmatrix} -\sin t \\ \cos t \\ -\cos t \end{pmatrix} \cdot \begin{pmatrix} -\sin t \\ \cos t \\ 0 \end{pmatrix} dt = 2\pi.$$

Hier ist übrigens $\nabla \times \vec{v}(x,y,z) = \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix}$.

(2) Sei G ein Integrationsbereich im $\mathbb{R}^2$ mit einer aus endlich vielen regulären Kurven zusammengesetzten doppelpunktfreien Randkurve. Dann gilt für die Fläche $A(G)$ von G:

$$A(G) = \frac{1}{2} \oint_{\partial G} \begin{pmatrix} -y \\ x \end{pmatrix} \cdot \, d\vec{s}.$$

(Man bette G in den $\mathbb{R}^3$ ein und beachte $\vec{N} = \vec{e}_3$ und $\left(\nabla \times \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}\right) \cdot \vec{e}_3 = 2$.)

(3) Anwendung von (2): Seien $\alpha, \beta \in \mathbb{R}$ mit $0 < \beta - \alpha \leq 2\pi$ und $r : [\alpha, \beta] \to \mathbb{R}$ eine C^1-Funktion. Sei $G \subseteq \mathbb{R}^2$ gegeben durch

$$G = \{(r \cos t, r \sin t) : t \in (\alpha, \beta), r \in (0, r(t))\}.$$

Dann gilt die Leibnizsche Sektorformel:

$$A(G) = \frac{1}{2} \int_\alpha^\beta r(t)^2 \, dt.$$

(Die Integrale über die Strecken $[(0,0), (r(\alpha)\cos\alpha, r(\alpha)\sin\alpha)]$ und $[(0,0), (r(\beta)\cos\beta, r(\beta)\sin\beta)]$ verschwinden, das Integral über $\gamma(t) = \binom{r(t)\cos t}{r(t)\sin t}$, $t \in [\alpha, \beta]$ rechne man aus.)

(4) Beispiel (2) legt nahe, $A(\mathcal{F})$ auch für gekrümmte Flächen durch

$$A(\mathcal{F}) = \frac{1}{2} \oint_{\partial \mathcal{F}} \vec{v} \cdot d\vec{s}.$$

berechnen zu können, wobei $\vec{v}$ ein C^1-Vektorfeld mit $(\nabla \times \vec{v}) \cdot \vec{N} = 1$ auf $\mathcal{F}$ ist. Dann sollte aber $\nabla \times \vec{v} = \vec{N}$ auf $\mathcal{F}$ sein und somit insbesondere $\nabla \cdot \vec{N} = 0$.

Ist jedoch z.B. $\mathcal{F}$ die obere Halbkugel mit Radius 1, dann ist $\nabla \cdot \vec{N} = 3$. Also findet man hier kein geeignetes Vektorfeld $\vec{v}$.

21.9. Der Divergenzsatz im $\mathbb{R}^3$: Sei $B \subseteq \mathbb{R}^3$ ein beschränkter und abgeschlossener Integrationsbereich 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 reguläre 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 $\vec{v} : V \to \mathbb{R}^3$ ein C^1-Vektorfeld. Dann gilt

$$\iiint\limits_G \nabla \cdot \vec{v}\, d\tau = \iint\limits_{\partial G} \vec{v} \cdot \vec{N}\, do,$$

wobei wir hier $d\tau$ für $d(x, y, z)$ geschrieben haben.

Beispiele: (1) Für $\vec{v} \in C^2(V, \mathbb{R}^3)$ gilt

$$\iint\limits_{\partial G} (\nabla \times \vec{v}) \cdot d\vec{o} = 0,$$

da ja $\operatorname{div} \operatorname{rot} \vec{v} = 0$ in G. Nach 21.8 ist das nicht verwunderlich, da die Oberfläche ∂G von G ja <u>geschlossen</u> ist und selbst keinen (Flächen-)Rand hat.

(2) Für $f \in C^2(V)$ gilt

$$\iiint\limits_G \Delta f\, d\tau = \iint\limits_{\partial G} \nabla f \cdot d\vec{o} = \iint\limits_{\partial G} \nabla f \cdot \vec{N}\, do = \iint\limits_{\partial G} \frac{\partial f}{\partial \vec{N}}\, do.$$

(3) **Greensche Formeln** Für $f, g \in C^2(V, \mathbb{R})$ und $h \in C^1(V, \mathbb{R})$ gilt:

$$\iint\limits_{\partial G} h \frac{\partial f}{\partial \vec{N}}\, do = \iiint\limits_G \Big(h \Delta f + \nabla h \cdot \nabla f \Big)\, d\tau,$$

$$\iint\limits_{\partial G} \Big(g \frac{\partial f}{\partial \vec{N}} - f \frac{\partial g}{\partial \vec{N}} \Big)\, do = \iiint\limits_G \Big(g \Delta f - f \Delta g \Big)\, d\tau.$$

(4) Setzt man $\vec{v}(\vec{x}) := \vec{x}$, so gilt $\nabla \cdot \vec{v} = 3$, und wir erhalten

$$\operatorname{vol}(G) = \frac{1}{3} \iint\limits_{\partial G} \vec{x} \cdot \vec{N} \, do.$$

Für $G = \{\vec{x} : \|\vec{x}\| < R\}$ ist also wegen $\vec{N} = \vec{x}/\|\vec{x}\|$:

$$\operatorname{vol}(G) = \frac{1}{3} \iint\limits_{\partial G} \|\vec{x}\| \, do = \frac{R}{3} \underbrace{A(\partial G)}_{=4\pi R^2} = \frac{4\pi}{3} R^3.$$

(5) Wir setzen $f(\vec{x}) = \|\vec{x}\|^{-1}$ für $\vec{x} \in \mathbb{R}^3 \setminus \{\vec{0}\}$ und $\vec{v} = \nabla f$. Dann ist

$$\begin{aligned}
\vec{v}(\vec{x}) &= \Big(-\frac{1}{2}(x_1^2 + x_2^2 + x_3^2)^{-3/2} \cdot (2x_j) \Big)_j = -\frac{\vec{x}}{\|\vec{x}\|^3} \\
\nabla \cdot \vec{v}(\vec{x}) &= -\sum_{j=1}^{3} \Big(\frac{1}{\|\vec{x}\|^3} - 3\frac{x_j^2}{\|\vec{x}\|^5} \Big) = 0
\end{aligned}$$

für $\vec{x} \neq \vec{0}$.

Nach dem Divergenzsatz ist also für G mit $0 \notin \overline{G}$:

$$\iint\limits_{\partial G} \frac{\vec{x}}{\|\vec{x}\|^3} \cdot d\vec{o} = 0.$$

Wir halten außerdem fest, dass

$$\Delta \frac{1}{\|\vec{x}\|} = 0 \quad \text{in } \mathbb{R}^3 \setminus \{\vec{0}\}.$$

(6) Sei $\varphi \in C^2(\mathbb{R}^3, \mathbb{R})$ mit $\varphi = 0$ außerhalb einer Kugel G um $\vec{0}$ mit Radius R. Wir wollen

$$\iiint\limits_{G} \frac{1}{\|\vec{x}\|} \Delta\varphi \, d\tau = -4\pi\varphi(\vec{0})$$

zeigen. Da der Integrand für $\vec{x} \to 0$ nicht beschränkt bleiben muss, verstehen wir unter der linken Seite

$$\lim_{\varepsilon \to 0} \iiint\limits_{G_\varepsilon} \frac{1}{\|\vec{x}\|} \Delta\varphi \, d\tau,$$

wobei $G_\varepsilon := G \setminus \{\|\vec{x}\| \leq \varepsilon\}$. Für festes $\varepsilon \in (0, R)$ ist jetzt nach der zweiten Greenschen Formel und nach (5):

$$\iiint\limits_{G_\varepsilon} \frac{1}{\|\vec{x}\|} \Delta\varphi \, d\tau = \int_{\|\vec{x}\| = \varepsilon} \frac{1}{\|\vec{x}\|} \frac{\partial \varphi}{\partial \vec{N}} \, do - \int_{\|\vec{x}\| = \varepsilon} \varphi \frac{-\vec{x}}{\|\vec{x}\|^3} \cdot \vec{N} \, do.$$

Dabei beachte man, dass für $\|\vec{x}\| = \varepsilon$ gilt: $\vec{N} = -\vec{x}/\varepsilon$. Also ist

$$-\int_{\|\vec{x}\|=\varepsilon} \varphi \frac{-\vec{x}}{\|\vec{x}\|^3} \cdot \vec{N}\, do = -\varepsilon^{-2} \int_{\|\vec{x}\|=\varepsilon} \varphi\, do \to -4\pi\varphi(0) \quad (\varepsilon \to 0).$$

Andererseits ist $\|\nabla\varphi\| \leq K$ für eine geeignete Konstante K und daher

$$\Big| \int_{\|\vec{x}\|=\varepsilon} \frac{1}{\|\vec{x}\|} \frac{\partial\varphi}{\partial\vec{N}}\, do \Big| \leq \frac{4\pi\varepsilon^2}{\varepsilon} K \to 0 \quad (\varepsilon \to 0).$$

Interpretation: Es gilt "$\Delta \frac{1}{\|\vec{x}\|} = -4\pi\delta_{\vec{0}}$", wenn wir gegen C^2-Funktionen ϕ integrieren, die außerhalb einer Kugel $G = K(\vec{0}, R)$ verschwinden. Dabei setzt man für solche ϕ:

$$\iiint\limits_G \Big(\Delta \frac{1}{\|\vec{x}\|}\Big) \phi\, d\tau := \iiint\limits_G \frac{1}{\|\vec{x}\|} (\Delta\phi)\, d\tau,$$

was durch die zweite Greensche Formel in (3) oben gerechtfertigt ist (die Randterme verschwinden), und definiert

$$\iiint\limits_G \phi\delta_{\vec{0}}\, d\tau := \phi(\vec{0}),$$

dh $\delta_{\vec{0}}$ ist formal als eine "Dichte" mit Integral 1 zu verstehen, die im Punkt $\vec{0}$ konzentriert ist.

Bemerkung: Ähnlich kann man im $\mathbb{R}^2$ zeigen, dass für $g(x, y) := -\frac{1}{4\pi} \ln(x^2 + y^2)$ gilt: $\Delta g = \delta_{(0,0)}$.