Mathematik für Physiker II

Skriptum des SS 2002

Prof. Dr. M.v.Golitschek

Institut für Angewandte Mathematik und Statistik

Universität Würzburg

Literatur:

Die Anzahl guter Lehrbücher über Analysis und Lineare Algebra wächst und wächst. Es ist für Sie einfach und umsonst, die Bibliothek der Mathematischen Institute im Erdgeschoß aufzusuchen. Im unteren Geschoß, gleich links neben der Treppe ist die **Lehrbuchsammlung Mathematik**.
Unter der Rubrik "Analysis" finden Sie zum Beispiel

1. H. Neunzert, W.G.Eschmann, et. al.: **Analysis 1, Analysis 2**.
   Springer-Verlag 1993.

   Diese beiden Bücher sind einfach zu lesen und decken sorgfältig und ausführlich den Inhalt der Kapitel I - IV meiner Vorlesung ab, insbesondere auch mein Kapitel III über Lineare Algebra.

2. Christian Blatter: **Analysis I, Analysis II**.
   Springer-Verlag 1997, 1995.

   Diese beiden Bände sind ebenfalls einfach und verständlich geschrieben. Sie decken meine Kapitel I, II, IV, V ab, also auch das Kapitel V über Vektoranalysis, aber nicht mein Kapitel III über Lineare Algebra.

Unter der Rubrik "Linearen Algebra" der **Lehrbuchsammlung Mathematik** finden Sie zum Beispiel

Howard Anton: **Lineare Algebra**.
Spektrum, Akademischer Verlag, 1998.

K. Jänisch: **Lineare Algebra**.
Springer-Verlag 2000.

Beide Bücher sind sehr klar und einfach geschrieben und decken mein Kapitel III ab.

Kapitel 3 : Lineare Algebra

§1. Linearer Raum

In diesem Kapitel bedeutet $\mathbf{K}$ den Körper der reellen Zahlen $\mathbb{R}$ oder den Körper der komplexen Zahlen $\mathbb{C}$. Sie können also bei jeder Definition oder jedem Satz $\mathbf{K}$ ersetzen durch $\mathbb{R}$ oder auch durch $\mathbb{C}$.

Es sei n eine natürliche Zahl. Mit $\mathbb{R}^n$ bezeichnen wir die Menge aller (geordneten) n-Tupel reeller Zahlen, also

$$\mathbb{R}^n := \Big\{ a = (a_1, a_2, \ldots, a_n) : \ \ a_i \in \mathbb{R}, \ \ i = 1, 2, \ldots, n \Big\},$$

mit $\mathbb{C}^n$ bezeichnen wir die Menge aller (geordneten) n-Tupel komplexer Zahlen, also

$$\mathbb{C}^n := \Big\{ a = (a_1, a_2, \ldots, a_n) : \ \ a_i \in \mathbb{C}, \ \ i = 1, 2, \ldots, n \Big\}.$$

Anstelle von $(a_1, a_2, \ldots, a_n)$ schreiben wir oft auch

$$\begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix}.$$

Im $\mathbf{K}^n$ definieren wir eine **Addition** $+ : \mathbf{K}^n \times \mathbf{K}^n \to \mathbf{K}^n$ wie folgt:
Seien

$$a = \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix} \quad , \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{pmatrix},$$

dann definieren wir

$$a + b := \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n + b_n \end{pmatrix},$$

daß heißt, wir addieren a und b **komponentenweise**.

Im $\mathbf{K}^n$ definieren wir die **skalare Multiplikation** $* : \mathbf{K} \times \mathbf{K}^n \to \mathbf{K}^n$ wie folgt:

$$\alpha * a := \begin{pmatrix} \alpha a_1 \\ \alpha a_2 \\ \cdot \\ \cdot \\ \cdot \\ \alpha a_n \end{pmatrix}, \quad \text{für } \alpha \in \mathbf{K}, \; a \in \mathbf{K}^n,$$

das heißt, wir multiplizieren jede **Komponente** von a mit α. Anstelle von $\alpha * a$ schreiben wir kürzer αa.

Wir können leicht nachprüfen, daß die folgenden Rechenregeln gelten:
(A1) $a + b = b + a$ für alle $a, b \in \mathbf{K}^n$, (kommutativ)
(A2) $(a + b) + c = a + (b + c)$ für alle $a, b, c \in \mathbf{K}^n$, (assoziativ)
(A3) Für das **Element** $\mathbf{0} = (0, \ldots, 0) \in \mathbf{K}^n$ gilt $a + \mathbf{0} = a$ für alle $a \in \mathbf{K}^n$,
(A4) Für jedes $a \in \mathbf{K}^n$ gibt es genau ein $b \in \mathbf{K}^n$, nämlich $b = (-a_1, -a_2, \ldots, -a_n)$ mit der Eigenschaft $a + b = \mathbf{0}$. Wir schreiben auch $b = -a$.

(B1) $\alpha(a + b) = \alpha a + \alpha b$ für alle $a, b \in \mathbf{K}^n$, $\alpha \in \mathbf{K}$,
(B2) $(\alpha + \beta)a = \alpha a + \beta a$ für alle $a \in \mathbf{K}^n$, $\alpha, \beta \in \mathbf{K}$,
(B3) $\alpha(\beta a) = (\alpha\beta)a$ für alle $a \in \mathbf{K}^n$, $\alpha, \beta \in \mathbf{K}$,
(B4) $1a = a$ für alle $a \in \mathbf{K}^n$.

Der $\mathbb{R}^n$ und der $\mathbb{C}^n$ mit der oben definierten Addition und skalaren Multiplikation sind wichtige Beispiele der folgenden allgemeinen Definition;

Definition 1.1. *Eine Menge V von Elementen $a, b, c, \ldots$ heißt* **linearer Raum über K** *oder* **Vektorraum über K**, *wenn eine Vektoraddition $+ : V \times V \to V$ und eine skalare Multiplikation $* : \mathbf{K} \times V \to V$ definiert sind mit den Eigenschaften (A1)-(A4) und (B1)-(B4).*

Beispiele:

1. Der $\mathbb{R}^n$ ist ein linearer Raum über $\mathbb{R}$, der $\mathbb{C}^n$ ist ein linearer Raum über $\mathbb{C}$.
2. Es sei $a < b$ gegeben. Es bezeichne $C[a, b]$ die Menge aller auf $[a, b]$ stetigen reellwertigen Funktionen. In $C[a, b]$ definieren wir die **Addition** durch

$$f + g : \quad (f + g)(x) := f(x) + g(x), \quad x \in [a, b],$$

und die **skalare Multiplikation** durch

$$\alpha f : \quad (\alpha f)(x) := \alpha f(x), \quad x \in [a, b].$$

Dann ist $C[a,b]$ ein linearer Raum über $\mathbb{R}$.

3. Zu $n \in \mathbb{N}$ sei $V = \mathcal{P}_n$ die Menge aller reellwertigen algebraischen Polynome vom Grade $\leq n$, versehen mit der Addition und skalaren Multiplikation des Beispiels 2.

4. Gegeben seien die Funktionen $c_0, c_1, c_2 \in C[a,b]$. Die Lösungen der lineare Differentialgleichung

$$c_0(x)y''(x) + c_1(x)y'(x) + c_2(x)y(x) = 0, \quad a \leq x \leq b,$$

bilden einen linearen Raum über $\mathbb{R}$ (bezüglich der Addition und skalaren Multiplikation des Beispiels 2).

Definition 1.2. *Es sei V ein linearer Raum über* $\mathbf{K}$. *Es heißt $b \in V$* **Linearkombination** *von $v_1, \ldots, v_n \in V$, falls es Zahlen $\alpha_1, \ldots, \alpha_n \in \mathbf{K}$ gibt, so daß*

$$b = \sum_{k=1}^{n} \alpha_k v_k.$$

Die Menge aller Linearkombinationen von $v_1, \ldots, v_n$ heißt **lineare Hülle** *von $v_1, \ldots, v_n$, in Zeichen span$\{v_1, \ldots, v_n\}$ oder $[v_1, \ldots, v_n]$.*

Beispiel 5:

$$\mathrm{span}\{1, \cos x, \sin x, \cos(2x), \sin(2x)\} = \mathcal{T}_2$$

sind die reellen trigonometrischen Polynome vom Grad ≤ 2.

Definition 1.3. *Es sei V ein linearer Raum über* $\mathbf{K}$. *Die Vektoren $v_1, \ldots, v_n \in V$ heißen* **linear abhängig**, *wenn es Zahlen $(\alpha_1, \ldots, \alpha_n) \in \mathbf{K}^n$ gibt, $(\alpha_1, \ldots, \alpha_n) \neq \mathbf{0}$, so daß*

$$\sum_{k=1}^{n} \alpha_k v_k = 0 \quad \textit{Nullvektor in } V.$$

Ansonsten heißen $v_1, \ldots, v_n \in V$ **linear unabhängig**.

Wie müssen wir vorgehen, wenn wir die lineare Unabhängigkeit von Vektoren $v_1, \ldots, v_n$ nachweisen wollen ?
Wir nehmen an, für die Zahlen $\alpha_1, \ldots, \alpha_n$ gelte

$$\sum_{k=1}^{n} \alpha_k v_k = \mathbf{0}$$

und beweisen, daß dann alle $\alpha_1, \ldots, \alpha_n$ gleich 0 sind.

Beispiel 6: Die Vektoren

$$\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \quad , \quad \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} \quad , \quad \begin{pmatrix} 4 \\ -2 \\ 2 \end{pmatrix}$$

des $\mathbb{R}^3$ sind linear unabhängig.

Beweis: Wir nehmen an, für die Zahlen $\alpha_1, \alpha_2, \alpha_3$ gelte

$$\alpha_1 \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 4 \\ -2 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Dies ist gleichbedeutend mit

$$\begin{aligned} 2\alpha_1 + \alpha_2 + 4\alpha_3 &= 0 \\ 3\alpha_2 - 2\alpha_3 &= 0 \\ 2\alpha_3 &= 0. \end{aligned}$$

Aus der 3.Gleichung folgt nun, daß $\alpha_3 = 0$ ist, aus der 2. Gleichung folgt $\alpha_2 = 0$ und aus der 1.Gleichung $\alpha_1 = 0$.

Beispiel 7: Es sei $n \in \mathbb{N}$, und $V = \mathcal{P}_n$ sei die Menge aller reellen algebraischen Polynome vom Grad $\leq n$. Die Polynome $P_0, P_1, \ldots, P_n$ seien von der Form

$$P_k(x) = \sum_{j=0}^{k} c_{j,k} x^j, \quad c_{j,k} \in \mathbb{R}, \quad c_{k,k} \neq 0.$$

Dann sind die Polynome $P_0, P_1, \ldots, P_n$ linear unabhängig.

Beweis: Wir nehmen an, für die Zahlen $\alpha_0, \ldots, \alpha_n$ gelte

$$\sum_{k=0}^{n} \alpha_k P_k(x) = 0 \quad \text{für alle } x \in \mathbb{R}.$$

Dies ist äquivalent zu

$$0 = \sum_{k=0}^{n} \alpha_k \Big(\sum_{j=0}^{k} c_{j,k} x^j \Big) = \sum_{j=0}^{n} x^j \Big(\sum_{k=j}^{n} \alpha_k c_{j,k} \Big) \quad \text{für alle } x \in \mathbb{R}.$$

Dies ist genau dann der Fall, wenn alle *Koeffizienten*

$$\beta_j := \sum_{k=j}^{n} \alpha_k c_{j,k}, \quad j = 0, 1, \ldots, n,$$

gleich Null sind. Aus $0 = \beta_n = \alpha_n c_{n,n}$ folgt $\alpha_n = 0$.
Aus

$$0 = \beta_{n-1} = \alpha_{n-1} c_{n-1,n-1} + \alpha_n c_{n-1,n}$$

folgt $\alpha_{n-1} c_{n-1,n-1} = 0$ und somit $\alpha_{n-1} = 0$, und so weiter.

Definition 1.4. *Es sei V ein linearer Raum über $\mathbf{K}$. Die Vektoren $v_1, \dots, v_n \in V$ bilden eine* **Basis** *von V, falls*

1. *$v_1, \dots, v_n$ linear unabhängig sind, und*
2. *$V = span\{v_1, \dots, v_n\}$.*

Beispiel 8: Es sei $n \in \mathbb{N}$. Die **Monome** $1, x, \dots, x^n$ bilden eine Basis von $\mathcal{P}_n$.

Satz 1.1. *Es sei $v_1, \dots, v_n$ eine Basis von V. Dann gilt:*

1. *Jede andere Basis von V besteht ebenfalls aus n Vektoren. Wir sagen: n ist die Dimension von V und schreiben $n = \dim V$.*
2. *Sei $m > n$. Je m Vektoren aus V sind linear abhängig.*
3. *Je n linear unabhängige Vektoren aus V bilden eine Basis von V.*
4. *Jedes $v \in V$ hat eine eindeutige Darstellung der Form*

$$v = \sum_{k=1}^{n} \alpha_k v_k, \quad \alpha_k \in \mathbb{R}.$$

Beweis: Satz 1.1 folgt sofort aus dem **Eliminationsverfahren von Gauß** des §3 und dem folgenden Beispiel 9. ■

Beispiel 9: Es sei $n \in \mathbb{N}$. Die Vektoren

$$v_k := \begin{pmatrix} a_{1,k} \\ a_{2,k} \\ \cdots \\ \cdots \\ \cdots \\ a_{k,k} \\ 0 \\ \cdots \\ \cdots \\ 0 \end{pmatrix}, \quad a_{k,k} \neq 0, \quad k = 1, 2, \dots, n,$$

bilden eine Basis des $\mathbf{K}^n$.
Beweis: Folge dem Beweis des Beispiels 7. ■

Definition 1.5. *Eine nichtleere Teilmenge V_0 eines linearen Raumes V heißt* **linearer Unterraum von** V*, falls V_0 selbst ein linearer Raum ist.*

Um zu zeigen, daß $V_0 \subset V$ ein linearer Unterraum ist, muß man wegen Definition 1.1 nachprüfen :

1. aus $u \in V_0$ und $v \in V_0$ folgt $u + v \in V_0$,
2. aus $u \in V_0$ und $\alpha \in \mathbf{K}$ folgt $\alpha u \in V_0$.
   Insbesondere muß $0 \in V_0$ liegen.

§2. Lineare Gleichungssysteme

In den Anwendungen treten häufig **lineare Gleichungssysteme** auf. Das Lösen von linearen Gleichungssystemen ist eine der wichtigsten Aufgaben der Angewandten Mathematik.

Ein $m \times n$ **lineares Gleichungssystem** besteht aus den m linearen Gleichungen

$$\sum_{k=1}^{n} a_{j,k} x_k = b_j, \quad j = 1, \dots, m, \tag{2.1}$$

für die n Unbekannten $x_1, \dots, x_n$.

Um das lineare Gleichungssystem (2.1) kompakter zu schreiben, führt man die $m \times n$-Matrix

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & \cdots & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & \cdots & \cdots & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & \cdots & \cdots & a_{3,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m,1} & a_{m,2} & a_{m,3} & \cdots & \cdots & \cdots & a_{m,n} \end{pmatrix}$$

mit den n **Spaltenvektoren**

$$\begin{pmatrix} a_{1,k} \\ a_{2,k} \\ a_{3,k} \\ \cdots \\ \cdots \\ \cdots \\ a_{m,k} \end{pmatrix} \in \mathbf{K}^m, \quad k = 1, \dots, n, \quad \Big(m \times 1 - \ \text{Matrix}\Big)$$

und den m **Zeilenvektoren**

$$(a_{j,1}, a_{j,2}, a_{j,3}, \dots, a_{j,n}) \in \mathbf{K}^n, \quad j = 1, 2, \dots, m. \quad \Big(1 \times n - \ \text{Matrix}\Big)$$

Das lineare Gleichungssystem können wir schreiben als

$$x_1 \begin{pmatrix} a_{1,1} \\ a_{2,1} \\ a_{3,1} \\ \cdots \\ \cdots \\ \cdots \\ a_{m,1} \end{pmatrix} + x_2 \begin{pmatrix} a_{1,2} \\ a_{2,2} \\ a_{3,2} \\ \cdots \\ \cdots \\ \cdots \\ a_{m,2} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1,n} \\ a_{2,n} \\ a_{3,n} \\ \cdots \\ \cdots \\ \cdots \\ a_{m,n} \end{pmatrix} = \begin{pmatrix} b_1 \\ \\ b_2 \\ \cdots \\ \cdots \\ \cdots \\ b_m \end{pmatrix}. \tag{2.2}$$

Aus (2.2) folgt sofort der nächste

Satz 2.1. *Wir schreiben $Ax = b$ mit*

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ \cdots \\ \cdots \\ x_n \end{pmatrix} \quad , \quad b = \begin{pmatrix} b_1 \\ \\ b_2 \\ \cdots \\ \cdots \\ \cdots \\ b_m \end{pmatrix}. \tag{2.3}$$

Das lineare Gleichungssystem (2.2) hat eine Lösung $x \in \mathbf{K}^n$ dann und nur dann, wenn die rechte Seite b als Linearkombination (2.2) der Spaltenvektoren von A darstellbar ist.

Über die Struktur der Lösungsgesamtheit eines linearen Gleichungssystem geben die beiden nächsten Sätze Auskunft:

Satz 2.2. *Die Lösungen $x \in \mathbf{K}^n$ des* **homogenen** *linearen Gleichungssystems*

$$\sum_{k=1}^{n} a_{j,k} x_k = 0, \quad j = 1, \ldots, m, \tag{2.4}$$

bilden einen linearen Unterraum des $\mathbf{K}^n$, genannt **Kern** *von A.*

Beweis: Es seien $x \in \mathbf{K}^n$ und $x' \in \mathbf{K}^n$ Lösungen von (2.4). Dann sind auch $x + x'$ und αx, $\alpha \in \mathbf{K}$, Lösungen von (2.4), denn es gilt

$$\sum_{k=1}^{n} a_{j,k}(x_k + x'_k) = \sum_{k=1}^{n} a_{j,k} x_k + \sum_{k=1}^{n} a_{j,k} x'_k = 0, \quad j = 1, \ldots, m,$$

sowie

$$\sum_{k=1}^{n} a_{j,k}(\alpha x_k) = \alpha \sum_{k=1}^{n} a_{j,k} x_k = 0, \quad j = 1, \ldots, m.$$

■

Satz 2.3. *Das lineare Gleichungssystem (2.1) sei lösbar. Sei x^0 eine Lösung von (2.1). Sei V_0 der Kern von A. Dann beschreibt*

$$\Big\{x^0 + v \ , \quad v \in V_0\Big\} \tag{2.5}$$

die Gesamtheit aller Lösungen von (2.1).
Insbesondere ist x^0 die einzige Lösung von (2.1), wenn $V_0 = \{\mathbf{0}\}$ ist.

Beweis:
Es sei $x = (x_1, \ldots, x_n) \in V_0$. Dann ist

$$\sum_{k=1}^{n} a_{j,k}(x_k^0 + x_k) = \sum_{k=1}^{n} a_{j,k}x_k^0 + \sum_{k=1}^{n} a_{j,k}x_k = b_j + 0, \quad j = 1, \ldots, m.$$

Seien umgekehrt $x^0 = (x_1^0, \ldots, x_n^0)$ und $x^1 = (x_1^1, \ldots, x_n^1)$ Lösungen von (2.1). Dann gilt für $x := x^1 - x^0$, daß

$$\begin{aligned}\sum_{k=1}^{n} a_{j,k}x_k &= \sum_{k=1}^{n} a_{j,k}(x_k^1 - x_k^0) \\ &= \sum_{k=1}^{n} a_{j,k}x_k^1 - \sum_{k=1}^{n} a_{j,k}x_k^0 \\ &= b_j - b_j = 0, \quad j = 1, \ldots, m.\end{aligned}$$

Also ist $x = x^1 - x^0$ Lösung des homogenen linearen Gleichungssystems. ■

Beispiel: Es sei

$$\begin{aligned} 4x_1 + 3x_2 + 7x_3 &= 7 \\ x_1 - \ \ x_2 \qquad\quad &= 0 \end{aligned}$$

Dies schreiben wir auch kurz als

$$\begin{pmatrix} 4 & 3 & 7 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 0 \end{pmatrix}.$$

Dann ist die Lösungsgesamtheit gegeben durch

$$\left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \alpha \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \quad \alpha \in \mathbf{K} \right\},$$

aber auch durch

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \quad \alpha \in \mathbf{K} \right\}.$$

§3. Das Eliminationsverfahren von Gauß

Es soll also das lineare Gleichungssystem (LGS)

$$\sum_{k=1}^{n} a_{j,k} x_k = b_k, \quad j = 1, \dots, m, \tag{3.1}$$

gelöst werden. Dieses LGS ist vollständig beschrieben durch die $m \times (n+1)$-Matrix

$$(A, b) = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & \cdots & \cdots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & \cdots & \cdots & a_{2,n} & b_2 \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & \cdots & \cdots & a_{3,n} & b_3 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m,1} & a_{m,2} & a_{m,3} & \cdots & \cdots & \cdots & a_{m,n} & b_m \end{pmatrix}. \tag{3.2}$$

Grundlage des Eliminationsverfahrens ist der folgende Satz:

Satz 3.1. *Die Lösungsmenge des LGS (3.1) ändert sich nicht bei den folgenden* **elementaren Zeilenumformungen***:*

1. *zwei Zeilen von (3.2) werden vertauscht,*
2. *Eine Zeile in (3.2) wird mit $\alpha \in \mathbf{K}$, $\alpha \neq 0$, multipliziert,*
3. *in (3.2) wird das Vielfache einer Zeile zu einer anderen Zeile addiert.*

Grundgedanke des Eliminationsverfahrens:
Durch elementare Zeilenumformungen wird (3.2) übergeführt in die Form

$$(R, c) = \begin{pmatrix} r_{1,1} & r_{1,2} & r_{1,3} & \cdots & \cdots & \cdots & r_{1,n} & c_1 \\ 0 & r_{2,2} & r_{2,3} & \cdots & \cdots & \cdots & r_{2,n} & c_2 \\ 0 & 0 & r_{3,3} & \cdots & \cdots & \cdots & r_{3,n} & c_3 \\ 0 & 0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \cdots & \cdots & r_{m,n} & c_n \end{pmatrix}, \tag{3.3}$$

wobei die $m \times n$-Matrix R die folgende Form erhalten soll:
$r_{j,k} = 0$ für $1 \le k < j \le m$, falls $n \ge m$
$r_{j,k} = 0$ für $1 \le k < j \le m$ und $k \le n$, falls $n < m$.

Wie man im Einzelnen vorgehen sollte, wollen wir an zwei Beispielen beschreiben.

$$(A, b) = \left(\begin{array}{ccccc|c} 1 & 2 & 0 & 1 & 1 & 1 \\ 2 & 2 & -2 & 3 & 2 & 0 \\ 2 & 3 & -1 & 1 & 2 & 2 \end{array}\right)$$

$$\begin{pmatrix} 1 & 2 & 0 & 1 & 1 & | & 1 \\ 0 & -2 & -2 & 1 & 0 & | & -2 \\ 0 & -1 & -1 & -1 & 0 & | & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 0 & 1 & 1 & | & 1 \\ 0 & -2 & -2 & 1 & 0 & | & -2 \\ 0 & 0 & 0 & -1.5 & 0 & | & 1 \end{pmatrix}$$

Eine Lösung ist gegeben durch

$x_3^0 = 0$, $x_4^0 = -2/3$, $x_5^0 = 0$, $x_2^0 = 2/3$, $x_1^0 = 1/3$.

Die Lösungsgesamtheit ist

$$x = \frac{1}{3}\begin{pmatrix} 1 \\ 2 \\ 0 \\ -2 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} 2 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \alpha, \beta \in \mathbf{K}.$$

Sei

$$(A, b) = \begin{pmatrix} 0 & 2 & 1 & | & 1 \\ 1 & 2 & 1 & | & 0 \\ 2 & 4 & c & | & 2 \end{pmatrix}$$

mit irgendeiner Zahl $c \in \mathbf{K}$.

$$\begin{pmatrix} 1 & 2 & 1 & | & 0 \\ 0 & 2 & 1 & | & 1 \\ 2 & 4 & c & | & 2 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 1 & | & 0 \\ 0 & 2 & 1 & | & 1 \\ 0 & 0 & c-2 & | & 2 \end{pmatrix}.$$

Ist $c = 2$, so ist das LGS nicht lösbar. Ist $c \neq 2$, so folgt

$x_3 = 2/(c-2)$,

$x_2 = \frac{1}{2}\Big(1 - x_3\Big) = \frac{1}{2} - \frac{1}{c-2}$,

$x_1 = -2x_2 - x_3 = -1 + \frac{2}{c-2} - \frac{2}{c-2} = -1$.

Wir wollen nun ein **C**-Programm für das Eliminationsverfahren für **reelle quadratische** lineare Gleichungssysteme angeben (*quadratisch* bedeutet $m = n$). Aus Gründen der numerischen Stabilität enthält das Programm die **Spaltenpivotsuche** verbunden mit Zeilentausch.

```
/* gauss.c Eliminationsverfahren */
/* Die Datei "daten" wird eingelesen */
/*  n <= 100  */

#include<stdio.h>
#include<math.h>
#include<stdlib.h>
main(){
  FILE *datei1;
  int j,k,n;
  float x[101];
  float b[101];
  float matrix[101][101];

  /* Beginn der GaussElimination */
    void gauss(int d)
    {
      float s,t, factor;
      int i,j,k,r;
      float aa[101];

      for (k = 1; k <= d - 1; k++)
      {
        r=k;
        s = fabs(matrix[k][k]);
        for (i = k + 1; i <= d; i++)
            if (fabs(matrix[i][k])>s)
            {
                r=i;
                s=fabs(matrix[i][k]);
            }

            for (j = k; j <= d; j++)
            {
                aa[j]=matrix[k][j];
                matrix[k][j]=matrix[r][j];
                matrix[r][j]=aa[j];
            }

            t=b[k];
```

```
            b[k]=b[r];
            b[r]=t;

            for (j = k + 1; j <= d; j ++)
            {
                  factor=matrix[j][k]/matrix[k][k];
                  for (i = k; i <= d; i ++)
                        matrix[j][i] = matrix[j][i]-factor*matrix[k][i];
                  b[j] = b[j]-factor*b[k];
            }
      }

      x[d]=b[d]/matrix[d][d];
      for (j = d - 1; j >= 1; j - -)
      {
            x[j]=b[j];
            for (i = j + 1; i <= d; i ++)
                  x[j]=x[j]-matrix[j][i]*x[i];
            x[j]=x[j]/matrix[j][j];
      }
}            /* Ende der GaussElimination */

      datei1=fopen("daten","r");
      fscanf(datei1,"%10d\n ",&n);

      for (j = 1; j <= n; j ++)
      {
            for (k = 1; k <= n; k ++)
                  fscanf(datei1,"%10f ",&matrix[j][k]);
            fscanf(datei1,"%10f\n ",&b[j]);
      }
      printf("\n");
      gauss(n);
      for (j = 1; j <= n; j ++)
            printf(" x[%2d] %3.4f \n",j,x[j]);
      fclose(datei1);
}
```

§4. Der Rang einer Matrix

Definition 4.1. *Es sei A eine $m \times n$-Matrix. Die maximale Anzahl linear unabhängiger Spaltenvektoren von A heißt* **Spaltenrang** *von A, die maximale Anzahl linear unabhängiger Zeilenvektoren von A heißt* **Zeilenrang** *von A.*

Beispiel: *Die Matrix*

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 2 & 2 & 1 & 2 \\ 0 & 0 & 2 & 4 \end{pmatrix}$$

hat den Zeilenrang $= 2$ und den Spaltenrang $= 2$.

Satz 4.1. *Für jede $m \times n$-Matrix gilt* **Zeilenrang = Spaltenrang**. *Daher sprechen wir vom* **Rang** *einer Matrix.*

Beweis: Der Rang einer Matrix ändert sich nicht bei den elementaren Zeilenumformungen des Satzes 3.1, aber auch nicht bei Spaltentausch.
Ausgehend von der $m \times n$-Matrix A kann man durch elementare Zeilenumformungen und Spaltentausch stets die folgende Matrixform erreichen

$$R = \begin{pmatrix} r_{1,1} & r_{1,2} & r_{1,3} & \cdots & \cdots & \cdots & r_{1,n} \\ 0 & r_{2,2} & r_{2,3} & \cdots & \cdots & \cdots & r_{2,n} \\ 0 & 0 & r_{3,3} & \cdots & \cdots & \cdots & r_{3,n} \\ 0 & 0 & 0 & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & r_{q,q} & \cdots & r_{q,n} \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

mit $r_{j,j} \neq 0$ für $j = 1, \ldots, q$, $r_{j,k} = 0$ für $j > k$, $r_{j,k} = 0$ für $j > q$.
Für die Matrix R aber ist Zeilenrang = Spaltenrang = q. ■

Hieraus und aus Satz 3.1 folgt sofort der folgende wichtige

Satz 4.2. *Gegeben sei das LGS*

$$\sum_{k=1}^{n} a_{j,k} x_k = b_j, \quad j = 1, \ldots, m. \tag{4.1}$$

1. Das LGS ist lösbar dann und nur dann, wenn

$$Rang(A) = Rang(A, b).$$

2. Ist das LGS lösbar, dann hat die Lösungsmenge V_0 des homogenen linearen Gleichungssystems

$$\sum_{k=1}^{n} a_{j,k} x_k = 0, \quad j = 1, \ldots, m \tag{4.2}$$

die Dimension

$$n - Rang(A). \tag{4.3}$$

§5. Determinanten

Problem 1: Bestimme den Flächeninhalt des Parallelogramms, das von den beiden Vektoren

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \in \mathbb{R}^2 \quad , \quad \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{R}^2$$

aufgespannt wir.

Antwort: Der Flächeninhalt ist

$$a_1 b_2 - a_2 b_2 \quad =: \quad \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}.$$

Problem 2: Bestimme das Volumen des schiefen Würfels, der von den drei Vektoren

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \in \mathbb{R}^3 \quad , \quad \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \in \mathbb{R}^3 \qquad \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \in \mathbb{R}^3$$

aufgespannt wir.

Antwort: Das Volumen ist

$$a_1 b_2 c_3 - a_1 b_3 c_2 - a_2 b_3 c_1 - a_2 b_1 c_3 - a_3 b_1 c_2 - a_3 b_2 c_1 \quad =: \quad \det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}.$$

Der Flächeninhalt oder das Volumen sind dabei negativ, wenn die Vektoren negativ orientiert sind

Für allgemeines $n \in \mathbb{N}$ definiert man die Determinante einer $n \times n$-Matrix

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & \cdots & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & \cdots & \cdots & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & \cdots & \cdots & a_{3,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \cdots & \cdots & \cdots & a_{n,n} \end{pmatrix}$$

mit den n **Spaltenvektoren**

$$\mathbf{a}_k := \begin{pmatrix} a_{1,k} \\ a_{2,k} \\ a_{3,k} \\ \cdots \\ \cdots \\ \cdots \\ a_{n,k} \end{pmatrix} \in \mathbf{K}^n, \quad k = 1, \ldots, n,$$

in folgender Weise:
Es bezeichne $\mathcal{M}_n$ die Menge aller $n \times n$-Matrizen. Gesucht ist eine Abbildung det : $\mathcal{M}_n \to \mathbf{K}$ mit den Eigenschaften

(D1) Die Abbildung det ist linear bezüglich jeder Spalte von $A = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$; hierbei bedeutet die Linearität bezüglich der Spalte j, daß

$$\det\Big(\mathbf{a}_1, \ldots, \mathbf{a}_{j-1}, \mathbf{a}_j + \mathbf{b}_j, \mathbf{a}_{j+1}, \ldots, \mathbf{a}_n\Big) = \det\Big(\mathbf{a}_1, \ldots, \mathbf{a}_{j-1}, \mathbf{a}_j, \mathbf{a}_{j+1}, \ldots, \mathbf{a}_n\Big) + \det\Big(\mathbf{a}_1, \ldots, \mathbf{a}_{j-1}, \mathbf{b}_j, \mathbf{a}_{j+1}, \ldots, \mathbf{a}_n\Big)$$

$$\det\Big(\mathbf{a}_1, \ldots, \mathbf{a}_{j-1}, \alpha\mathbf{a}_j, \mathbf{a}_{j+1}, \ldots, \mathbf{a}_n\Big) = \alpha\det\Big(\mathbf{a}_1, \ldots, \mathbf{a}_{j-1}, \mathbf{a}_j, \mathbf{a}_{j+1}, \ldots, \mathbf{a}_n\Big)$$

für alle $\alpha \in \mathbf{K}$.

(D2) Beim Vertauschen zweier Spalten in $A = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$ ändert sich das Vorzeichen, aber der Betrag bleibt derselbe;

(D3) Die Determinante der $n \times n$-**Einheitsmatrix**

$$I_n = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n) = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

ist gleich 1.

Im Falle $n = 1$ ist $\det(\alpha) = \alpha\ \det(1) = \alpha, \quad \alpha \in \mathbf{K}$.

Im Falle $n = 2$ ist

$$\begin{aligned}
&\det\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} = \det\begin{pmatrix} a_{1,1}+0 & a_{1,2} \\ 0+a_{2,1} & a_{2,2} \end{pmatrix} \\
&= \det\begin{pmatrix} a_{1,1} & a_{1,2} \\ 0 & a_{2,2} \end{pmatrix} + \det\begin{pmatrix} 0 & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \\
&= a_{1,1}\det\begin{pmatrix} 1 & a_{1,2} \\ 0 & a_{2,2} \end{pmatrix} + a_{2,1}\det\begin{pmatrix} 0 & a_{1,2} \\ 1 & a_{2,2} \end{pmatrix} \\
&= a_{1,1}\det\begin{pmatrix} 1 & a_{1,2} \\ 0 & 0 \end{pmatrix} + a_{1,1}\det\begin{pmatrix} 1 & 0 \\ 0 & a_{2,2} \end{pmatrix} \\
&\quad + a_{2,1}\det\begin{pmatrix} 0 & a_{1,2} \\ 1 & 0 \end{pmatrix} + a_{2,1}\det\begin{pmatrix} 0 & 0 \\ 1 & a_{2,2} \end{pmatrix} \\
&= a_{1,1}a_{1,2}\det\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + a_{1,1}a_{2,2}\det\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
&\quad + a_{2,1}a_{1,2}\det\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + a_{2,1}a_{2,2}\det\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \\
&= 0 + a_{1,1}a_{2,2} - a_{2,1}a_{1,2} + 0.
\end{aligned}$$

Allgemein gilt

Satz 5.1. *Für jedes $n \in I\!N$ ist durch die Eigenschaften (D1), (D2), (D3) genau eine Abbildung det bestimmt, und zwar ist*

$$det\begin{pmatrix}
a_{1,1} & a_{1,2} & a_{1,3} & \cdots & \cdots & \cdots & a_{1,n} \\
a_{2,1} & a_{2,2} & a_{2,3} & \cdots & \cdots & \cdots & a_{2,n} \\
a_{3,1} & a_{3,2} & a_{3,3} & \cdots & \cdots & \cdots & a_{3,n} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
a_{n,1} & a_{n,2} & a_{n,3} & \cdots & \cdots & \cdots & a_{n,n}
\end{pmatrix} = \sum_{\sigma\in\Pi_n} sign(\sigma)\cdot a_{\sigma_1,1}\cdot a_{\sigma_2,2}\cdot\cdots\cdot a_{\sigma_n,n}.$$

Hierbei definieren wir

$$\Pi_n := \Big\{\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n) : \ \sigma \ \textit{ist Permutation der Zahlen } 1, 2, \ldots, n\Big\}$$

und

sign(σ) := 1, falls $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ durch eine **gerade** *Anzahl von Vertauschungen zweier Komponenten aus $(1, 2, \ldots, n)$ entstanden ist, und*
sign(σ) := −1, falls $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ durch eine **ungerade** *Anzahl von Vertauschungen zweier Komponenten aus $(1, 2, \ldots, n)$ entstanden ist.*

Wir prüfen sofort nach, daß durch Satz 5.1 im Falle $n = 3$ gerade die Lösungsformel des Problems 2 herauskommt.

Die Determinante hat die folgenden weiteren Eigenschaften:

Satz 5.2. *Es sei A eine $n \times n$-Matrix. Es bezeichne*

$$A^T = \begin{pmatrix} a_{1,1} & a_{2,1} & a_{3,1} & \cdots & \cdots & \cdots & a_{n,1} \\ a_{1,2} & a_{2,2} & a_{3,2} & \cdots & \cdots & \cdots & a_{n,2} \\ a_{1,3} & a_{2,3} & a_{3,3} & \cdots & \cdots & \cdots & a_{n,3} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{1,n} & a_{2,n} & a_{3,n} & \cdots & \cdots & \cdots & a_{n,n} \end{pmatrix}$$

die **transponierte Matrix** *der $n \times n$-Matrix A. Es gilt*

$$det(A^T) = det(A). \tag{5.1}$$

Beweis: Dies folgt sofort aus der Formel des Satzes 5.1. ■

Satz 5.3. *Es sei A eine $n \times n$-Matrix. Es ist sind die folgenden Aussagen äquivalent:*

$det(A) = 0$
$det(A^T) = 0$
die Spaltenvektoren von A sind linear abhängig
die Zeilenvektoren von A sind linear abhängig
$Rang(A) = Rang(A^T) \leq n - 1.$

Zur **Berechnung der Determinante** $\det(A)$ benutzen wir das Gaußsche Eliminationsverfahren unter Verwendung des folgenden Satzes.

Satz 5.4. *Es sei*

$$A = \begin{pmatrix} a^1 \\ a^2 \\ \cdots \\ \cdots \\ a^n \end{pmatrix}$$

eine $n \times n$-Matrix mit den Zeilenvektoren

$$a^j = (a_{j,1}, a_{j,2}, \cdots, a_{j,n}), \quad j = 1, \ldots, n.$$

Es gelten die Rechenregeln

$$det \begin{pmatrix} a^1 \\ a^2 \\ \cdots \\ \alpha a^j \\ \cdots \\ a^n \end{pmatrix} = \alpha \;\; det(A), \quad \alpha \in \mathbf{K}.$$

Für $k \neq j$ ist

$$\det \begin{pmatrix} a^1 \\ \cdots \\ a^{j-1} \\ a^j + \alpha a^k \\ a^{j+1} \\ \cdots \\ a^n \end{pmatrix} = \det(A), \quad \alpha \in \mathbf{K},$$

$$\det \begin{pmatrix} a^1 \\ \cdots \\ a^j \\ \cdots \\ a^k \\ \cdots \\ a^n \end{pmatrix} = -\det \begin{pmatrix} a^1 \\ \cdots \\ a^k \\ \cdots \\ a^j \\ \cdots \\ a^n \end{pmatrix}.$$

Beweis: Wir verwenden die Aussage $\det(A^T) = \det(A)$ des Satzes 5.2 und verwenden für A^T die Eigenschaften (D1) und (D2). ■

Wir wollen diese Regeln an einem Beispiel anwenden:

$$\det \begin{pmatrix} 0 & 3 & 3 & 0 \\ 1 & 5 & 0 & 2 \\ 2 & 2 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} = -\det \begin{pmatrix} 1 & 5 & 0 & 2 \\ 0 & 3 & 3 & 0 \\ 2 & 2 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$= -\det \begin{pmatrix} 1 & 5 & 0 & 2 \\ 0 & 3 & 3 & 0 \\ 0 & -8 & 1 & -3 \\ 0 & 1 & 0 & 0 \end{pmatrix} = -3\det \begin{pmatrix} 1 & 5 & 0 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & -8 & 1 & -3 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$= -3\det \begin{pmatrix} 1 & 5 & 0 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 9 & -3 \\ 0 & 0 & -1 & 0 \end{pmatrix} = 3\det \begin{pmatrix} 1 & 5 & 0 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 9 & -3 \end{pmatrix}$$

$$= 3\det \begin{pmatrix} 1 & 5 & 0 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix} = 9,$$

denn eine **obere** $n \times n$ **Dreiecksmatrix**

$$R = \begin{pmatrix} r_{1,1} & r_{1,2} & r_{1,3} & \cdots & \cdots & \cdots & r_{1,n} \\ 0 & r_{2,2} & r_{2,3} & \cdots & \cdots & \cdots & r_{2,n} \\ 0 & 0 & r_{3,3} & \cdots & \cdots & \cdots & r_{3,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \cdots & 0 & r_{n,n} \end{pmatrix}$$

hat nach Satz 5.1 die Determinante

$$\det(R) = \prod_{j=1}^{n} r_{j,j}.$$

Enthält eine $n \times n$- Matrix A viele Nullen, so kann man zur Berechnung der Determinante den folgenden **Entwicklungssatz** mit Erfolg anwenden:

Satz 5.5. *Es sei A eine $n \times n$-Matrix. Für $1 \leq j,k \leq n$ bezeichnen wir mit $A_{j,k}$ die $(n-1) \times (n-1)$-Matrix, die aus A durch Streichen der Zeile j und der Spalte k hervorgeht.*

1. **Entwickeln nach Zeile** j: *Es gilt*

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

2. **Entwickeln nach Spalte** k: *Es gilt*

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

Wenden wir den Entwicklungssatz auf das letzte Beispiel an, so erhalten wir

$$\det \begin{pmatrix} 0 & 3 & 3 & 0 \\ 1 & 5 & 0 & 2 \\ 2 & 2 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \det \begin{pmatrix} 0 & 3 & 0 \\ 1 & 0 & 2 \\ 2 & 1 & 1 \end{pmatrix}$$
$$= -3\det \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = -3(1-4) = 9.$$

Von theoretischer, aber nicht von numerischer Bedeutung ist die **Cramersche Regel** zur Lösung von LGS:

Satz 5.6. *Gegeben sei das LGS*

$$\sum_{k=1}^{n} a_{j,k} x_k = b_j, \quad j = 1, \dots, n.$$

Erfüllt die Matrix $A = (a_{j,k})_{j,k=1}^{n}$ *die Bedingung* $det(A) \neq 0$, *so ist die Lösung gegeben durch*

$$x_j = \frac{1}{det(A)} det \begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1,n} \\ a_{2,1} & \cdots & a_{2,j-1} & b_2 & a_{2,j+1} & \cdots & a_{2,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}, \quad j = 1, \dots, n.$$

Beweis: Es bezeichne

$$b := \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_n \end{pmatrix}, \quad \mathbf{a}_k := \begin{pmatrix} a_{1.k} \\ a_{2.k} \\ \cdot \\ \cdot \\ a_{n,k} \end{pmatrix}, \quad k = 1, \dots, n.$$

Das LGS ist äquivalent zu

$$x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \cdots + x_n \mathbf{a}_n = b.$$

Daher ist für $j = 1, 2, \dots, n$

$$\begin{aligned} &\det\Big(\mathbf{a}_1, \dots, \mathbf{a}_{j-1}, b, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n\Big) \\ &= \det\Big(\mathbf{a}_1, \dots, \mathbf{a}_{j-1}, \sum_{k=1}^{n} x_k \mathbf{a}_k, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n\Big) \\ &= \sum_{k=1}^{n} x_k \det\Big(\mathbf{a}_1, \dots, \mathbf{a}_{j-1}, \mathbf{a}_k, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n\Big) \\ &= x_j \det\Big(\mathbf{a}_1, \dots, \mathbf{a}_{j-1}, \mathbf{a}_j, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n\Big) = x_j \det(A), \end{aligned}$$

quod erat demonstrandum. ■

Man nennt eine $n \times n$-Matrix A **regulär** oder **nicht-singulär**, falls $\det(A) \neq 0$ ist. Ist $\det(A) = 0$, so heißt A **singulär** .

Aus den bisherigen Diskussionen können wir den folgenden **sehr wichtigen** Satz über LGS gewinnen:

Satz 5.7. *Gegeben sei die $n \times n$-Matrix A. Dann sind die folgenden Aussagen äquivalent:*

1. *$det(A) \neq 0$,*
2. *Das LGS $Ax = b$ ist für jede rechte Seite $b \in \mathbf{K}^n$ eindeutig lösbar,*
3. *Das homogene LGS $Ax = 0$ hat nur die Lösung $x = \mathbf{0}$.*

Aus dem letzten Satz können wir eine wichtige Anwendung aus der Interpolationstheorie herleiten:

Interpolationsproblem:
Sei $d \in \mathbb{N}$ und $\Omega \subset \mathbb{R}^d$. Es seien n linear unabhängige Funktionen $u_k : \Omega \to \mathbb{R}$, $k = 1, \ldots, n$, gegeben, die einen linearen Raum $V := \text{span}\{u_1, u_2, \ldots, u_n\}$ aufspannen. Zu gegebenen n verschiedenen Punkten (Orten) $x_i \in \Omega$, $i = 1, \ldots, n$ und n reellen Zahlen y_i, $i = 1, \ldots, n$, bestimme man $v \in V$ mit der Eigenschaft

$$v(x_i) = y_i, \quad i = 1, \ldots, n.$$

Lösung:
Das Interpolationsproblem kann keine, eine oder viele Lösungen haben. Wir machen den Ansatz

$$v = \sum_{k=1}^{n} c_k u_k.$$

Dann können wir das Interpolationsproblem schreiben als lineares Gleichungssystem

$$\sum_{k=1}^{n} c_k u_k(x_i) = y_i, \quad i = 1, \ldots, n.$$

Oder in Matrixschreibweise :

$$Ac = b, \quad \begin{pmatrix} u_1(x_1) & u_2(x_1) & u_3(x_1) & \cdots & \cdots & u_n(x_1) \\ u_1(x_2) & u_2(x_2) & u_3(x_2) & \cdots & \cdots & u_n(x_2) \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ u_1(x_n) & u_2(x_n) & u_3(x_n) & \cdots & \cdots & u_n(x_n) \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ c_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_n \end{pmatrix}.$$

Wegen Satz 5.7 sind daher die folgenden Aussagen äquivalent:

(a) $\det(A) \neq 0$,
(b) Das Interpolationsproblem ist für jedes n-Tupel $y_1, y_2, \ldots, y_n$ eindeutig lösbar,
(c) Das homogene Interpolationsproblem $v(x_i) = 0$, $i = 1, \ldots, n$, hat nur die Lösung $v = 0$ (Nullfunktion).

Wir wollen dies für die Interpolation durch algebraische Polynome $V = \mathcal{P}_n$ vom Grade $\leq n$ anwenden. (**Siehe auch Übungsaufgabe 11**). Nach dem Fundamentalsatz der Algebra wissen wir, daß zu gegebenen $n + 1 = \dim(\mathcal{P}_n)$ verschiedenen reellen oder komplexen Stellen x_i, $i = 0, \ldots, n$, das homogene Interpolationsproblem nur das Nullpolynom als Lösung hat. Also gilt die Eigenschaft (c). Hierzu ist (b) äquivalent, womit bewiesen ist, daß auch das inhomogene Interpolationsproblem für $V = \mathcal{P}_n$ und beliebige Werte y_i, $i = 0, \ldots, n$, eine eindeutige Lösung besitzt.

§6. Euklidische Vektorräume. Inneres Produkt

Beispiele linearer Räume mit innerem Produkt:

1. Der $\mathbb{R}^2$ mit dem inneren Produkt

$$\langle a, b\rangle = a_1 b_1 + a_2 b_2, \quad a = (a_1, a_2), \;\; b = (b_1, b_2).$$

Die Zahl

$$\|a\|_2 := \sqrt{\langle a, a\rangle} = \sqrt{a_1^2 + a_2^2}$$

nennen wir die **Euklidische Norm** von $a = (a_1, a_2) \in \mathbb{R}^2$.

2. Der $\mathbb{R}^3$ mit dem inneren Produkt

$$\langle a, b\rangle = a_1 b_1 + a_2 b_2 + a_3 b_3, \quad a = (a_1, a_2, a_3), \;\; b = (b_1, b_2, b_3).$$

Die Zahl

$$\|a\|_2 := \sqrt{\langle a, a\rangle} = \sqrt{a_1^2 + a_2^2 + a_3^2}$$

nennen wir die **Euklidische Norm** von $a = (a_1, a_2, a_3) \in \mathbb{R}^3$.
In einem rechtwinkligen Kooordinatensystem mit den Einheits-Achsen $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ ist der Winkel γ zwischen den Ortsvektoren

$$\mathbf{x} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_3\mathbf{e}_3 \quad , \quad \mathbf{y} = b_1\mathbf{e}_1 + b_2\mathbf{e}_2 + b_3\mathbf{e}_3$$

gegeben durch

$$\cos\gamma \;\; = \;\; \frac{\langle a, b\rangle}{\|a\|_2 \|b\|_2}.$$

3. Der $\mathbb{R}^n$, $n \in \mathbb{N}$, mit dem inneren Produkt

$$\langle a, b\rangle = \sum_{k=1}^{n} a_k b_k, \quad a = (a_1, \ldots, a_n), \;\; b = (b_1, \ldots, b_n).$$

Die Zahl

$$\|a\|_2 := \sqrt{\langle a, a\rangle} = \Big(\sum_{k=1}^{n} a_k^2\Big)^{1/2} \tag{6.1}$$

nennen wir die **Euklidische Norm** von $a \in \mathbb{R}^n$.

4. In $C[a, b]$, dem linearen Raum aller stetigen reellwertigen Funktionen auf dem Intervall $[a, b]$, definieren wir das innere Produkt

$$\langle f, g\rangle := \int_a^b f(x)g(x)dx$$

und die L_2-Norm

$$\|f\|_2 := \sqrt{\langle a, a\rangle} = \Big(\int_a^b f(x)^2 dx\Big)^{1/2}.$$

Um eine einheitliche mathematische Theorie für diese und viele andere Beispiele zu entwickeln, definieren wir abstrakt und axiomatisch:

Definition 6.1. *Es sei V ein linearer Raum über $\mathbb{R}$. Eine Abbildung $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ heißt* **inneres Produkt** *in V, falls*

(L) (Linearität in der 1.−ten Komponente von $\langle\cdot,\cdot\rangle$)

$\langle f_1 + f_2, g\rangle = \langle f_1, g\rangle + \langle f_2, g\rangle$ *für alle* $f_1, f_2, g \in V$,

$\langle \alpha f, g\rangle = \alpha\langle f, g\rangle$ *für alle* $f, g \in V$ *und* $\alpha \in \mathbb{R}$,

(S) (Symmetrie)

$\langle f, g\rangle = \langle g, f\rangle$ *für alle* $f, g \in V$,

(P) (Positivität)

$\langle f, f\rangle \geq 0$ *für alle* $f \in V$,

(P)* $\langle f, f\rangle = 0$ *genau dann, wenn* $f = 0$.

Aus (L) und (S) folgt sofort die Linearität des inneren Produkts bezüglich der 2. Komponente:

(L*) (Linearität in der 2.Komponente von $\langle\cdot,\cdot\rangle$)

$\langle f, g_1 + g_2\rangle = \langle f, g_1\rangle + \langle f, g_2\rangle$ für alle $f, g_1, g_2 \in V$,

$\langle f, \alpha g\rangle = \alpha\langle f, g\rangle$ für alle $f, g \in V$ und $\alpha \in \mathbb{R}$,

Definition 6.2. *Es sei V ein linearer Raum über $\mathbb{R}$. Eine Abbildung $\|\cdot\| : V \to \mathbb{R}$ heißt* **Norm** *in V, falls*

(N1) $\|\alpha f\| = |\alpha|\|f\|$ *für alle* $f \in V$ *und alle* $\alpha \in \mathbb{R}$,

(N2) (Dreiecksungleichung)

$\|f + g\| \leq \|f\| + \|g\|$ *für alle* $f, g \in V$,

(N3) (Positivität)

$\|f\| \geq 0$ *für alle* $f \in V$,

$\|f\| = 0$ *genau dann, wenn* $f = 0$.

Satz 6.1. *Es sei V ein linearer Raum über $\mathbb{R}$ mit dem inneren Produkt $\langle\cdot,\cdot\rangle$. Dann wird durch*

$$\|f\| := \Big(\langle f, f\rangle\Big)^{1/2}, \quad f \in V,$$

eine Norm in V definiert.

Beweis: Es ist einfach, die Eigenschaften (N1) und (N3) nachzuprüfen. Um die Dreiecksungleichung (N2) zu beweisen, müssen wir zuerst die **Cauchy-Schwarzsche Ungleichung**

$$|\langle f, g\rangle| \leq \|f\|\|g\|, \quad f, g \in V, \tag{6.2}$$

beweisen, was in den Übungen geschehen soll. Denn mit Hilfe von (6.2) folgt

$$\begin{aligned}\|f+g\|^2 &= \langle f+g, f+g\rangle = \langle f,f\rangle + \langle f,g\rangle + \langle g,f\rangle + \langle g,g\rangle \\ &= \|f\|^2 + 2\langle f,g\rangle + \|g\|^2 \\ &\le \|f\|^2 + 2\|f\|\|g\| + \|g\|^2 \\ &= (\|f\| + \|g\|)^2,\end{aligned}$$

womit die Dreiecksungleichung bewiesen ist. ■

Definition 6.3.
Sind in Definition 6.1 die Bedingungen (L), (S), (P), aber nicht (P) erfüllt, so heißt* $\langle\cdot,\cdot\rangle$ *ein* **semidefinites** *inneres Produkt und* $\|\cdot\| := \langle\cdot,\cdot\rangle^{1/2}$ *eine* **Semi-Norm**, *denn aus* $\|f\| = 0$ *folgt nicht* $f = 0$.

Ein wichtiges Beispiel eines semi-definiten inneren Produkts Im linearen Raum $V = C[a,b]$ ist gegeben durch m verschiedene Punkte $t_i \in [a,b]$, $i = 1,2,\ldots,m$, und

$$\langle f,g\rangle := \frac{1}{m}\sum_{i=1}^{m} f(t_i)g(t_i), \quad f,g \in C[a,b].$$

Die zugehörige Semi-Norm ist

$$\|f\| := \Big(\frac{1}{m}\sum_{i=1}^{m} f(t_i)^2\Big)^{1/2}.$$

In einem rechtwinkligen Kooordinatensystem des Raumes mit den Einheits-Achsen $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ stehen die Ortsvektoren

$$\mathbf{x} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_3\mathbf{e}_3 \quad , \quad \mathbf{y} = b_1\mathbf{e}_1 + b_2\mathbf{e}_2 + b_3\mathbf{e}_3$$

senkrecht aufeinander, wenn

$$a_1b_1 + a_2b_2 + a_3b_3 = 0$$

ist. Allgemein definieren wir

Definition 6.4. *Es sei V ein linearer Raum über $\mathbb{R}$ mit dem inneren Produkt* $\langle\cdot,\cdot\rangle$.

1. *Zwei Vektoren $f,g \in V$ heißen* **orthogonal**, *falls $\langle f,g\rangle = 0$ erfüllt ist.*
2. *Eine Basis $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ von V heißt* **Orthonormalbasis** *von V, falls*

$$\langle v_j, v_k\rangle = \delta_{j,k} := \begin{cases} 1, & \text{für } j = k, \\ 0, & \text{für } j \ne k \end{cases}$$

(Kroneckersymbol δ). *$\mathcal{B}$ heißt* **orthogonal**, *falls*

$$\langle v_j, v_k \rangle = 0 \quad \textit{für alle } j \neq k.$$

Beispiel 1:
Es sei $n \in \mathbb{N}$. Es sei $V = \mathbb{R}^n$ versehen mit dem inneren Produkt (6.1). Dann bilden die **Einheitsvektoren**

$$\begin{pmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \quad \cdots, \begin{pmatrix} 0 \\ 0 \\ \cdot \\ 0 \\ 1 \end{pmatrix}$$

eine Orthonormalbasis.

Beispiel 2: Vergleiche §21 in Kapitel II des WS 2001/02 über **Fourierreihen**!
Es sei $n \in \mathbb{N}$. Es sei $V = \mathcal{T}_n$ der lineare Raum aller reellwertigen trigonometrischen Polynome vom Grade $\leq n$. Bezüglich des inneren Produkts

$$\langle f, g \rangle := \int_0^{2\pi} f(t)g(t)dt, \quad f, g \in V,$$

bilden die $2n + 1$ trigonometrischen Polynome

$$1, \cos t, \sin t, \cos(2t), \sin(2t), \ldots, \cos(nt), \sin(nt) \tag{6.3}$$

eine **orthogonale Basis** von $\mathcal{T}_n$. Diese Basis ist nicht orthonormal, denn es gilt

$$\int_0^{2\pi} 1^2 dt = 2\pi, \quad \int_0^{2\pi} \cos(kt)^2 dt = \pi, \quad \int_0^{2\pi} \sin(kt)^2 dt = \pi, \quad k \in \mathbb{N}.$$

Eine Verallgemeinerung des Satz 21.2 in Kapitel II des WS 2001/02 über Fourierreihen ist der wichtige **Projektionssatz**, den wir nun ohne Beweis formulieren wollen:

Satz 6.2.
Es sei V ein linearer Raum über $\mathbb{R}$ mit dem **semidefiniten inneren** *Produkt $\langle \cdot, \cdot \rangle$ und der zugehörige Semi-Norm $\|\cdot\| = \langle \cdot, \cdot \rangle|^{1/2}$. Es sei $U \subset V$ ein linearer Unterraum von V, und die Einschränkung von $\langle \cdot, \cdot \rangle$ auf $U \times U$ sei ein inneres Produkt auf U. Sei $\mathcal{B}_U = \{u_1, u_2, \ldots, u_n\}$ eine Basis von U.*
Dann hat zu gegebenem $f \in V$ das Minimumproblem

$$\min \left\{ \|f - \sum_{k=1}^{n} \alpha_k u_k\| : \quad \alpha_k \in \mathbb{R} \right\} \tag{6.4}$$

eine eindeutige Lösung $u_f = \sum_{k=1}^n c_k u_k$. *Diese ist charakterisiert durch die Orthogonalitätsbedingungen*

$$\langle f - u_f, u\rangle = 0 \quad \text{für alle } u \in U. \tag{6.5}$$

Die Orthogonalitätsbedingungen (6.5) sind äquivalent zu

$$\langle f - u_f, u_j\rangle = 0, \quad j = 1, 2, \ldots, n, \tag{6.6}$$

und somit zu dem linearen Gleichungssystem

$$\sum_{k=1}^n c_k \langle u_j, u_k\rangle = \langle f, u_j\rangle, \quad j = 1, 2, \ldots, n. \tag{6.7}$$

Anwendung : Die Regressionsgerade (Ausgleichsgerade)

Aufgabe:
Zu den Zeiten $t_1 < t_2 < \cdots < t_m$ seien die reellen Zahlen $y_1, \ldots, y_m$ gemessen worden. Man berechne zu den Zahlenpaaren $\{t_i, y_i\}_{i=1}^m$ die "beste Gerade", die sogenannte "Regressionsgerade". Darunter versteht man die Gerade $g(t)$, $t \in \mathbb{R}$, für die der quadratische Fehler

$$\sum_{i=1}^m (y_i - g(t_i))^2$$

so klein wie möglich ist.

Lösung:
Um diese Regressionsgerade auszurechnen, berechnen wir zunächst den Schwerpunkt der t-Werte,

$$s := \frac{1}{m}\sum_{i=1}^m t_i$$

und stellen die Gerade in der Form

$$g(t) = c_1 + c_2(t - s), \quad t \in \mathbb{R},$$

dar. Das heißt, wir wählen die Basis

$$u_1(t) = 1 \quad , \quad u_2(t) = t - s$$

des Unterraumes $U = \mathcal{P}_1$ der algebraischen Polynome vom Grad ≤ 1.
Der quadratische Fehler ist nun

$$Q(\alpha, \beta) := \sum_{i=1}^m (y_i - \alpha - \beta(t_i - s))^2.$$

Wir können nun sofort Satz 6.2 anwenden für den linearen Raum $V = C[t_1, t_m]$ und das semidefinite innere Produkt

$$\langle f, g\rangle = \sum_{i=1}^{m} f(t_i)g(t_i), \quad f, g \in V.$$

Sei $f \in C[t_1, t_m]$ irgendeine Funktion mit $f(t_i) = y_i$, $i = 1, \ldots, m$, so führen die Ortogonalitätsbedingungen (6.7) wegen

$$\langle u_1, u_1\rangle = m \quad , \quad \langle u_2, u_2\rangle = \sum_{i=1}^{m}(t_i - s)^2 \quad , \quad \langle u_1, u_2\rangle = \langle u_2, u_1\rangle = \sum_{i=1}^{m}(t_i - s) = 0$$

zu der Lösung

$$c_1 = \frac{1}{m}\langle f, u_1\rangle = \frac{1}{m}\sum_{i=1}^{m} y_i,$$

$$c_2 = \frac{\langle f, u_2\rangle}{\langle u_2, u_2\rangle} = \frac{\sum_{i=1}^{m} y_i(t_i - s)}{\sum_{i=1}^{m}(t_i - s)^2}.$$

Zum Abschluß dieses Paragraphen kehren wir zurück zu **Orthonormalbasen** eines linearen Raumes V, die in Definition 6.4 eingeführt wurden. Der nächste Satz zeigt, daß jeder **endlich dimensionale** lineare Raum V mit innerem Produkt eine Orthonormalbasis besitzt.

Satz 6.3. (Orthogonalisierungsverfahren von Gram - Schmidt)
Es sei V ein linearer Raum über $\mathbb{R}$ mit dem inneren Produkt $\langle\cdot,\cdot\rangle$ und der Basis $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$. Dann existiert eine Orthonormalbasis $\mathcal{B}^* = \{u_1, u_2, \ldots, u_n\}$ von V, so daß

$$\mathit{span}\{u_1, \ldots, u_k\} = \mathit{span}\{v_1, \ldots, v_k\} \quad \text{für } k = 1, 2, \ldots, n$$

erfüllt ist.

Beweis: Wir setzen

$$u_1 := \frac{v_1}{\|v_1\|}$$

und definieren für $k = 2, 3, \ldots, n$ die Vektoren w_k und u_k aus $\mathrm{span}\{v_1, \ldots, v_k\}$ rekursiv durch

$$w_k := v_k - \sum_{j=1}^{k-1}\langle v_k, u_j\rangle u_j,$$

$$u_k := \frac{w_k}{\|w_k\|},$$

und prüfen nach, daß die so definierten Vektoren $u_1, \ldots, u_n$ die gewünschten Eigenschaften des Satzes besitzen. ■

Bemerkung: Der Einfachheit wegen haben wir in diesem Paragraphen die Diskussion auf lineare Räume über $\mathbb{R}$ beschränkt. Bei linearen Räumen über $\mathbb{C}$ müssen wir die Definition 6.1 des inneren Produkts wie folgt abändern:

Definition 6.5. *Es sei V ein linearer Raum über $\mathbb{C}$. Eine Abbildung $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{C}$ heißt* **inneres Produkt** *in V, falls*

(L) *(Linearität in der 1.Komponente von $\langle\cdot,\cdot\rangle$)*
$\langle f_1 + f_2, g\rangle = \langle f_1, g\rangle + \langle f_2, g\rangle$ *für alle* $f_1, f_2, g \in V$,
$\langle \alpha f, g\rangle = \alpha\langle f, g\rangle$ *für alle* $f, g \in V$ *und* $\alpha \in \mathbb{C}$,

(S) *(hermitesch)*
$\langle f, g\rangle = \overline{\langle g, f\rangle}$ *für alle* $f, g \in V$,

(P) *(Positivität)*
$\langle f, f\rangle \geq 0$ *für alle* $f \in V$,

(P*) $\langle f, f\rangle = 0$ *genau dann, wenn* $f = 0$.

Sei zum Bespiel n eine natürliche Zahl und V die Menge aller komplexwertigen trigonometrischen Polynome vom Grade $\leq n$, dann bilden für das innere Produkt

$$\langle f, g\rangle := \int_0^{2\pi} f(t)\overline{g(t)}dt$$

die trigonometrischen Polynome

$$\left\{e^{ikt} : \quad -n \leq k \leq n\right\}$$

eine **orthogonale Basis** von V.

§7. Matrizenkalkül

Zur Lösung von linearen Gleichungssystemen, von linearen **Eigenwertproblemen** (siehe §8), von linearen Optimierungsproblemen (durch die Simplexmethode), und so weiter, ist die Darstellung dieser Probleme in Matrixform und das **Rechnen mit Matrizen** sehr bequem. Wir wollen in diesem und auch in dem nächsten Paragraphen **Matrizen mit reellen oder komplexen Zahlen** $a_{j,k}, b_{j,k}, c_{j,k}$ betrachten:

Es seien A und B $m \times n$-Matrizen, also Tableaus der Form

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & \cdots & a_{2,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m,1} & a_{m,2} & \cdots & \cdots & a_{m,n} \end{pmatrix} \quad , \quad B = \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & \cdots & b_{1,n} \\ b_{2,1} & b_{2,2} & \cdots & \cdots & b_{2,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ b_{m,1} & b_{m,2} & \cdots & \cdots & b_{m,n} \end{pmatrix}.$$

Dann definieren wir die **Addition** durch

$$A + B := \begin{pmatrix} a_{1,1} + b_{1,1} & a_{1,2} + b_{1,2} & \cdots & \cdots & a_{1,n} + b_{1,n} \\ a_{2,1} + b_{2,1} & a_{2,2} + b_{2,2} & \cdots & \cdots & a_{2,n} + b_{2,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m,1} + b_{m,1} & a_{m,2} + b_{m,2} & \cdots & \cdots & a_{m,n} + b_{m,n} \end{pmatrix},$$

also durch **komponentenweise** Addition der Elemente der Matrizen.

Für $\alpha \in \mathbb{C}$ definieren wir die **skalare Multiplikation**

$$\alpha A := \begin{pmatrix} \alpha a_{1,1} & \alpha a_{1,2} & \cdots & \cdots & \alpha a_{1,n} \\ \alpha a_{2,1} & \alpha a_{2,2} & \cdots & \cdots & \alpha a_{2,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha a_{m,1} & \alpha a_{m,2} & \cdots & \cdots & \alpha a_{m,n} \end{pmatrix},$$

also durch **komponentenweise** Multiplikation der Elemente der Matrizen mit α.

Sie können sicher sofort zeigen, daß die Menge aller reellen $m \times n$-Matrizen mit dieser Addition und skalaren Multiplikation einen linearen Raum über $\mathbb{R}$ der Dimension mn bilden.
Ebenso bilden die komplexen $m \times n$-Matrizen mit dieser Addition und skalaren Multiplikation einen linearen Raum über $\mathbb{C}$ der Dimension mn.

Nur etwas komplizierter ist das **Matrizenprodukt** zweier Matrizen A und B.
Hierbei ist ganz wichtig:
wir können eine $m \times n$-Matrix A mit einer $\ell \times q$-Matrix B nur dann multiplizieren, wenn

$$n = \ell$$

ist, also wenn die Spaltenanzahl von A mit der Zeilenanzahl von B übereinstimmt. Seien also m, n, q natürliche Zahlen, und seien

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & \cdots & a_{2,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m,1} & a_{m,2} & \cdots & \cdots & a_{m,n} \end{pmatrix} \quad , \quad B = \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & \cdots & b_{1,q} \\ b_{2,1} & b_{2,2} & \cdots & \cdots & b_{2,q} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ b_{n,1} & b_{n,2} & \cdots & \cdots & b_{n,q} \end{pmatrix}.$$

Dann definieren wir das **Matrizenprodukt** AB als die $m \times q$-Matrix

$$C = AB = \begin{pmatrix} c_{1,1} & c_{1,2} & \cdots & \cdots & c_{1,q} \\ c_{2,1} & c_{2,2} & \cdots & \cdots & c_{2,q} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{m,1} & c_{m,2} & \cdots & \cdots & c_{m,q} \end{pmatrix}$$

mit

$$c_{i,k} := \sum_{j=1}^{n} a_{i,j} b_{j,k}, \quad i = 1, \dots, m, \;\; k = 1, \dots, q.$$

Beispiel: Rechnen Sie nach, daß

$$\begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 4 & 1 & 2 \\ 3 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 7 & 1 & 3 \\ 8 & 2 & 4 \end{pmatrix}.$$

Als Spezialfall der Matrizenmultiplikation sei A eine $n \times n$-Matrix und

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ \cdots \\ x_n \end{pmatrix} \quad , \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ \cdots \\ b_n \end{pmatrix}$$

Spaltenvektoren des $\mathbb{C}^n$ ($= n \times 1$-Matrizen), so können wir das lineare Gleichungssystem

$$\sum_{k=1}^{n} a_{j,k} x_k = b_j, \quad j = 1, \ldots, n,$$

auch kurz schreiben als

$$Ax = b,$$

wobei Ax das Matrizenprodukt zwischen A und x bedeutet.

Eine andere wichtige Anwendung des Matrizenprodukts ist die **Hintereinanderausführung** zweier linearer Abbildungen

$$B : \mathbb{R}^q \to \mathbb{R}^n \quad , \quad A : \mathbb{R}^n \to \mathbb{R}^m.$$

Denn bilden wir zunächst $x \in \mathbb{R}^q$ ab auf $y = Bx \in \mathbb{R}^n$ und danach y auf $z = Ay \in \mathbb{R}^m$, dann führt dies zu der linearen Abbildung $x \to z$, die durch die Matrix $C = AB$ beschrieben wird :

$$z = A(Bx) = (AB)x.$$

Denn es gilt, wie man (zum Teil mühsam) nachrechnen kann,

Satz 7.1.
Ist A eine $m \times n$-Matrix, B eine $n \times q$ Matrix und C eine $q \times \ell$-Matrix, so gilt

$$(AB)C = A(BC).$$

Daher sagen wir: Die Matrizenmultiplikation ist **assoziativ** *und schreiben anstelle von $(AB)C$ oder $A(BC)$ oft nur ABC.*

Satz 7.2.
Ist A eine $m \times n$-Matrix und B eine $n \times q$ Matrix, dann gilt für die **transponierten** *Matrizen*

$$(AB)^T = B^T A^T.$$

Satz 7.3.
Sind A_1 und A_2 $m \times n$-Matrizen, B_1 und B_2 $n \times q$ Matrizen, so gilt

$$(A_1 + A_2)(B_1 + B_2) = A_1 B_1 + A_1 B_2 + A_2 B_1 + A_2 B_2,$$

man kann also **kreuzweise ausmultiplizieren**.

Sehr nützlich ist auch der folgende

Satz 7.4.
Es seien A und B zwei $n \times n$-Matrizen. Dann gilt

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

§8. Das Eigenwertproblem

Wir beginnen mit einem Beispiel aus der Physik:
Ein gekoppeltes System von Schwingungen (von Federn, Stromkreis) kann man beschreiben durch ein **homogenes System von Differentialgleichungen 1.Ordnung mit konstanten Koeffizienten**:

$$\begin{array}{l} y_1'(t) = a_{1,1}y_1(t) + a_{1,2}y_2(t) + \cdots + a_{1,n}y_n(t) \\ y_2'(t) = a_{2,1}y_1(t) + a_{2,2}y_2(t) + \cdots + a_{2,n}y_n(t) \\ \quad \cdots\cdots\cdots\cdots\cdots\cdots \\ \quad \cdots\cdots\cdots\cdots\cdots\cdots \\ y_n'(t) = a_{n,1}y_1(t) + a_{n,2}y_2(t) + \cdots + a_{n,n}y_n(t) \end{array} . \tag{8.1}$$

Sei $A = (a_{j,k})_{j,k=1}^n$ die zugehörige $n \times n$-Matrix und

$$\mathbf{y(t)} := \begin{pmatrix} y_1(t) \\ y_2(t) \\ \cdot \\ \cdot \\ y_n(t) \end{pmatrix}, \quad \textbf{mit} \;\; \mathbf{t} \in \mathbb{R}$$

die zugehörigen Spaltenvektoren des $\mathbb{C}^n$ ($= n \times 1$-Matrizen), so können wir mit Hilfe des Matrizenprodukts das System (8.1) kurz durch

$$\mathbf{y}'(\mathbf{t}) = \mathbf{A}\mathbf{y}(\mathbf{t}) \tag{8.2}$$

beschreiben.

Der Ansatz $\mathbf{y(t)} = \mathbf{e^{\lambda t}x}$ mit $\lambda \in \mathbb{C}$ und zeitunabhängigem $\mathbf{x} \in \mathbf{C^n}$ führt wegen $e^{\lambda t} \neq 0$ zu

$$A\mathbf{x} = \lambda\mathbf{x}. \tag{8.3}$$

Definition 8.1.
Es heißt $\lambda \in C$ **Eigenwert** *der $n \times n$-Matrix A, falls ein $\mathbf{x} \in \mathbf{C^n}$, $\mathbf{x} \neq (0,0,\cdots,0)^T$, existiert, so daß (8.3) erfüllt ist. Dann heißt* $\mathbf{x}$ **Eigenvektor** *zur Matrix A und Eigenwert λ.*

Ist A eine reelle $n \times n$-Matrix und λ ein reeller Eigenwert von A, so gibt es stets einen reellen Eigenvektor $\mathbf{x} \in \mathbb{R}^\mathbf{n}$. Können Sie diese Aussage schnell beweisen ?

Die Bedingung (8.3) ist äquivalent zu

$$\begin{pmatrix} a_{1,1}-\lambda & a_{1,2} & \cdots & \cdots & \cdots & a_{1,n} \\ a_{2,1} & a_{1,2}-\lambda & \cdots & \cdots & \cdots & a_{2,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & a_{n,2} & \cdots & \cdots & a_{n,n-1} & a_{n,n}-\lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}$$

oder, kurz geschrieben,

$$(A-\lambda I)\mathbf{x} = \mathbf{0}$$

mit der $n \times n$-Einheitsmatrix

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

Daher folgt aus Satz 5.3 und Satz 4.2

Satz 8.1.

1. *Die Zahl $\lambda \in \mathbb{C}$ ist Eigenwert der Matrix A dann und nur dann, wenn $\det(A-\lambda I) = 0$ ist.*
2. *Die Zahl $\lambda \in \mathbb{C}$ ist Eigenwert der Matrix A dann und nur dann, wenn $\mathrm{Rang}(A-\lambda I) \leq n-1$ ist.*
3. *Die maximale Anzahl linear unabhängiger Eigenvektoren zum Eigenwert λ von A ist*

$$n - \mathrm{Rang}(A-\lambda I).$$

Wir erkennen sofort aus Satz 5.1, daß

$$\det(A-zI) = (-1)^n z^n + z^{n-1}(-1)^{n-1}\sum_{k=1}^{n} a_{k,k} + \cdots + \det(A), \quad z \in \mathbb{C}, \tag{8.4}$$

ein algebraisches Polynom vom Grad n ist. Es heißt **charakteristisches Polynom** der Matrix A. Aus Satz 8.1 lesen wir sofort die folgende Charakterisierung der Eigenwerte der $n \times n$-Matrix A ab:

Satz 8.2.
Die Eigenwerte der Matrix A sind gerade die Nullstellen des charakteristischen Polynoms von A.

Beispiel 1.
Es sei $n = 2$ und $A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. Dann ist

$$\det(A - \lambda I) = \det \begin{pmatrix} 1-\lambda & 1 \\ 1 & -1-\lambda \end{pmatrix} = \lambda^2 - 2.$$

Somit sind die Eigenwerte $\lambda_1 = \sqrt{2}$ und $\lambda_2 = -\sqrt{2}$. Die Eigenvektoren zu $\lambda_1 = \sqrt{2}$ sind

$$x = \alpha \begin{pmatrix} \sqrt{2}+1 \\ 1 \end{pmatrix}, \quad \alpha \in \mathbb{C},$$

die Eigenvektoren zu $\lambda_2 = -\sqrt{2}$ sind

$$x = \alpha \begin{pmatrix} -1 \\ 1+\sqrt{2} \end{pmatrix}, \quad \alpha \in \mathbb{C}.$$

Beispiel 2.
Es sei $n = 2$ und $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Dann ist

$$\det(A - \lambda I) = \det \begin{pmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{pmatrix} = (1-\lambda)^2.$$

Dann sind $\lambda_1 = \lambda_2 = 1$ die Eigenwerte, und

$$x = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \alpha \in \mathbb{C},$$

sind die zugehörigen Eigenvektoren.

In der Mathematik und ihren Anwendungen treten häufig symmetrische, orthogonale, hermitesche und unitäre $n \times n$- Matrizen auf.

Definition 8.1.

1. *Eine reelle $n \times n$-Matrix A heißt* **symmetrisch**, *falls $A^T = A$* *(A^T ist die transponierte Matrix von A).*
2. *Eine reelle $n \times n$-Matrix A heißt* **orthogonal**, *falls $AA^T = A^T A = I$*
3. *Eine komplexe $n \times n$-Matrix A heißt* **hermitesch**, *falls $A^H = A$ erfüllt ist. Hierbei definieren wir A^H und sagen "A hermitesch" durch*

$$A^H = \overline{A^T} = (\overline{A})^T = \begin{pmatrix} \overline{a_{1,1}} & \overline{a_{2,1}} & \cdots & \cdots & \overline{a_{n,1}} \\ \overline{a_{1,2}} & \overline{a_{2,2}} & \cdots & \cdots & \overline{a_{n,2}} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \overline{a_{1,n}} & \overline{a_{2,n}} & \cdots & \cdots & \overline{a_{n,n}} \end{pmatrix}.$$

4. *Eine komplexe $n \times n$-Matrix A heißt* **unitär**, *falls $AA^H = A^H A = I$ erfüllt ist.*

Bemerkung:
Ist A eine reelle $n \times n$-Matrix, so ist
A symmetrisch genau dann, wenn A hermitesch ist,
A orthogonal genau dann, wenn A unitär ist.

Über die Lage der Eigenwerte von symmetrischen, orthogonalen, hermiteschen und unitären $n \times n$- Matrizen kann man folgendes sagen:

Satz 8.3.
1. *Die Eigenwerte einer reell symmetrischen oder einer hermiteschen Matrix sind* **alle reell**.
2. *Die Eigenwerte λ einer reellen orthogonalen oder einer unitären Matrix haben* **alle** *den Betrag $|\lambda| = 1$.*

Beweis: Der Beweis ist nicht schwer und wird in den Übungen behandelt.

§9. Die inverse Matrix

Definition 9.1. *Es sei A eine $n \times n$-Matrix.*
1. *A heißt* **nichtsingulär**, *falls $det(A) \neq 0$ ist. Ist $det(A) = 0$, so heißt A* **singulär**.
2. *Eine $n \times n$-Matrix C heißt* **Inverse** *der Matrix A, wenn*

$$AC = I, \qquad I \;\; n \times n - \textit{Einheitsmatrix}.$$

Wir bezeichen die Inverse von A mit A^{-1}.

Aus Zeitgründen wollen wir einige Aussagen über inverse Matrizen auflisten, aber nicht beweisen. Die Beweise sollten Sie aber in der Literatur nachlesen.

Satz 9.1. *Sei A eine $n \times n$-Matrix.*
1. *Die Matrix A besitzt eine Inverse dann und nur dann, wenn A nichtsingulär ist.*
2. *Es gibt zu jeder nichtsingulären Matrix A eine eindeutige Inverse. Diese erfüllt die Gleichung*

$$A^{-1}A = I.$$

Für nichtsinguläre Matrizen gelten die folgenden Rechenregeln:

Satz 9.2. *Es seien A und B nichtsinguläre $n \times n$-Matrizen. Dann gilt*

$$\begin{aligned}
(\overline{A})^{-1} &= \overline{A^{-1}} \\
(A^{-1})^{-1} &= A \\
(AB)^{-1} &= B^{-1}A^{-1}\,. \\
(A^T)^{-1} &= (A^{-1})^T \\
(A^H)^{-1} &= (A^{-1})^H
\end{aligned}$$

Satz 9.3. *Ist λ ein Eigenwert einer nichtsingulären $n \times n$-Matrix A, so ist $\lambda \neq 0$, und $\frac{1}{\lambda}$ ist ein Eigenwert von A^{-1}.*

Satz 9.4. *Sei A eine reell symmetrische $n \times n$-Matrix. Sei*

$$det(A - zI) = (-1)^n \prod_{j=1}^{n} (z - \lambda_j)$$

das charakteristische Polynom von A.

1. *Dann gibt es zu jedem $j = 1, 2, \ldots, n$ einen Eigenvektor $x^{(j)} \in \mathbb{R}^n$ zum Eigenwert λ_j, so daß*
$$(x^{(j)})^T x^{(k)} = \delta_{j,k}, \quad j, k = 1, \ldots, n.$$
Mit anderen Worten: Es gibt eine Orthonormalbasis $x^{(1)}, \ldots, x^{(n)}$ des $\mathbb{R}^n$ bestehend nur aus Eigenvektoren der Matrix A.
2. *Sei*
$$C = \left(x^{(1)}, \ldots, x^{(n)}\right)$$
die $n \times n$-Matrix, deren Spalten aus den Vektoren $x^{(1)}, \ldots, x^{(n)}$ besteht. Dann gilt
$$C^T C = C C^T = I, \quad C \text{ ist orthogonal}$$
und
$$C^T A C = \begin{pmatrix} \lambda_1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda_2 & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix} =: diag(\lambda_1, \ldots, \lambda_n).$$

Definition 9.2. *Zwei $n \times n$-Matrizen heißen* **ähnlich**, *falls es eine nichtsinguläre $n \times n$-Matrix C gibt, so daß*

$$B = C^{-1} A C$$

erfüllt ist.

Der letzte Satz besagt also, daß eine reelle symmetrische Matrix A ähnlich ist zu der Diagonalmatrix $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ der Eigenwerte von A.
Für ähnliche Matrizen gilt der folgende

Satz 9.5. *Ähnliche Matrizen haben dasselbe charakteristische Polynom und somit dieselben Eigenwerte (inklusive deren Vielfachheiten).*

Zum Abschluß dieses Kapitels wollen wir eine interessante Anwendung des Satzes 9.4 über reell symmetrische Matrizen diskutieren, die

Hauptachsentransformation quadratischer Formen:

Gegeben sei eine reelle symmetrische $n \times n$-Matrix $A = (a_{j,k})_{j,k=1}^n$ und die zugehörige **quadratische Form** $\mathbf{Q} : \mathbb{R}^n \to \mathbb{R}$ definiert durch

$$\mathbf{Q}(x_1, x_2, \ldots, x_n) = \sum_{j,k=1}^{n} a_{j,k} x_j x_k = x^T A x.$$

Frage: Für welche **orthogonale Basis** des $\mathbb{R}^n$ (= **Richtungen der Hauptachsen**) wird die Darstellung von $\mathbf{Q}$ besonders einfach ?

Antwort: Sei $C = (x^{(1)}, x^{(2)}, \ldots, x^{(n)})$ eine orthogonale Basis des Satzes 9.4, so gilt für die neuen Koordinaten $y \in \mathbb{R}^n$, definiert durch

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ \cdots \\ y_n \end{pmatrix} := C^{-1} x = C^T x \quad , \quad x = Cy$$

und die durch diese Transformation entstehende quadratische Form $\mathbf{Q}^*$, daß

$$\mathbf{Q}^*(y) = x^T A x = y^T C^T A C y = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2.$$

Hierbei haben wir ausgenutzt, daß die Matrizenmultiplikation **assoziativ** ist.

Beispiel 3: $Q(x_1, x_2) = x_1^2 + 4x_1x_2 + 4x_2^2 = 1, \quad (x_1, x_2) \in \mathbb{R}^2.$
Schreibe dies in der Form

$$Q(x_1, x_2) = (x_1, x_2) \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 1, \quad (x_1, x_2) \in \mathbb{R}^2.$$

Beispiel 4: $Q(x_1, x_2) = 2x_1^2 + 4x_1x_2 - x_2^2 + 2x_1 - 4x_2 = 2, \quad (x_1, x_2) \in \mathbb{R}^2.$
Schreibe dies in der Form

$$Q(x_1, x_2) = (x_1, x_2) \begin{pmatrix} 2 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + (2, -4) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 2, \quad (x_1, x_2) \in \mathbb{R}^2.$$

Kapitel IV
Differential- und Integralrechnung von Funktionen in mehreren Variablen

§1. Reellwertige Funktionen in n reellen Variablen

In §5 des Kapitel II haben wir den Begriff des metrischen Raumes D mit Metrik $d : D \to \mathbb{R}$, sowie Cauchyfolgen und deren Grenzwerte in metrischen Räumen definiert. Wir können sofort nachprüfen, daß für $D \subseteq \mathbb{R}^n$ und jede Norm $\|\cdot\|$ des $\mathbb{R}^n$ durch

$$d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|, \quad \mathbf{x}, \mathbf{y} \in D$$

eine Metrik definiert wird. Vektoren des $\mathbb{R}^n$ werden ab jetzt **fett** gedruckt !!!
Besonders wichtige Normen des $\mathbb{R}^n$ sind gegeben durch

$$\|\mathbf{x}\|_1 := \sum_{k=1}^{n} |x_k| \quad , \quad \|\mathbf{x}\|_2 := \Big(\sum_{k=1}^{n} |x_k|^2\Big)^{1/2} \quad , \quad \|\mathbf{x}\|_\infty := \max_{1 \le k \le n} |x_k|,$$

wobei $\mathbf{x} = (x_1, \ldots, x_n)$.
Sie können sicher sofort beweisen, daß für alle $\mathbf{x} \in \mathbb{R}^n$ die folgenden Ungleichungen richtig sind:

$$\begin{aligned} \|\mathbf{x}\|_2 &\le \|\mathbf{x}\|_1 \le \sqrt{n}\|\mathbf{x}\|_2, \\ \|\mathbf{x}\|_\infty &\le \|\mathbf{x}\|_1 \le n\|\mathbf{x}\|_\infty, \\ \|\mathbf{x}\|_\infty &\le \|\mathbf{x}\|_2 \le \sqrt{n}\|\mathbf{x}\|_\infty. \end{aligned}$$

Man sagt : "Diese Normen sind äquivalent". Hieraus folgt auch sofort der

Satz 1.1. *Eine Folge*

$$\Big\{\mathbf{x}^m = (x_1^m, x_2^m, \ldots, x_n^m) : \quad m \in \mathbb{N}\Big\} \subset \mathbb{R}^n$$

konvergiert **komponentenweise** *gegen* $\mathbf{x} \in \mathbb{R}^n$*, das heißt*

$$\lim_{m \to \infty} x_k^m = x_k, \quad k = 1, \ldots, n,$$

genau dann, wenn die Folge $(\mathbf{x}^m)_{m=1}^\infty$ *in jeder der Normen* $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$ *gegen* $\mathbf{x}$ *konvergiert, das heißt, wenn* $\lim_{m\to\infty} \|\mathbf{x}^m - \mathbf{x}\|_p = 0$, $p = 1, 2, \infty$.

Definition 1.1.

1. *Es sei* $D \subseteq \mathbb{R}^n$*. Ein Punkt* $\mathbf{x}^0 \in D$ *heißt* **innerer Punkt** *von* D*, wenn es eine* ε*-Umgebung*
$$U_\varepsilon(\mathbf{x}^0) := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{x}^0\|_2 < \varepsilon\}$$
(mit $\varepsilon > 0$*) von* $\mathbf{x}^0$ *gibt, die vollständig in* D *liegt.*
2. $D \subseteq \mathbb{R}^n$ *heißt* **offen**, *falls jeder Punkt von* D *innerer Punkt von* D *ist.*
3. $D \subseteq \mathbb{R}^n$ *heißt* **abgeschlossen**, *wenn das Komplement* $\mathbb{R}^n \setminus D$ *eine offene Menge ist.*
4. $\mathbf{x}^0 \in \mathbb{R}^n$ *heißt* **Häufungspunkt** *von* D*, wenn es eine Folge* $(\mathbf{x}^m)_{m=1}^\infty \subset D \setminus \{\mathbf{x}^0\}$ *gibt, die gegen* $\mathbf{x}^0$ *konvergiert.*
5. $\mathbf{x}^0 \in \mathbb{R}^n$ *heißt* **Randpunkt** *von* D*, wenn jede* ε*-Umgebung* $U_\varepsilon(\mathbf{x}^0)$ *von* $\mathbf{x}^0$ *Punkte aus* D *und aus* $\mathbb{R}^n \setminus D$ *enthält.*

Nach dieser Definition ist die leere Menge $\emptyset$ und $D = \mathbb{R}^n$ offen und abgeschlossen.

Ähnlich wie in Definition 5.5 des Kapitel II definieren wir die Stetigkeit einer Funktion $f : D \to \mathbb{R}$.

Definition 1.2. *Sei $D \subseteq \mathbb{R}^n$. Eine Funktion $f : D \to \mathbb{R}$ heißt stetig in $\mathbf{x}^0 \in D$, wenn es zu jedem $\varepsilon > 0$ ein $\delta = \delta(\varepsilon, \mathbf{x}^0, f) > 0$ gibt, so daß*

$$|f(\mathbf{x}) - f(\mathbf{x}^0)| < \varepsilon \quad \text{für alle } \mathbf{x} \in D \text{ mit } \|\mathbf{x} - \mathbf{x}^0\|_2 < \delta.$$

Eine Funktion $f = \begin{pmatrix} f_1 \\ f_2 \\ \cdot \\ \cdot \\ f_N \end{pmatrix} : D \to \mathbb{R}^N$ heißt stetig in $\mathbf{x}^0 \in D$, wenn jede der Komponentenfunktionen $f_1, f_2, \dots, f_N$ stetig in $\mathbf{x}^0$ ist.

Ähnlich wie bei Funktionen einer Variablen beweist man die folgenden drei Sätze.

Satz 1.2. *Es sei $D \subseteq \mathbb{R}^n$. Folgende Aussagen sind äquivalent:*

1. *Eine Funktion $f : D \to \mathbb{R}$ ist stetig in $\mathbf{x}^0$.*
2. *Für jede Folge $(\mathbf{x}^m)_{m=1}^{\infty} \subset D$ mit $\lim_{m\to\infty} \mathbf{x}^m = \mathbf{x}^0$ gilt*

$$\lim_{m\to\infty} f(\mathbf{x}^m) = f(\mathbf{x}^0).$$

Satz 1.3. *Es sei $D \subseteq \mathbb{R}^n$ und $F : D \to \mathbb{R}^N$, sei $\tilde{D} \subseteq \mathbb{R}^N$ mit $\{F(\mathbf{x}) : \mathbf{x} \in D\} \subseteq \tilde{D}$ und $G : \tilde{D} \to \mathbb{R}^q$. Sei $f := G \circ F : D \to \mathbb{R}^q$, daß heißt*

$$f(\mathbf{x}) := G(F(\mathbf{x})), \quad \mathbf{x} \in D.$$

Ist F in $\mathbf{x}^0 \in D$ stetig und G in $F(\mathbf{x}^0)$ stetig, dann ist auch die zusammengesetzte Funktion $f := G \circ F$ in $\mathbf{x}^0$ stetig.

Satz 1.4. *Sind $f : D \to \mathbb{R}$ und $g : D \to \mathbb{R}$ stetig in $\mathbf{x}^0 \in D$, dann sind auch die folgenden Funktionen stetig in $\mathbf{x}^0$:*
$f + g$, αf mit $\alpha \in \mathbb{R}$, $|f|$, fg, sowie $\frac{f}{g}$ falls $g(\mathbf{x}^0) \neq 0$.

Wir wollen dieses Paragraphen mit einem Beispiel abschließen:
Die Funktion $f : D := \mathbb{R}^3 \setminus \{(0,0,0)\} \to \mathbb{R}$, gegeben durch

$$f(x_1, x_2, x_3) := \frac{1}{\sqrt{x_1^2 + x_2^2 + x_3^2}},$$

ist stetig in D, aber nicht **gleichmäßig stetig** in D.

§2. Partielle Ableitungen. Richtungsableitungen

In Kapitel II haben wir Funktionen $f : (a, b) \to \mathbb{R}$ einer reellen Variablen differenziert, und zwar ist für $x \in (a, b)$

$$f'(x) := \lim_{h \to 0, h \in \mathbb{R}} \frac{f(x+h) - f(x)}{h} \tag{2.1}$$

die 1.Ableitung von f in x, falls dieser Grenzwert existiert. Bei Funktionen $f : D \to \mathbb{R}$ mit $D \subseteq \mathbb{R}^n$ mit $n \geq 2$ Variablen definieren wir die **Richtungsableitungen**:

Definition 2.1. *Sei* $\mathbf{x} = (x_1, \ldots, x_n)$ *ein innerer Punkt von* $D \subseteq \mathbb{R}^n$ *und* $\mathbf{a} = (a_1, \ldots, a_n) \subset \mathbb{R}^n$ *eine* **Richtung**, *das heißt von der Länge* $\sqrt{a_1^2 + \cdots + a_n^2} = 1$. *Dann definieren wir die* **Richtungsableitung** *von* f *in* $\mathbf{x}$ *in Richtung* $\mathbf{a}$ *durch*

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{a}} := \lim_{h \to 0, h \in \mathbb{R}} \frac{f(\mathbf{x} + h\mathbf{a}) - f(\mathbf{x})}{h} \quad , \tag{2.2}$$

falls dieser Grenzwert existiert.

Ist die Richtung $\mathbf{a}$ ein Einheitsvektor $\mathbf{a} = \mathbf{e}_k = (0, \ldots, 0, 1, 0, \ldots, 0)$ mit einer 1 in der k-ten Komponente, dann schreiben wir

$$\frac{\partial f(\mathbf{x})}{\partial x_k} \text{ anstelle von } \frac{\partial f(\mathbf{x})}{\partial \mathbf{e}_k}$$

und nennen $\frac{\partial f(\mathbf{x})}{\partial x_k}$ die **partielle Ableitung** von f nach x_k an der Stelle $\mathbf{x}$. Mit anderen Worten,

$$\frac{\partial f(\mathbf{x})}{\partial x_k} = \lim_{h \to 0, h \in \mathbb{R}} \frac{f(x_1, \ldots, x_{k-1}, x_k + h, x_{k+1}, \ldots, x_n) - f(\mathbf{x})}{h}. \tag{2.3}$$

Partielle Ableitungen höherer Ordnung definieren wir rekursiv durch

$$\frac{\partial^m f(\mathbf{x})}{\partial x_{k_1} \cdots \partial x_{k_m}} := \frac{\partial}{\partial x_{k_m}} \left(\frac{\partial^{m-1} f(\mathbf{x})}{\partial x_{k_1} \cdots \partial x_{k_{m-1}}} \right).$$

Für eine Abbildung $f : D \to \mathbb{R}$ mit $D \subseteq \mathbb{R}^n$ schreiben wir $f \in C(D)$ oder auch $f \in C^0(D)$, falls f in jedem Punkt von D stetig ist.

Definition 2.2. *Sei* $m \in \mathbb{N}$. *Es sei* $D \subseteq \mathbb{R}^n$ *eine offene Menge und* $f \in C(D)$. *Wir schreiben* $f \in C^m(D)$, *falls alle partiellen Ableitungen von* f *der Ordnung* $\leq m$ *in jedem Punkt von* D *existieren und stetig sind.*

Natürlich ist $C^m(D) \subset C^{m-1}(D) \subset \cdots \subset C(D)$. Ist $f \in C^m(D)$, so brauchen wir uns keine Gedanken darüber machen, in welcher Reihenfolge wir differenzieren. Denn es gilt

Satz 2.1. *Sei $m \in \mathbb{N}$, $D \subseteq \mathbb{R}^n$ offen und $f \in C^m(D)$. Dann gilt für jedes $\mathbf{x} \in D$ und jede Permutation $\nu_1, \ldots, \nu_m$ der Zahlen $k_1, \ldots, k_m \in \{1, 2, \ldots, n\}$*

$$\frac{\partial^m f(\mathbf{x})}{\partial x_{\nu_1} \cdots \partial x_{\nu_m}} = \frac{\partial^m f(\mathbf{x})}{\partial x_{k_1} \cdots \partial x_{k_m}}.$$

Insbesondere gilt für jede offene Menge $D \subseteq \mathbb{R}^2$ und $f \in C^2(D)$, daß

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} = \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1}, \quad \mathbf{x} = (x_1, x_2) \in D.$$

Beispiel 1. $f(x_1, x_2) := \frac{1}{x_1^2 + x_2^2}$, $\mathbf{x} = (x_1, x_2) \neq (0, 0)$.
Dann ist

$$\begin{aligned}
\frac{\partial f(\mathbf{x})}{\partial x_1} &= \frac{-2x_1}{(x_1^2 + x_2^2)^2} \quad , \\
\frac{\partial f(\mathbf{x})}{\partial x_2} &= \frac{-2x_2}{(x_1^2 + x_2^2)^2} \quad , \\
\frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} &= \frac{8x_1 x_2}{(x_1^2 + x_2^2)^3} \quad , \\
\frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} &= \frac{8x_1 x_2}{(x_1^2 + x_2^2)^3} \quad .
\end{aligned}$$

Sehr wichtig und nützlich ist der folgende Satz über Richtungsableitungen, dessen Beweis wir aus Zeitgründen nicht führen wollen. Sie sollten den Beweis aber in einem Lehrbuch nachlesen !!

Satz 2.2. *Es sei $D \subseteq \mathbb{R}^n$ offen und $f \in C^1(D)$. Es sei $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ eine Richtung im $\mathbb{R}^n$. Dann ist*

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{a}} = \sum_{k=1}^{n} a_k \frac{\partial f(\mathbf{x})}{\partial x_k}, \quad \text{für alle } \mathbf{x} \in D.$$

Beispiel 2. Sei $f \subset C^\infty(D)$ die Funktion des Beispiels 1 und $\mathbf{a} := (1/2, -\sqrt{3}/2)$. Dann ist

$$\frac{\partial f(x_1, x_2)}{\partial \mathbf{a}} = \frac{-2a_1 x_1 - 2a_2 x_2}{(x_1^2 + x_2^2)^2} = \frac{-x_1 + \sqrt{3} x_2}{(x_1^2 + x_2^2)^2} \quad , \quad (x_1, x_2) \neq (0, 0).$$

§3. Wichtige Begriffe aus der Vektoranalysis

Beachte: Wir unterscheiden nicht zwischen Zeilen- und Spaltenvektoren des $\mathbb{R}^n$, es sei denn, wir fassen diese als $(1 \times n)-$ oder als $(n \times 1)$-Matrizen auf.
Sei zum Beispiel $D \subseteq \mathbb{R}^n$ offen und $f : D \to \mathbb{R}$ eine "skalare Funktion" mit $f \in C^1(D)$, so unterscheiden wir nicht die Vektoren des $\mathbb{R}^n$

$$\text{grad } f(\mathbf{x}) = \Big(\frac{\partial f(\mathbf{x})}{\partial x_1}, \frac{\partial f(\mathbf{x})}{\partial x_2}, \dots, \frac{\partial f(\mathbf{x})}{\partial x_n}\Big)$$

von

$$\text{grad } f(\mathbf{x}) = \begin{pmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \frac{\partial f(\mathbf{x})}{\partial x_2} \\ \cdots \\ \cdots \\ \cdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{pmatrix}, \qquad \text{(Gradient von } f\text{)}$$

es sei denn, wir schreiben

$$(\text{grad } f(\mathbf{x}))^T \text{grad } f(\mathbf{x}) = \|\text{grad } f(\mathbf{x})\|_2^2$$

und meinen damit das Quadrat der **Euklidischen Länge** des **Gradienten** grad $f(\mathbf{x})$ **der Funktion** f an der Stelle $\mathbf{x} \in D$. Unter Verwendung des inneren Produkts $\langle \cdot, \cdot \rangle$ des $\mathbb{R}^n$ schreiben wir aber auch

$$\langle \text{grad } f(\mathbf{x}) \ , \ \text{grad } f(\mathbf{x}) \rangle \ = \ \|\text{grad } f(\mathbf{x})\|_2^2 \ .$$

Satz 3.1. *Es ist im Falle grad* $f(\mathbf{x}) \neq (0, \dots, 0)$ *die Richtung*

$$\frac{\text{grad } f(\mathbf{x})}{\|\text{grad } f(\mathbf{x})\|_2} \tag{3.1}$$

die Richtung des **steilsten Anstiegs** *der Funktion* f *im Punkt* $\mathbf{x} \in D$.

Beweis:
Der Vektor in (3.1) ist eine Richtung, das heißt, von der Euklidischen Länge 1. Es sei $\mathbf{a} = (a_1, a_2, \dots, a_n) \in \mathbb{R}^n$ eine Richtung, das heißt, $\|\mathbf{a}\|_2 = \sqrt{a_1^2 + \cdots + a_n^2} = 1$. Dann ist nach Satz 2.2 die "Anstiegsgeschwindigkeit" in $\mathbf{x}$ in Richtung $\mathbf{a}$ gegeben durch

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{a}} = \sum_{k=1}^{n} a_k \frac{\partial f(\mathbf{x})}{\partial x_k}.$$

Wenden wir hierauf die Cauchy-Schwarzsche Ungleichung an, so folgt

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{a}} \le \Big(\sum_{k=1}^{n} a_k^2\Big)^{1/2} \left(\sum_{k=1}^{n} \Big(\frac{\partial f(\mathbf{x})}{\partial x_k}\Big)^2\right)^{1/2} = \left(\sum_{k=1}^{n} \Big(\frac{\partial f(\mathbf{x})}{\partial x_k}\Big)^2\right)^{1/2}$$

mit Gleichheit genau dann, wenn

$$\mathbf{a} = \frac{\text{grad } f(\mathbf{x})}{\|\text{grad } f(\mathbf{x})\|_2}$$

erfüllt ist. ■

Definition 3.1. *Sei $D \subseteq \mathbb{R}^n$. Eine Abbildung*

$$\mathbf{V} = \begin{pmatrix} V_1 \\ V_2 \\ \cdot \\ \cdot \\ V_n \end{pmatrix} : D \to \mathbb{R}^n \tag{3.2}$$

heißt **Vektorfeld** *auf D.*

Ist zum Beispiel $D \subseteq \mathbb{R}^n$ offen und $f \in C^1(D)$, dann ist durch den Gradienten grad $f : D \to \mathbb{R}^n$ ein **stetiges** Vektorfeld auf D gegeben.

Wir schreiben $\mathbf{V} \in C^1(D)$, falls jede Komponente V_k von $\mathbf{V}$ in $C^1(D)$ liegt.

Definition 3.2. *Sei $D \subseteq \mathbb{R}^n$ offen. Das Vektorfeld $\mathbf{V}$ liege in $C^1(D)$. Dann definieren wir dessen* **Divergenz** *div $\mathbf{V} : D \to \mathbb{R}$ durch*

$$\mathit{div}\ \mathbf{V}(\mathbf{x}) := \sum_{k=1}^{n} \frac{\partial V_k(\mathbf{x})}{\partial x_k}.$$

Liegt zum Beispiel die skalare Funktion f in $C^2(D)$, dann ist

$$\text{div}\ \Big(\text{grad}\ f(\mathbf{x})\Big) = \sum_{k=1}^{n} \frac{\partial}{\partial x_k}\Big(\frac{\partial f(\mathbf{x})}{\partial x_k}\Big) = \sum_{k=1}^{n} \frac{\partial^2 f(\mathbf{x})}{\partial x_k^2} := \Delta f(\mathbf{x}).$$

Als **Laplace-Operator** bezeichnen wir die lineare Abbildung $\Delta : C^2(D) \to C(D)$ gegeben durch

$$\Delta f(\mathbf{x}) := \sum_{k=1}^{n} \frac{\partial^2 f(\mathbf{x})}{\partial x_k^2}, \quad \mathbf{x} \in D.$$

Also haben wir die wichtige Aussage bewiesen, daß

$$\Delta = \text{div}\ \Big(\text{grad}\ \Big)$$

und schreiben auch

$$\Delta = \text{div grad}\ .$$

Beachte: Es sind grad : $C^1(D) \to \mathbb{R}^n$ und div ebenfalls **lineare Operatoren** (lineare Abbildungen).

Im Falle $n = 3$ definieren wir einen weiteren linearen Operator, die **Rotation** eines Vektorfeldes $\mathbf{V}$.

Definition 3.3. *Sei $D \subseteq \mathbb{R}^3$ offen. Das Vektorfeld $\mathbf{V} = \begin{pmatrix} V_1 \\ V_2 \\ V_3 \end{pmatrix}$ liege in $C^1(D)$. Dann ist die* **Rotation** *rot $\mathbf{V} : D \to \mathbb{R}^3$ definiert durch*

$$\mathit{rot}\ \mathbf{V} = \begin{pmatrix} \frac{\partial V_3}{\partial x_2} - \frac{\partial V_2}{\partial x_3} \\[2ex] \frac{\partial V_1}{\partial x_3} - \frac{\partial V_3}{\partial x_1} \\[2ex] \frac{\partial V_2}{\partial x_1} - \frac{\partial V_1}{\partial x_2} \end{pmatrix}. \tag{3.3}$$

Ebenfalls im $\mathbb{R}^3$ definieren wir ein wichtiges Produkt, das **Vektorprodukt** im $\mathbb{R}^3$:
Für $\mathbf{v} = (v_1, v_2, v_3) \in \mathbb{R}^3$ und $\mathbf{w} = (w_1, w_2, w_3) \in \mathbb{R}^3$ definieren wir

$$\mathbf{v} \times \mathbf{w} := \Big(v_2 w_3 - v_3 w_2 \ , \ v_3 w_1 - v_1 w_3 \ , \ v_1 w_2 - v_2 w_1 \Big) \ \in \ \mathbb{R}^3. \tag{3.4}$$

Das Vektorprodukt hat eine wichtige geometrische Bedeutung: Es ist $\mathbf{v} \times \mathbf{w}$ der Vektor des $\mathbb{R}^3$,

(a) dessen Euklidische Länge $\|\mathbf{v} \times \mathbf{w}\|_2$ gleich ist der Fläche des von $\mathbf{v}$ und $\mathbf{w}$ aufgespannten Parallelogramms und

(b) der senkrecht steht auf den Vektoren $\mathbf{v}$ und $\mathbf{w}$, wobei

$$\mathbf{v}, \ \mathbf{w}, \ \mathbf{v} \times \mathbf{w}$$

ein Rechtssystem bilden.

Aus der Definition des Vektorprodukts folgt sofort, daß

$$\mathbf{w} \times \mathbf{v} = -\mathbf{v} \times \mathbf{w},$$

aber auch bezüglich des inneren Produkts im $\mathbb{R}^3$, daß

$$\langle \mathbf{v}, \mathbf{v} \times \mathbf{w} \rangle \ = \ \langle \mathbf{w}, \mathbf{v} \times \mathbf{w} \rangle \ = \ 0.$$

Beliebt bei den Physikern ist die symbolische Schreibweise

$$\mathrm{grad} \times \mathbf{V}$$

für ein Vektorfeld $\mathbf{V} = \begin{pmatrix} V_1 \\ V_2 \\ V_3 \end{pmatrix}$ des $C^1(D)$ mit $D \subseteq \mathbb{R}^3$. Diese Schreibweise bedeutet, daß man in der Definition (3.4) des Vektorprodukts stets v_k durch $\frac{\partial}{dx_k}$, $k = 1, 2, 3$, ersetzen muß. Das Ergebnis dieser Schreibweise ergibt dann die bekannte Aussage

$$\mathrm{grad} \times \mathbf{V} = \mathrm{rot}\ \mathbf{V}.$$

Es gibt eine lange Liste von Beziehungen zwischen den linearen Operatoren

$$\text{grad, div, rot im } \mathbb{R}^3,$$

die Sie mehr oder weniger leicht nachprüfen können. Hier ist eine kleine Auswahl:
Es sei $D \subseteq \mathbb{R}^3$ offen und $f : D \to \mathbb{R}$, $\mathbf{V} : D \to \mathbb{R}^3$. dann gilt

$$\begin{aligned}
&\text{rot}\Big(f\ \mathbf{V}\Big) = f\ \text{rot}\ \mathbf{V} - \mathbf{V} \times \text{grad}\ f \quad ,\\
&\text{div}\Big(\text{rot}\ \mathbf{V}\Big) = 0 \quad ,\\
&\text{rot}\Big(\text{rot}\ \mathbf{V}\Big) = \text{grad}\Big(\text{div}\ \mathbf{V}\Big) - \begin{pmatrix} \Delta V_1 \\ \Delta V_2 \\ \Delta V_3 \end{pmatrix} \quad .
\end{aligned}$$

§4. Differenzierbarkeit im $\mathbb{R}^n$

Bisher haben wir für Funktionen f mit n reellen Variablen **partielle Ableitungen** und **Richtungsableitungen** eingeführt. Wenn eine Funktion $f : (a, b) \to \mathbb{R}$ einer Variablen in $x_0 \in (a, b)$ differenzierbar ist, kann man durch

$$g(t) := f(x_0) + (t - x_0)f'(x_0), \quad t \in \mathbb{R},$$

eine Tangente ("Berührende") finden mit der Eigenschaft, daß

$$\lim_{t \to x_0} \frac{f(t) - g(t)}{|t - x_0|} = 0.$$

Wir wollen nun die Frage diskutieren, wann für eine Funktion $f : D \to \mathbb{R}$ mit $D \subseteq \mathbb{R}^n$ und einen inneren Punkt $\mathbf{x}^0 \in D$ eine **Tangentialebene** an f in $\mathbf{x}^0 = (x_1^0, x_2^0, \ldots, x_n^0)$ existiert. Hierbei verstehen wir unter einer solchen Ebene eine Funktion

$$g(\mathbf{x}) = f(\mathbf{x}^0) + \sum_{k=1}^{n} c_k(x_k - x_k^0), \quad \mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n,$$

die die Eigenschaft

$$\lim_{\mathbf{x} \to \mathbf{x}^0} \frac{f(\mathbf{x}) - g(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}^0\|_2} = 0 \tag{4.1}$$

besitzt. Marschieren wir mit $\mathbf{x} \in D$ parallel zur x_k-Achse gegen $\mathbf{x}^0$, so erkennen wir sofort, daß (4.1) nur möglich ist, falls

$$c_k = \frac{\partial f(\mathbf{x}^0)}{\partial x_k}, \quad k = 1, \ldots, n,$$

existiert.

Definition 4.1. *Sei $D \subseteq I\!R^n$ und $\mathbf{x}^0$ ein innerer Punkt von D. Eine Funktion $f : D \to I\!R$ heißt* **differenzierbar** *in $\mathbf{x}^0$, falls*

$$\lim_{\mathbf{x}\to\mathbf{x}^0} \frac{f(\mathbf{x}) - f(\mathbf{x}^0) - \sum_{k=1}^n \frac{\partial f(\mathbf{x}^0)}{\partial x_k}(x_k - x_k^0)}{\|\mathbf{x}-\mathbf{x}^0\|_2} = 0 \tag{4.2}$$

erfüllt ist.

Es ist sofort zu sehen, daß eine in $\mathbf{x}^0$ differenzierbare Funktion f in $\mathbf{x}^0$ stetig sein muß. Aber reicht die Stetigkeit und die Existenz der partiellen Ableitungen $\frac{\partial f(\mathbf{x}^0)}{\partial x_k}$, $k = 1, \ldots, n$, für die Differenzierbarkeit aus? Hierzu betrachten wir die Funktion

$$f(x_1, x_2) := \begin{cases} \frac{x_1 x_2}{\sqrt{x_1^2 + x_2^2}}, & (x_1, x_2) \in I\!R^2,\ (x_1, x_2) \neq (0,0), \\ 0, & (x_1, x_2) = (0,0). \end{cases} \tag{4.3}$$

In Polarkoordinaten $x_1 = r\cos\phi$, $x_2 = r\sin\phi$ geschrieben, ist dies die Funktion

$$f(x_1, x_2) = r \sin\phi \cos\phi$$

mit den Eigenschaften $f(x_1, 0) = f(0, x_2) = 0$, $x_1 \in I\!R$, $x_2 \in I\!R$. Also ist

$$c_k = \frac{\partial f(\mathbf{0})}{\partial x_k} = 0, \quad k = 1, 2,$$

aber es gilt nicht (4.2) wegen $g(\mathbf{x}) = 0$ und

$$\lim_{\mathbf{x}\to\mathbf{x}^0} \frac{f(\mathbf{x}) - g(\mathbf{x})}{\|\mathbf{x}\|_2} = \lim_{\mathbf{x}\to\mathbf{x}^0} \frac{f(\mathbf{x})}{\|\mathbf{x}\|_2} = \sin\phi\cos\phi \neq 0 \quad \text{für} \quad \phi \neq 0 \pmod{\frac{\pi}{2}}.$$

Doch oft können wir den folgenden Satz anwenden, um die *Differenzierbarkeit* einer Funktion f in n Variablen nachzuweisen.

Satz 4.1. *Sei $D \subseteq I\!R^n$ und $\mathbf{x}^0$ ein innerer Punkt von D. Eine Funktion $f : D \to I\!R$ ist in $\mathbf{x}^0$ differenzierbar, falls es ein $\varepsilon > 0$ und eine ε-Umgebung $U_\varepsilon(\mathbf{x}^0)$ gibt, in der die partiellen Ableitungen*

$$\frac{\partial f(\mathbf{x})}{\partial x_k}, \quad k = 1, \ldots, n, \quad \mathbf{x} \in U_\varepsilon(\mathbf{x}^0),$$

existieren und stetig sind.

Als Beispiel betrachten wir die obige Funktion (4.3). Sie ist nach Satz 4.1 in jedem Punkt $\mathbf{x}^0 \neq (0,0)$ differenzierbar.

§5. Die Kettenregel

Es sei $F = \begin{pmatrix} F_1 \\ F_2 \\ \cdot \\ \cdot \\ F_N \end{pmatrix} : D \to \mathbb{R}^N$ mit $D \subseteq \mathbb{R}^n$. Ist $\mathbf{x} \in D$ ein innerer Punkt, dann heißt die $N \times n$-Matrix

$$J(\mathbf{x}) = \begin{pmatrix} \frac{\partial F_1(\mathbf{x})}{\partial x_1} & \frac{\partial F_1(\mathbf{x})}{\partial x_2} & \dots & \dots & \frac{\partial F_1(\mathbf{x})}{\partial x_n} \\ \frac{\partial F_2(\mathbf{x})}{\partial x_1} & \frac{\partial F_2(\mathbf{x})}{\partial x_2} & \dots & \dots & \frac{\partial F_2(\mathbf{x})}{\partial x_n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \frac{\partial F_N(\mathbf{x})}{\partial x_1} & \frac{\partial F_N(\mathbf{x})}{\partial x_2} & \dots & \dots & \frac{\partial F_N(\mathbf{x})}{\partial x_n} \end{pmatrix}$$

die **Jacobische Matrix** von F an der Stelle $\mathbf{x}$, falls die obigen partiellen Ableitungen existieren. Anstelle von $J(\mathbf{x})$ schreibt man auch

$$\frac{\partial(F_1, \dots, F_N)(\mathbf{x})}{\partial(x_1, \dots, x_n)} \quad \text{oder} \quad \frac{dF(\mathbf{x})}{d\mathbf{x}}.$$

Eine in der Praxis sehr wichtige Regel beim Differenzieren ist die **Kettenregel**:

Satz 5.1. *Es sei $D \subseteq \mathbb{R}^n$ und $F : D \to \mathbb{R}^N$, sei $\tilde{D} \subseteq \mathbb{R}^N$ mit $\{F(\mathbf{x}) : \mathbf{x} \in D\} \subseteq \tilde{D}$ und $G : \tilde{D} \to \mathbb{R}^q$. Sei $f := G \circ F : D \to \mathbb{R}^q$, daß heißt*

$$f(\mathbf{x}) := G(F(\mathbf{x})), \quad \mathbf{x} \in D.$$

Ist F im inneren Punkt $\mathbf{x}^0 \in D$ differenzierbar und G im inneren Punkt $\mathbf{y}^0 := F(\mathbf{x}^0) \in \tilde{D}$ differenzierbar, dann ist auch $f := G \circ F$ in $\mathbf{x}^0$ differenzierbar, und es gilt

$$\frac{df(\mathbf{x^0})}{d\mathbf{x}} = \frac{dG(\mathbf{y}^0)}{d\mathbf{y}} \frac{dF(\mathbf{x}^0)}{d\mathbf{x}}, \tag{5.1}$$

wobei auf der rechten Seite in (5.1) das **Matrizenprodukt** *der $(q \times N)$-Matrix $\frac{dG(\mathbf{y}^0)}{d\mathbf{y}}$ mit der $(N \times n)$-Matrix $\frac{dF(\mathbf{x}^0)}{d\mathbf{x}}$ gebildet wird.*

Anstelle des Beweises, den Sie in einem Lehrbuch nachlesen sollten, wollen wir uns ein Beispiel ansehen: Es sei

$$f(x_1, x_2) = (x_1^2 + x_2) \sin(x_1 x_2), \quad (x_1, x_2) \in \mathbb{R}^2.$$

Wir können

$$\frac{df(\mathbf{x})}{d\mathbf{x}} = \text{grad } f(\mathbf{x}) = \Big(\frac{\partial f(\mathbf{x})}{\partial x_1} \;,\; \frac{\partial f(\mathbf{x})}{\partial x_2}\Big)$$

direkt berechnen, nämlich

$$\begin{aligned}\frac{\partial f(\mathbf{x})}{\partial x_1} &= 2x_1 \sin(x_1x_2) + (x_1^2+x_2)x_2\cos(x_1x_2)\\ \frac{\partial f(\mathbf{x})}{\partial x_2} &= \sin(x_1x_2) + (x_1^2+x_2)x_1\cos(x_1x_2)\end{aligned}.$$

Wollen wir hingegen die Kettenregel anwenden, so ist

$$F(x_1,x_2) = \begin{pmatrix} x_1^2+x_2 \\ x_1x_2 \end{pmatrix} \quad , \quad G(y_1,y_2) = y_1 \sin y_2$$

und $f = G \circ F$. Nach der Kettenregel ist dann

$$\frac{df(\mathbf{x})}{d\mathbf{x}} = (\sin y_2, y_1 \cos y_2) \begin{pmatrix} 2x_1 & 1 \\ x_2 & x_1 \end{pmatrix},$$

was natürlich zu demselben Ergebnis führt.

Wir können die Kettenregel auch verwenden, um die Jacobische Matrix einer Umkehrabbildung zu berechnen:

Satz 5.2. (**Umkehrsatz**)
Es sei

$$F = \begin{pmatrix} F_1 \\ F_2 \\ \cdot \\ \cdot \\ F_n \end{pmatrix} : D \to \mathbb{R}^n$$

ein Vektorfeld in $D \subseteq \mathbb{R}^n$. Sei D offen und $F \in C^1(D)$. Dann gibt es zu jedem $\mathbf{x}^0 \in D$ mit $\det \frac{dF(\mathbf{x}^0)}{d\mathbf{x}} \neq 0$ ein $\varepsilon > 0$ und genau eine Abbildung

$$F^{-1} : \tilde{U} \to \mathbb{R}^n \quad , \quad \tilde{U} := \{F(\mathbf{x}) : \mathbf{x} \in U_\varepsilon(\mathbf{x}^0)\},$$

mit der Eigenschaft

$$F^{-1}(F(\mathbf{x})) = \mathbf{x} \quad , \quad F(F^{-1}(\mathbf{y})) = \mathbf{y} \quad , \quad \mathbf{x} \in U_\varepsilon(\mathbf{x}^0) \quad , \quad \mathbf{y} \in \tilde{U}.$$

Des weiteren gilt dann

$$\frac{dF^{-1}(\mathbf{y})}{d\mathbf{y}} = \left(\frac{dF(\mathbf{x})}{d\mathbf{x}}\right)^{-1} \quad , \quad \mathbf{y} = F(\mathbf{x}) \quad , \quad \mathbf{x} \in U_\varepsilon(\mathbf{x}^0). \tag{5.2}$$

Beweis: Für die Identität $f(\mathbf{x}) = \mathbf{x}$ ist

$$\frac{df(\mathbf{x})}{d\mathbf{x}} = I \quad (n \times n - \text{Einheitsmatrix}).$$

Wir wenden die Kettenregel auf die rechte Seite der Gleichung

$$\mathbf{x} = F^{-1}(F(\mathbf{x}))$$

an und so erhalten aus Satz 5.1

$$I = \frac{dF^{-1}(\mathbf{y})}{d\mathbf{y}} \frac{dF(\mathbf{x})}{d\mathbf{x}},$$

woraus sofort (5.2) folgt.

§6. Das Riemannsche Integral für Intervalle des $\mathbb{R}^n$

Zu $\mathbf{a} = (a_1, \ldots, a_n)$, $\mathbf{b} = (b_1, \ldots, b_n) \in \mathbb{R}^n$ mit $a_k < b_k$, $k = 1, \ldots, n$, nennen wir die Menge

$$I = [\mathbf{a}, \mathbf{b}] = \{x = (x_1, \ldots, x_n) \in \mathbb{R}^n,\ \ a_k \le x_k \le b_k,\ \ k = 1, \ldots, n\}$$

ein **abgeschlossenes Intervall**. Das **Maß** von I ist

$$\mu(I) = \prod_{k=1}^{n} (b_k - a_k).$$

Eine Menge $\Delta = \{I_1, \ldots, I_m\}$ von nicht überlappenden Intervallen heißt **Partition** von I, falls $I = \cup\{I_j : j = 1, \ldots, m\}$.
Die maximale Kantenlänge eines Intervalls I_j bezeichnen wir mit $\delta(I_j)$. Dann ist

$$\|\Delta\| := \max_{1 \le j \le m} \delta(I_j)$$

ein Maß für die **Feinheit** der Partition Δ.
Wählen wir Punkte $\xi_j \in I_j$, $j = 1, \ldots, m$, so wird zu einer Funktion $f : I \to \mathbb{R}$ durch

$$J(\xi, \Delta) := \sum_{j=1}^{m} f(\xi_j)\mu(I_j)$$

die zu $\xi := (\xi_1, \ldots, \xi_n)$ und Δ gehörende **Riemannsche Zwischensumme** definiert.

Definition 6.1. *Eine Funktion $f : I \to \mathbb{R}$ heißt* **integrierbar** *auf dem Intervall $I = [\mathbf{a}, \mathbf{b}]$, falls für jede Folge $(\xi^m, \Delta^m)_{m=1}^{\infty}$ mit $\lim_{m\to\infty} \|\Delta^m\| = 0$ der Grenzwert*

$$\lim_{m\to\infty} J(\xi^m, \Delta^m)$$

existiert und derselbe ist. Dieser Grenzwert heißt dann **(Riemannsches) Integral** *von f auf I, in Zeichen*

$$\int_I f \quad \textit{oder} \quad \int_I f(x)dx \quad \textit{oder} \quad \int_I f(x_1, \ldots, x_n) d(x_1, \ldots, x_n).$$

Ohne Beweis geben wir den wichtigen

Satz 6.1. *Ist $f : I \to \mathbb{R}$ stetig im Intervall $I = [\mathbf{a}, \mathbf{b}]$, dann ist f auf I integrierbar.*

§7. Iterierte Integrale

Die Berechnung von Integralen über Intervallen $I = [\mathbf{a}, \mathbf{b}]$ des $\mathbb{R}^n$ kann häufig auf die Berechnung von n Integralen einer Variablen zurückgeführt werden.
Im Falle $n = 2$ wollen wir, um Sie zu verwirren, anstelle von (x_1, x_2) einfacher (x, y) schreiben. Das Intervall I des $\mathbb{R}^2$ wird nun zu

$$I = \{(x, y) \in \mathbb{R}^2 : \quad a \leq x \leq b, \;\; c \leq y \leq d\} = [a, b] \times [c, d].$$

Wie wird im nächsten Satz das Intervall I des $\mathbb{R}^2$ "aufgeschnitten"?

Satz 7.1. *Es sei $f : I \to \mathbb{R}$ auf dem Intervall I des $\mathbb{R}^2$ integrierbar.*

1. *Falls für jedes $x \in [a, b]$ das Integral $\int_c^d f(x, y)dy$ existiert, so existiert auch das* **iterierte Integral**

$$\int_a^b \Big(\int_c^d f(x, y)dy \Big) dx, \qquad \textit{(auch } \textbf{Mehrfach-Integral} \textit{ genannt)}$$

und es gilt

$$\int_I f(x, y)d(x, y) = \int_a^b \Big(\int_c^d f(x, y)dy \Big) dx. \tag{7.1}$$

2. Falls für jedes $y \in [c, d]$ das Integral $\int_a^b f(x, y)dx$ existiert, so existiert auch das **iterierte Integral**

$$\int_c^d \Big(\int_a^b f(x, y)dx \Big) dy,$$

und es gilt

$$\int_I f(x, y)d(x, y) = \int_c^d \Big(\int_a^b f(x, y)dx \Big) dy. \tag{7.2}$$

Korollar 7.1.

1. *Ist $f \in C(I)$, dann gelten (7.1) und (7.2).*
2. *Ist f von der Form $f(x, y) = g(x)h(y)$ mit $g \in C[a, b]$, $h \in C[c, d]$, so gilt*

$$\int_I f(x, y)d(x, y)) = \int_a^b g(x)dx \int_c^d h(y)dy.$$

Aber man kann einfache Beispiele von Funktionen f konstruieren, wo die iterierten Integrale $\int_c^d f(x, y)dy$ für alle $x \in [a, b]$ existieren, aber f nicht integrierbar ist, **und umgekehrt**.

Dieses **iterierte Integrieren** klappt auch für integrierbare Funktionen $f : I \to \mathbb{R}$ auf Intervallen des $\mathbb{R}^n$ mit $n \geq 3$. Wir wollen auf die Formulierung für allgemeines n verzichten und nur den Fall $n = 3$ genauer beschreiben :

Satz 7.2. *Es sei $I = [a_1, b_1]\times[a_2, b_2]\times[a_3, b_3]$ ein Intervall des $\mathbb{R}^3$ und $I_1 := [a_2, b_2]\times[a_3, b_3]$ das Intervall des $\mathbb{R}^2$. Es sei $f : I \to \mathbb{R}$ auf I integrierbar. Dann gilt:*

1. *Es ist*
$$\int_I f(x, y, z)d(x, y, z) = \int_{a_1}^{b_1} \Big(\int_{I_1} f(x, y, z)d(y, z)\Big)dx, \tag{7.3}$$
falls alle Integrale in (7.3) existieren;
2. *Es ist*
$$\int_I f(x, y, z)d(x, y, z) = \int_{I_1} \Big(\int_{a_1}^{b_1} f(x, y, z)dx\Big)d(y, z), \tag{7.4}$$
falls alle Integrale in (7.4) existieren.

Nun ist es aber höchste Zeit für einige **Beispiele**:

1. Sei $I = [0, 2] \times [0, 2]$. Dann ist

$$\begin{aligned}\int_I y \sin(\pi xy)d(x, y) &= \int_0^2 \Big(\int_0^2 y \sin(\pi xy)dx\Big)dy \\ &= \int_0^2 \Big[\frac{-y}{\pi y} \cos(\pi xy)\Big]_{x=0}^{x=2} dy \\ &= \frac{1}{\pi} \int_0^2 \Big(1 - \cos(2\pi y)\Big)dy = \frac{2}{\pi}.\end{aligned}$$

2. Sei $I = [0, 1]^3$. Dann ist

$$\begin{aligned}\int_I (x^3 + y^2 + z)d(x, y, z) &= \int_0^1 \Big(\int_{[0,1]^2} (x^3 + y^2 + z)d(y, z)\Big)dx \\ &= \int_0^1 \Big(\int_0^1 \Big(\int_0^1 (x^3 + y^2 + z)dz\Big)dy\Big)dx \\ &= \int_0^1 \Big(\int_0^1 \Big(x^3 + y^2 + \frac{1}{2}\Big)dy\Big)dx \\ &= \int_0^1 \Big(x^3 + \frac{1}{3} + \frac{1}{2}\Big)dx = \frac{1}{4} + \frac{1}{3} + \frac{1}{2}.\end{aligned}$$

§8. Das Riemannsche Integral auf beschränkten Mengen

Gegeben sei nun ein beschränkte Menge $D \subset \mathbb{R}^n$ und eine beschränkte Funktion $f : D \to \mathbb{R}$. Um das (Riemannsche) Integral auf beliebigen beschränkten Menge $D \subset \mathbb{R}^n$ einzuführen, bezeichnen wir

1. mit I_D das kleinste abgeschlossene Intervall des $\mathbb{R}^n$, in dem D enthalten ist und definieren
2. die Funktion $f_D : I_D \to \mathbb{R}$ durch

$$f_D(x) := \begin{cases} f(x), & x \in D, \\ 0, & x \in I_D \setminus D \end{cases} .$$

Definition 8.1.
Es heißt $f : D \to \mathbb{R}$ **Riemann-integrierbar** *auf D, falls f_D Riemann-integrierbar auf I_D ist, und wir schreiben dann*

$$\int_D f(\mathbf{x})d\mathbf{x} := \int_{I_D} f_D(\mathbf{x})d\mathbf{x}.$$

Im Wintersemester haben wir eine ganze Reihe von Rechenregeln für Integrale von Funktionen einer Variablen kennengelernt. Diese gelten meist auch für Integrale von Funktionen mehrerer Variablen und sollen im nächsten Satz aufgelistet werden.

Satz 8.1. *Es sei $D \subset \mathbb{R}^n$ beschränkt, und $f : D \to \mathbb{R}$, $g : D \to \mathbb{R}$ beschränkt und integrierbar auf D. Dann sind die folgenden Funktionen ebenfalls auf D integrierbar, und es gilt:*

$$\begin{aligned}
&\int_D \alpha f(\mathbf{x})d\mathbf{x} = \alpha \int_D f(\mathbf{x})d\mathbf{x}, \quad \alpha \in \mathbb{R}, \\
&\int_D (f(\mathbf{x}) + g(\mathbf{x}))d\mathbf{x} = \int_D f(\mathbf{x})d\mathbf{x} + \int_D g(\mathbf{x})d\mathbf{x}, \\
&\left| \int_D f(\mathbf{x})d\mathbf{x} \right| \le \int_D |f(\mathbf{x})|d\mathbf{x}, \\
&\left| \int_D f(\mathbf{x})g(\mathbf{x})d\mathbf{x} \right| \le \left(\int_D |f(\mathbf{x})|^2 d\mathbf{x} \right)^{1/2} \left(\int_D |g(\mathbf{x})|^2 d\mathbf{x} \right)^{1/2}, \quad \textit{(Cauchy-Schwarz)}.
\end{aligned}$$

Satz 8.2. *Sei $D = D_1 \cup D_2$, wobei D_1 und D_2 sich nicht überlappen. Ist f auf D beschränkt und auf D_1 und D_2 integrierbar, dann ist f auch auf D integrierbar, und es gilt*

$$\int_D f(\mathbf{x})d\mathbf{x} = \int_{D_1} f(\mathbf{x})d\mathbf{x} + \int_{D_2} f(\mathbf{x})d\mathbf{x}.$$

Anwendung:
Berechnung des Volumens und des Schwerpunkts eines Kreiskegels $D \subset \mathbb{R}^3$ mit Grundradius 1 und Höhe h,

$$D := \left\{ (x,y,z) \in \mathbb{R}^3 : \ x^2 + y^2 \le \left(\frac{h-z}{h} \right)^2 \right\}.$$

Dann ist $I_D = [-1,1]^2 \times [0,h]$, und das Volumen V ist

$$V = \int_0^h \Big(\int_{[-1,1]^2} f_D(x,y,z) d(x,y) \Big) dz$$

mit $f_D(x,y,z) = 1$ für $(x,y,z) \in D$, $f_D(x,y,z) = 0$ für $(x,y,z) \notin D$. Also ist

$$V = \int_D d(x,y,z) = \int_0^h \pi \Big(\frac{h-z}{h} \Big)^2 dz = \frac{\pi h}{3}.$$

Um den Schwerpunkt (x_S, y_S, z_S) zu berechnen, sehen wir sofort, daß $x_S = y_S = 0$ ist. Des weiteren ist

$$z_S = \frac{1}{V} \int_D z \ \ d(x,y,z) = \frac{3}{\pi h} \int_0^h z\pi \Big(\frac{h-z}{h} \Big)^2 dz = \frac{h}{4}.$$

Aufgabe:
Derselbe Kreiskegel $D \subset \mathbb{R}^3$ mit Grundradius 1 und Höhe h habe nun die Massendichte

$$\rho(x,y,z) := h - z, \quad (x,y,z) \in D.$$

Berechne dessen Schwerpunkt (x_S, y_S, z_S) !

Die Masse des Kreiskegels ist $M = \int_D \rho(x,y,z) d(x,y,z)$. Es ist $x_S = y_S = 0$, sowie

$$z_S = \frac{1}{M} \int_D z \ \ \rho(x,y,z) \ \ d(x,y,z) = \frac{1}{M} \int_0^h z(h-z)\pi \Big(\frac{h-z}{h} \Big)^2 dz.$$

§9. Die Substitutionsregel

Sie erinnern sich sicher an die Substitutionsregel bei der Integration von Funktionen einer Variablen. Wir wollen nun untersuchen, wie diese Regel zur Integration von Funktionen mehrerer Variablen zu formulieren ist. Zunächst legen wir fest, welche Eigenschaften die zur Substitution verwendeten Funktionen haben dürfen:

Definition 9.1. *Es seien D und $\tilde{D}$ offene Mengen des $\mathbb{R}^n$. Wir sagen, auf D werden durch die Funktion*

$$g = \begin{pmatrix} g_1 \\ g_2 \\ \cdot \\ \cdot \\ g_n \end{pmatrix} : \tilde{D} \to \mathbb{R}^n$$

neue Koordinaten *eingeführt, falls*

(1) g *bildet* $\tilde{D}$ *eineindeutig auf* D *ab,*

(2) $g \in C^1(\tilde{D})$,

(3) $\det\left(\frac{dg(\tilde{\mathbf{x}})}{d\tilde{\mathbf{x}}}\right) \neq 0$ *für alle* $\tilde{\mathbf{x}} \in \tilde{D}$.

Beispiel 1: Polarkoordinaten der Ebene

Es ist

$$\tilde{\mathbf{x}} = \begin{pmatrix} r \\ \phi \end{pmatrix} \quad , \quad g(\tilde{\mathbf{x}}) = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\cos\phi \\ r\sin\phi \end{pmatrix}.$$

$$\tilde{D} = \{(r,\phi): \;\; 0 < r < \infty, \;\; 0 < \phi < 2\pi\} \quad , \quad D = \{(x,y) \in \mathbb{R}^2\} \setminus \{(x,0): \;\; 0 \le x < \infty\}.$$

Und die Determinante wird

$$\det\left(\frac{dg(\tilde{\mathbf{x}})}{d\tilde{\mathbf{x}}}\right) = \det\begin{pmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{pmatrix} = r. \tag{9.1}$$

Beispiel 2: Sphärische Polarkoordinaten im $\mathbb{R}^3$

Es ist

$$\tilde{\mathbf{x}} = \begin{pmatrix} r \\ \phi \\ \theta \end{pmatrix} \quad , \quad g(\tilde{\mathbf{x}}) = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r\cos\phi\cos\theta \\ r\sin\phi\cos\theta \\ r\sin\theta \end{pmatrix}.$$

$$\begin{aligned} \tilde{D} &= \{(r,\phi,\theta): \;\; 0 < r < \infty, \;\; 0 < \phi < 2\pi, \;\; -\frac{\pi}{2} < \theta < \frac{\pi}{2}\} \\ D &= \{(x,y,z) \in \mathbb{R}^3\} \setminus \{(x,y,z): \;\; x \ge 0, \;\; y = 0, \;\; z \in \mathbb{R}\} \end{aligned}.$$

Und die Determinante ist nach einer Übungsaufgabe

$$\det\left(\frac{dg(\tilde{\mathbf{x}})}{d\tilde{\mathbf{x}}}\right) = r^2\cos\theta. \tag{9.2}$$

Beispiel 3: Zylinderkoordinaten

Es ist

$$\tilde{\mathbf{x}} = \begin{pmatrix} r \\ \phi \\ \tilde{z} \end{pmatrix} \quad , \quad g(\tilde{\mathbf{x}}) = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r\cos\phi \\ r\sin\phi \\ \tilde{z} \end{pmatrix}.$$

$$\begin{aligned}\tilde{D} &= \{(r,\phi,\tilde{z}):\ \ 0<r<\infty,\ \ 0<\phi<2\pi,\ \ \tilde{z}\in\mathbb{R}\} \\ D &= \{(x,y,z)\in\mathbb{R}^3\}\setminus\{(x,y,z):\ \ x\geq 0,\ \ y=0,\ \ z\in\mathbb{R}\}\end{aligned}.$$

Und die Determinante ist

$$\det\left(\frac{dg(\tilde{\mathbf{x}})}{d\tilde{\mathbf{x}}}\right) = \det\begin{pmatrix}\cos\phi & -r\sin\phi & 0\\ \sin\phi & r\cos\phi & 0\\ 0 & 0 & 1\end{pmatrix} = r. \tag{9.3}$$

Wir kennen bereits die geometrische Bedeutung der Determinante:
In einem rechtwinkligen Koordinatensystem $(\mathbf{e}_1, \mathbf{e}_2)$ seien die Vektoren

$$a_1\mathbf{e}_1 + a_2\mathbf{e}_2 \quad , \quad b_1\mathbf{e}_1 + b_2\mathbf{e}_2$$

gegeben. Sie spannen ein Parallelogramm der Fläche

$$\left|\det\begin{pmatrix}a_1 & b_1\\ a_2 & b_2\end{pmatrix}\right|$$

auf.
In einem rechtwinkligen Koordinatensystem $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ seien drei Vektoren

$$a_{1k}\mathbf{e}_1 + a_{2k}\mathbf{e}_2 + a_{3k}\mathbf{e}_3 \quad , \quad k = 1,2,3,$$

gegeben. Sie spannen einen schiefen Würfel des Volumens

$$\left|\det\begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix}\right|$$

auf.
Ist $\tilde{I} = [a_1,b_1]\times[a_2,b_2]\times\cdots\times[a_n,b_n]$ ein kleines Intervall des $\mathbb{R}^n$, und bildet die Funktion g das Intervall $\tilde{I}$ eineindeutig auf die Menge $B\subset\mathbb{R}^n$ ab, so gilt näherungsweise für jedes $\tilde{\mathbf{x}}\in\tilde{I}$, daß

$$\text{Volumen}(B) = \left|\det\Big(\frac{dg(\tilde{\mathbf{x}})}{d\tilde{\mathbf{x}}}\Big)\right| \text{Volumen}(\tilde{I}).$$

Dies macht die folgende Regel plausibel.

Satz 9.1. (Substitutionsregel.)
Es sei $K\subset\mathbb{R}^n$ eine abgeschlossene, beschränkte, meßbare (siehe §10) Menge. Auf einer offenen Menge $D\subset\mathbb{R}^n$ mit $K\subset D$ seien durch die Funktion $g:\tilde{D}\to\mathbb{R}^n$ neue Koordinaten eingeführt. Ist $f\in C(D)$, so gilt

$$\int_K f(\mathbf{x})d\mathbf{x} = \int_{g^{-1}(K)} f(g(\tilde{\mathbf{x}}))\left|\det\Big(\frac{dg(\tilde{\mathbf{x}})}{d\tilde{\mathbf{x}}}\Big)\right| d\tilde{\mathbf{x}}.$$

Dabei ist

$$g^{-1}(K) := \{\tilde{\mathbf{x}} \in \mathbb{R}^n : \; g(\tilde{\mathbf{x}}) \in K\} \quad \subset \tilde{D}$$

das Urbild von K bezüglich g.

Beispiel 4: Das Volumen der Kugel vom Radius R
Wir erhalten 1/8 des Volumens der Kugel vom Radius R, wenn wir das Volumen des Kugeloktanten

$$K = \{(x,y,z) \in \mathbb{R}^3 : \; x \geq 0, \; y \geq 0, \; z \geq 0, \; 0 \leq x^2 + y^2 + z^2 \leq R^2\}$$

berechnen. Um den Voraussetzungen des Satzes 9.1 voll Genüge zu tun, müßten wir zunächst für kleines $\varepsilon > 0$ das Volumen von

$$K_\varepsilon = \{(x,y,z) \in \mathbb{R}^3 : \; x \geq 0, \; y \geq 0, \; z \geq 0, \; \varepsilon^2 \leq x^2 + y^2 + z^2 \leq R^2\}$$

ausrechnen und danach $\varepsilon \to 0$ gehen lassen. Dies haben wir im Kopf, aber verfahren etwas legerer:
Unter Verwendung der sphärischen Polarkoordinaten des Beispiels 2 erhalten wir

$$\begin{aligned}
\text{Volumen}(K) &= \int_K 1 \cdot d(x,y,z) = \int_{g^{-1}(K)} r^2 |\cos\theta| \; d(r,\phi,\theta) \\
&= \int_0^R \Big(\int_0^{\pi/2} \Big(\int_0^{\pi/2} r^2 \cos\theta \; d\theta \Big) d\phi \Big) dr \\
&= \int_0^R \Big(\int_0^{\pi/2} r^2 d\phi \Big) dr = \frac{\pi R^3}{6}.
\end{aligned}$$

Also ist das Kugelvolumen gleich $V(R) = \frac{4\pi R^3}{3}$.

Können Sie mit Hilfe des Kugelvolumens die Oberfläche einer Kugel vom Radius R ausrechnen? Klar! Begünden Sie, warum die Oberfläche gegeben ist durch

$$\frac{dV(R)}{dR} = 4\pi R^2.$$

Beispiel 5:
Es sei $f \in C[1,2]$ und K der Kreisring $K = \{(x,y) \in \mathbb{R}^2 : \; 1 \leq x^2 + y^2 \leq 4\}$. Man berechne das Integral

$$\int_K f(\sqrt{x^2+y^2}) d(x,y) \quad !$$

Wir benutzen die Polarkoordinaten des Beispiels 1 : Dann ist

$$\int_K f(\sqrt{x^2+y^2}) d(x,y) = \int_1^2 \Big(\int_0^{2\pi} f(r) \; r \; d\phi \Big) dr = 2\pi \int_1^2 f(r) \; r \; dr.$$

Ist zum Beispiel $f(r) = r$, so ist das Integral gleich

$$2\pi \int_1^2 r^2 \ dr = \frac{2\pi(2^3 - 1)}{3} = \frac{14\pi}{3}.$$

Beachten Sie : Um die Voraussetzungen des Satzes 9.1 exakt einzuhalten, dürften wir zunächst nur über den Winkel $\varepsilon \le \phi \le 2\pi$ integrieren und müßten danach ε gegen 0 gehen lassen. Aber wir sind doch nicht so kleinlich, oder ?

§10. Das Maß von Teilmengen des $\mathbb{R}^n$

Uns allen ist aus §6 das Maß von Intervallen $I = [\mathbf{a}, \mathbf{b}]$ des $\mathbb{R}^n$ bekannt. Es ist

$$\mu(I) = \prod_{k=1}^{n} (b_k - a_k).$$

Für allgemeine beschränkte Mengen $D \subset \mathbb{R}^n$ definieren wir

Definition 10.1.
Es sei $D \subset \mathbb{R}^n$ eine beschränkte Menge und I_D das kleinste Intervall mit $D \subseteq I_D \subset \mathbb{R}^n$. Mit $\Delta = \{I_1, I_2, \ldots, I_m\}$ bezeichnen wir die Partitionen von I_D.
Es heißt D **Riemann meßbar**, *wenn*

$$\mu(D) := \sup_{\Delta} \sum_{I_k \subset D} \mu(I_k) = \inf_{\Delta} \sum_{I_k \cap D \neq \emptyset} \mu(I_k),$$

das heißt, wenn das **innere Maß** *von D gleich dem* **äußeren Maß** *von D ist. Dieser Wert heißt dann Riemannsches* **Maß** *von D.*

Aus dieser Definition der Meßbarkeit folgt der

Satz 10.1.
Es sei $D \subset \mathbb{R}^n$ beschränkt. Es ist D Riemann meßbar genau dann, wenn die charakteristische Funktion $c_D : I_D \to \mathbb{R}$,

$$c_D(\mathbf{x}) := \begin{cases} 1, & \mathbf{x} \in D \\ 0, & \mathbf{x} \notin D \end{cases}$$

von D Riemann integrierbar ist, und es ist

$$\mu(D) = \int_D 1 \ d\mathbf{x} = \int_{I_D} c_D(\mathbf{x}) d\mathbf{x}.$$

Kapitel V
Vektoranalysis. Integralsätze

§1. Kurven im $\mathbb{R}^n$

Eine Kurve im $\mathbb{R}^2$ kann dargestellt sein durch eine Gleichung, zum Beispiel beschreibt die Gleichung
$$\frac{(x-x_0)^2}{a^2}+\frac{(y-y_0)^2}{b^2}=1$$
eine Ellipse mit Mittelpunkt (x_0, y_0) und den Halbachsen $a > 0$ und $b > 0$.
In **Parameterdarstellung** wird dieselbe Ellipse beschrieben durch
$$\mathbf{f}(t) = (x(t), y(t)) = (a\cos t, b\sin t), \quad 0 \le t \le 2\pi,$$
und hat in dieser Darstellung eine **Durchlaufrichtung** mit **Anfangspunkt** $\mathbf{f}(0)$ und **Endpunkt** $\mathbf{f}(2\pi)$.

Eine elliptische Schraubenlinie im $\mathbb{R}^3$ um die Achse $\{(x_0, y_0, z) \in \mathbb{R}^3 : \ z \in \mathbb{R}\}$ mit Anstiegsgeschwindigkeit $c > 0$ und 5 Windungen hat in Parameterdarstellung die Form
$$\mathbf{f}(t) = (x(t), y(t), z(t)) = (x_0 + a\cos t,\ \ y_0 + b\sin t,\ \ ct), \quad 0 \le t \le 10\pi.$$
Ihr Anfangspunkt ist $\mathbf{f}(0)$, ihr Endpunkt $\mathbf{f}(10\pi)$.

Definition 1.1. *Sei $[a, b] \subset \mathbb{R}$ ein endliches reelles Intervall. Sei*
$$\mathbf{f} = (f_1, f_2, \ldots, f_n) : [a, b] \to \mathbb{R}^n$$
stetig auf $[a, b]$. Die Punktmenge
$$K := \{\mathbf{f}(t) \in \mathbb{R}^n : \ \ a \le t \le b\}$$
heißt **Kurve** *im $\mathbb{R}^n$, mit Parameterdarstellung $(\mathbf{f}, [a, b])$.*
Ist $\mathbf{f}(a) = \mathbf{f}(b)$. so heißt die Kurve **geschlossen**. *Sind alle Punkte $\mathbf{f}(t)$, $a \le t < b$, verschieden, so heißt sie* **Jordankurve**.

Die Ellipse ist also eine geschlossene Jordankurve, die Schraubenlinie ist ebenfalls eine Jordankurve, aber nicht geschlossen.

Existieren in $t \in [a,b]$ die Ableitungen $f_k'(t)$, $k = 1, \ldots, n$, und ist die Euklidische Norm

$$\|\mathbf{f}'(t)\|_2 := \left(\sum_{k=1}^{n} f_k'(t)^2 \right)^{1/2}$$

von Null verschieden, so nennen wir den Vektor

$$\mathbf{T}_f(t) := \frac{1}{\|\mathbf{f}'(t)\|_2} \mathbf{f}'(t)$$

der (Euklidischen) Länge $\|\mathbf{T}_f(t)\|_2 = 1$ den **Tangentenvektor** der Kurve K im Punkt $\mathbf{f}(t)$ und die Gerade

$$T := \{\mathbf{g}(\tau) := \mathbf{f}(t) + (\tau - t)\mathbf{f}'(t), \quad \tau \in \mathbb{R}\}$$

die **Tangente** an die Kurve K im Punkt $\mathbf{f}(t)$.
Hierbei bedeuten $\mathbf{f}'(a)$ die rechtsseitigen und $\mathbf{f}'(b)$ die linksseitigen Ableitungen.
Die Kurve K heißt **glatt**, falls $\mathbf{T}_f(t)$ für jedes $t \in [a,b]$ existiert und auf $[a,b]$ eine stetige Funktion ist.
Die Kurve K heißt **stückweise glatt**, falls es eine Partition $a = t_0 < t_1 < \cdots < t_m = b$ des Intervalls $[a,b]$ gibt, so daß die m Teilkurven in jedem Intervall $[t_{k-1}, t_k]$, $k = 1, \ldots, m$, glatt sind.

Beachte: Es kann von der Parameterdarstellung abhängen, ob ein Punktmenge als glatte oder als nichtglatte Kurve dargestellt wird. Zum Bespiel beschreiben die beiden Kurven

$$\left\{(t,\ t^2) \in \mathbb{R}^2 : \ -1 \le t \le 1\right\}$$

und

$$\left\{(t^3,\ t^6) \in \mathbb{R}^2 : \ -1 \le t \le 1\right\}$$

dieselbe Parabel, aber in der ersten Darstellung ist die Kurve glatt, in der zweiten Darstellung existiert der Tangentenvektor im Punkt $(0,0)$ nicht wegen $f_1'(0) = f_2'(0) = 0$.

§2. Die Bogenlänge von Kurven im $\mathbb{R}^n$

Die Kurve $K \subset \mathbb{R}^n$ sei dargestellt durch $(\mathbf{f}, [a,b])$, mit stetiger Funktion $\mathbf{f}$. Zu jeder Partition $\Delta : a = t_0 < t_1 < \cdots < t_m = b$ von $[a,b]$ ist die Länge des **Polygonzugs** durch die $m+1$ Punkte $\mathbf{f}(t_0), \mathbf{f}(t_1), \ldots, \mathbf{f}(t_m)$ der Kurve durch den Ausdruck

$$L_\Delta(K) = \sum_{j=1}^{m} \left(\sum_{k=1}^{n} \Big(f_k(t_j) - f_k(t_{j-1}) \Big)^2 \right)^{1/2} \tag{2.1}$$

gegeben. Die Kurve K heißt **rektifizierbar**, falls

$$L(K) := \sup_\Delta L_\Delta(K)$$

endlich ist, und $L(K)$ heißt die **Bogenlänge** der Kurve K.

Satz 2.1. *Ist K eine stückweise glatte Kurve, dargestellt durch $(\mathbf{f}, [a,b])$, so ist K rektifizierbar, und*

$$L(K) = \int_a^b \sqrt{\sum_{k=1}^{n} (f_k'(t))^2} \; dt. \tag{2.2}$$

Beweis: Die Formel (2.2) folgt sofort aus der Formel (2.1) und dem Mittelwertsatz der Differentialrechnung. ■

Man kann nachprüfen, daß die so definierte Bogenlänge nicht von der Parameterdarstellung der Kurve K abhängt. Um dies nachzuvollziehen, muß man die Substitutionregel bei der Integration von Funktionen einer reellen Variablen anwenden.

Beispiel 1: Die Bogenlänge des Graphen einer Funktion
Es sei $g \in C^1[a,b]$ und

$$\mathbf{f}(t) := (t, \;\; g(t)), \quad a \le t \le b,$$

der zu g gehörende **Graph**. Dann ist die Bogenlänge dieses Graphen nach Satz 2.1 zu berechnen nach der Formel

$$L(K) = \int_a^b \sqrt{1 + g'(t)^2} \; dt. \tag{2.3}$$

Beispiel 2: Die Bogenlänge einer Ellipse
Wie schon öfters erwähnt, hat die Ellipse mit Mittelpunkt $(0,0)$ und den Halbachsen $a > 0$, $b > 0$ die Parameterdarstellung

$$\mathbf{f}(t) = (a \cos t, \;\; b \sin t) \quad , \quad 0 \le t \le 2\pi.$$

Die Formel des Satzes 2.1 führt zu

$$L(K) = \int_0^{2\pi} \sqrt{a^2 \sin^2 t + b^2 \cos^2 t} \; dt = 4b \int_0^{\pi/2} \sqrt{1 - \alpha \sin^2 t} \; dt \tag{2.4}$$

mit $\alpha := (b^2 - a^2)/b^2$.
Für den Kreis ist $a = b$, also $\alpha = 0$ und somit $L(K) = 2\pi b$. Ist $a \ne b$, so sollten Sie nicht zu lange in der Formelsammlung nach einer Stammfunktion suchen. Sie werden eine solche nicht finden. Die Werte dieser Integrale sind tabelliert unter dem Stichwort **elliptische Integrale**. Im Zeitalter der schnellen Computer können Sie diese Integrale aber mit Hilfe der **Trapezregel** oder besser der **Simpsonregel** berechnen. Die Trapezregel haben wir bereits im Wintersemester kennengelernt. Die Simpsonregel zur näherungsweisen Berechnung des Integrals $\int_a^b f(t)dt$ hat die Form :

Wähle $n \in \mathbb{N}$ und berechne $h := (b-a)/(2n)$. Je größer n, um so näher liegt die folgende Zahl $S(h)$ bei $\int_a^b f(t)dt$:

$$\begin{aligned} S(h) &:= \frac{b-a}{6n}\Big(f(a) + f(b) + 4\sum_{k=1}^{n} f(a+(2k-1)h) + 2\sum_{k=1}^{n-1} f(a+2kh)\Big) \\ &= \frac{b-a}{6n}\Big(f(a) + 4f(a+h) + 2f(a+2h) + 4f(a+3h) + \cdots + 4f(b-h) + f(b)\Big). \end{aligned}$$

In **Pascal** könnte das Computer-Programm zur Berechnung des Integrals $\int_a^b f(t)dt = \int_0^{2\pi} \sqrt{1+\cos^2 t}\ \ dt$ durch die **Simpson-Regel** wie folgt aussehen:

```
program Simpson;
const a = 0;
      b = 2*pi;
      n = 1000;
var
      k       : integer;
      h, s, integral, fehler :     extended;
      ende     : char;
      function f(x:real):  real;
      begin
            f:= sqrt(1 + sqr(cos(x)));
      end;
begin
      h:=(b-a)/(2*n);
      s:= (f(a)+f(b))/2+2*f(b-h);
      for k:=1 to n-1 do
            s:=s+2*f(a+(2*k-1)*h)+ f(a+2*k*h);
      integral:=2*s*h/3;
      write('n=',n,' Integral=',integral:2:10);
      read(ende);
end.
```

§3. Kurvenintegrale im $\mathbb{R}^n$

In diesem Paragraphen ist K eine stückweise glatte Kurve des $\mathbb{R}^n$, dargestellt durch $(\mathbf{f}, [a,b])$. Es ist $\mathbf{V} = (V_1, V_2, \ldots, V_n) : K \to \mathbb{R}^n$ ein stetiges Vektorfeld. Wir nennen

$$\int_a^b \mathbf{V}(\mathbf{f}(t)) \cdot \mathbf{f}'(t)dt = \int_a^b \Big(\sum_{k=1}^{n} V_k(\mathbf{f}(t)) f_k'(t)\Big)dt \tag{3.1}$$

das **Kurvenintegral** von $\mathbf{V}$ längs K. Hierbei haben wir die in der Vektoranalysis gebräuchliche Schreibweise $\mathbf{x} \cdot \mathbf{y}$ anstelle von $\langle \mathbf{x}, \mathbf{y} \rangle$ verwendet.

Wegen der Substitutionsregel (in Kapitel II) ist der Wert (3.1) unabhängig von der Darstellung $(\mathbf{f}, [a, b])$ der Kurve K, das heißt, unabhängig davon, mit welcher "Geschwindigkeit" die Kurve K durchlaufen wird. Daher schreibt man anstelle von (3.1) auch

$$\int_K (V_1 dx_1 + V_2 dx_2 + \cdots + V_n dx_n). \tag{3.1*}$$

Wir können das Kurvenintegral auch physikalisch interpretieren:
Sei $\mathbf{V} = (V_1, V_2, V_3) : D \to \mathbb{R}^3$ ein Kraftfeld in einer offenen Menge $D \subseteq \mathbb{R}^3$. In D durchlaufe ein Körper der Masse 1 die Kurve K, wobei $\mathbf{f}(t) = (f_1(t), f_2(t), f_3(t))$ die Lage des Körpers zur Zeit t angebe, $a \leq t \leq b$. Dann wird dem Körper beim Durchlaufen der Kurve K die **Energie**

$$E = \int_a^b \mathbf{V}(\mathbf{f}(t)) \cdot \mathbf{f}'(t) dt$$

zugeführt.

§4. Potentiale

Sei $D \subseteq \mathbb{R}^n$ offen und $\mathbf{V} : D \to \mathbb{R}^n$ ein stetiges Vektorfeld. Wir nennen $\mathbf{V}$ ein **Gradientenfeld**, falls eine Funktion $\phi \in C^1(D)$ existiert, so daß

$$\text{grad } \phi(\mathbf{x}) = \mathbf{V}(\mathbf{x}), \quad \mathbf{x} \in D.$$

ϕ heißt dann **Potential** von $\mathbf{V}$.

Gradientenfelder haben folgende Eigenschaft.

Satz 4.1.
Es sei $D \subseteq \mathbb{R}^n$ offen und zusammenhängend. Sei $\mathbf{V} : D \to \mathbb{R}^n$ ein stetiges Vektorfeld. Dann sind folgende Aussagen äquivalent:

1. $\mathbf{V}$ *besitzt ein Potential* ϕ *in* D.
2. *Das Kurvenintegral bezüglich* $\mathbf{V}$ *ist wegunabhängig, das heißt: Zu beliebigen* $\mathbf{x}^0 \in D$, $\mathbf{x}^1 \in D$ *und beliebiger stückweiser glatten Jordankurve* $K := K(\mathbf{x}^0, \mathbf{x}^1)$, *die innerhalb* D *von* $\mathbf{x}^0$ *nach* $\mathbf{x}^1$ *verläuft, hat das Kurvenintegral*

   $$\int_K (V_1 dx_1 + V_2 dx_2 + \cdots + V_n dx_n) = \phi(\mathbf{x}^1) - \phi(\mathbf{x}^0)$$

   stets denselben Wert.

Beweis: Aus (1.) folgt (2.):

Sei $K := K(\mathbf{x}^0, \mathbf{x}^1)$ glatt und habe die Parameterdarstellung $(\mathbf{f}, [a,b])$. Dann ist

$$\int_K (V_1 dx_1 + V_2 dx_2 + \cdots + V_n dx_n) = \int_a^b \Big(\mathbf{V}(\mathbf{f}(t)) \cdot \mathbf{f}'(t)\Big) dt$$
$$= \int_a^b \Big(\sum_{k=1}^n V_k(\mathbf{f}(t)) f_k'(t)\Big) dt = \int_a^b \Big(\sum_{k=1}^n \frac{\partial \phi(\mathbf{f}(t))}{\partial x_k} f_k'(t)\Big) dt = \phi(\mathbf{f}(b)) - \phi(\mathbf{f}(a)),$$

wobei wir zum Schluß den Hauptsatz der Infinitesimalrechnung verwendet haben, denn es ist

$$\frac{\partial \phi(\mathbf{f}(t))}{\partial t} = \sum_{k=1}^n \frac{\partial \phi(\mathbf{f}(t))}{\partial x_k} f_k'(t).$$

Eine Menge $D \subset \mathbb{R}^n$ heißt **einfach zusammenhängend**, wenn sie keine Löcher hat. Zum Beispiel sind die Mengen

$$\mathbb{R}^n \setminus \{\mathbf{0}\} \quad , \quad \{\mathbf{x} \in \mathbb{R}^n,\ 1 < \|\mathbf{x}\|_2 < 2\}$$

zusammenhängend, aber **nicht** einfach zusammenhängend.
Für Vektorfelder im $\mathbb{R}^3$ gilt der folgende interessante

Satz 4.2.
Sei $D \subseteq \mathbb{R}^3$ offen und **einfach** *zusammenhängend. Es sei $\mathbf{V} = (V_1, V_2, V_3) : D \to \mathbb{R}^3$ ein Vektorfeld, sowie $\mathbf{V} \in C^1(D)$. Die drei folgenden Aussagen sind äquivalent:*

1. *$\mathbf{V}$ besitzt ein Potential ϕ.*
2. *Es ist*

$$\mathit{rot}\ \mathbf{V}(\mathbf{x}) \;=\; \mathbf{0} \quad \text{für alle } \mathbf{x} \in D.$$

   (wirbelfreies Vektorfeld $\mathbf{V}$).
3. *Es gelten die* **Integrabilitätsbedingungen** *in D,*

$$\frac{\partial V_3}{\partial x_2} = \frac{\partial V_2}{\partial x_3} \quad , \quad \frac{\partial V_1}{\partial x_3} = \frac{\partial V_3}{\partial x_1} \quad , \quad \frac{\partial V_2}{\partial x_1} = \frac{\partial V_1}{\partial x_2}.$$

§5. Flächen im $\mathbb{R}^3$

Es sei $S \subset \mathbb{R}^2$ offen, $\mathbf{f}(s,t) : S \to \mathbb{R}^3$, $\mathbf{f} = (f_1, f_2, f_3)$ mit $\mathbf{f} \in C^1(S)$ und

$$\text{Rang}\begin{pmatrix} \frac{\partial f_1(s,t)}{\partial s} & \frac{\partial f_1(s,t)}{\partial t} \\ \frac{\partial f_2(s,t)}{\partial s} & \frac{\partial f_2(s,t)}{\partial t} \\ \frac{\partial f_3(s,t)}{\partial s} & \frac{\partial f_3(s,t)}{\partial t} \end{pmatrix} = 2 \quad \text{für alle } (s,t) \in S.$$

Dann heißt die Punktmenge

$$F := \{(x,y,z) \subset \mathbb{R}^3 : \ x = f_1(s,t),\ y = f_2(s,t),\ z = f_3(s,t), \quad (s,t) \in S\}$$

eine **glatte Fläche** im $\mathbb{R}^3$ mit Parameterdarstellung $(\mathbf{f}, S)$.
Der Vektor

$$\mathbf{N}(s,t) := \frac{\frac{\partial \mathbf{f}(s,t)}{\partial s} \times \frac{\partial \mathbf{f}(s,t)}{\partial t}}{\left\| \frac{\partial \mathbf{f}(s,t)}{\partial s} \times \frac{\partial \mathbf{f}(s,t)}{\partial t} \right\|_2}$$

heißt **Normalenvektor** der Fläche F an der Stelle $\mathbf{f}(s,t)$.
Die Punktmenge

$$T := \{\mathbf{f}(s,t) + \alpha \frac{\partial \mathbf{f}(s,t)}{\partial s} + \beta \frac{\partial \mathbf{f}(s,t)}{\partial t}, \quad \alpha \in \mathbb{R},\ \beta \in \mathbb{R}\}$$

heißt **Tangentialebene** der Fläche F an der Stelle $\mathbf{f}(s,t)$.

Als Beispiel betrachten wir die Oberfläche der oberen Hälfte der Einheitskugel. Sie wird dargestellt durch

$$\mathbf{f}(s,t) = (s, t, \sqrt{1-s^2-t^2}), \quad S = \{(s,t) \in \mathbb{R}^2,\ s^2+t^2 < 1\}.$$

Es ist

$$\frac{\partial \mathbf{f}(s,t)}{\partial s} = \begin{pmatrix} 1 \\ 0 \\ -\frac{s}{\sqrt{1-s^2-t^2}} \end{pmatrix} \quad , \quad \frac{\partial \mathbf{f}(s,t)}{\partial t} = \begin{pmatrix} 0 \\ 1 \\ -\frac{t}{\sqrt{1-s^2-t^2}} \end{pmatrix}.$$

Da

$$\frac{\partial \mathbf{f}(s,t)}{\partial s} \times \frac{\partial \mathbf{f}(s,t)}{\partial t} = \begin{pmatrix} \frac{s}{\sqrt{1-s^2-t^2}} \\ \frac{t}{\sqrt{1-s^2-t^2}} \\ 1 \end{pmatrix} \quad , \quad \left\| \frac{\partial \mathbf{f}(s,t)}{\partial s} \times \frac{\partial \mathbf{f}(s,t)}{\partial t} \right\|_2 = \frac{1}{\sqrt{1-s^2-t^2}},$$

ist der Normalenvektor

$$\mathbf{N}(s,t) = \begin{pmatrix} s \\ t \\ \sqrt{1-s^2-t^2} \end{pmatrix}.$$

§6. Flächeninhalt von Flächenstücken. Oberflächenintegrale

Die Fläche F im $\mathbb{R}^3$ sei (wie in §5) dargestellt durch $(\mathbf{f}, S)$. Dann nennen wir

$$\int_F d\sigma := \int_S \left\| \frac{\partial \mathbf{f}(s,t)}{\partial s} \times \frac{\partial \mathbf{f}(s,t)}{\partial t} \right\|_2 d(s,t) \tag{6.1}$$

den **Oberflächeninhalt** der Fläche F.
Ist eine stetige Funktion $\rho : F \to \mathbb{R}$ (zum Beispiel: $\rho =$ Ladungsdichte) gegeben, dann nennen wir

$$\int_F \rho \; d\sigma := \int_S \rho(\mathbf{f}(s,t)) \left\| \frac{\partial \mathbf{f}(s,t)}{\partial s} \times \frac{\partial \mathbf{f}(s,t)}{\partial t} \right\|_2 d(s,t)$$

das **Oberflächenintegral** von ρ über F.

Als Beispiel berechnen wir die Oberfläche einer Kugel vom Radius 1: Es ist nach (6.1) und dem Beispiel in §5

$$\begin{aligned} \text{Oberfläche} &= 2\int_F d\sigma = 2\int_S \left\| \frac{\partial \mathbf{f}(s,t)}{\partial s} \times \frac{\partial \mathbf{f}(s,t)}{\partial t} \right\|_2 d(s,t) \\ &= 2\int_S \frac{1}{\sqrt{1-s^2-t^2}} d(s,t) = 2\int_0^1 \Big(\int_0^{2\pi} \frac{r}{\sqrt{1-r^2}} d\phi \Big) dr \\ &= 4\pi \int_0^1 \frac{r}{\sqrt{1-r^2}} dr = 4\pi. \end{aligned}$$

Ist die Ladungsdichte ρ nur von der Höhe $\sqrt{1-s^2-t^2}$ abhängig,
etwa $\rho(s,t) = \sqrt{1-s^2-t^2}$, dann wird

$$\int_F \rho \; d\sigma = \int_S \frac{\sqrt{1-s^2-t^2}}{\sqrt{1-s^2-t^2}} \; d(s,t) = \int_S d(s,t) = \int_0^1 \Big(\int_0^{2\pi} r d\phi \Big) dr = \pi.$$

§7. Die Integralsätze von Gauß, Stokes und Green

In Kapitel II haben wir den Hauptsatz der Infinitesimalrechnung kennengelernt: Für eine auf dem Intervall $[a, b]$ stetige und in (a, b) stetig differenzierbare Funktion f gilt

$$\int_a^b f'(x)dx = f(b) - f(a).$$

Die in diesem Paragraphen vorgestellten Sätze von Gauß, Stokes und Green sind die "Verwandten" dieses Hauptsatzes für Funktionen in mehreren Variablen.

Satz 7.1. (Gaußscher Integralsatz im $\mathbb{R}^3$.)
Sei $D \subset \mathbb{R}^3$ abgeschlossen. Der Rand ∂D von D bestehe aus endlich vielen glatten Flächenstücken $F_1, F_2, \ldots, F_r$. Sei $\mathbf{V} = (V_1, V_2, V_3) : D \to \mathbb{R}^3$ ein Vektorfeld auf D, das noch in einer offenen Umgebung $\mathcal{U}$ von D stetig differenzierbar ist. Es bezeichne $\mathbf{N}(x, y, z)$ den **äußeren** *Normalenvektor von D im Randpunkt $(x, y, z) \in \partial D$. Dann gilt*

$$\int_D \text{div } \mathbf{V}(x, y, z) d(x, y, z) = \int_{\partial D} \mathbf{V} \cdot \mathbf{N} \; d\sigma.$$

Ausführlich geschrieben sieht dies so aus :

$$\int_D \Big(\frac{\partial V_1}{\partial x} + \frac{\partial V_2}{\partial y} + \frac{\partial V_3}{\partial z} \Big) d(x, y, z) = \int_{\partial D} \Big(V_1 N_1 + V_2 N_2 + V_3 N_3 \Big) \; d\sigma. \tag{7.1}$$

Anschauliche Deutung von (7.1):
Stationäre (=zeitunabhängige) Strömung einer inkompressiblen Flüssigkeit im $\mathbb{R}^3$ der konstanten Dichte 1.
Es sei $\mathbf{V}(x, y, z) = \Big(V_1(x, y, z), V_2(x, y, z), V_3(x, y, z) \Big)$ der Geschwindigkeitsvektor.
Die Gesamtheit der aus $D \subset \mathbb{R}^3$ austretenden Flüssigkeit ist

$$M := \int_{\partial D} \mathbf{V} \cdot \mathbf{N} \; d\sigma.$$

Nach Satz 7.1 ist

$$M = \int_D \text{div } \mathbf{V} \; d(x, y, z).$$

Daher heißt div $\mathbf{V}(x, y, z) =: \sigma(x, y, z)$ die **Quelldichte** des **Vektorfeldes V** an der Stelle (x, y, z).
Ist insbesondere $\mathbf{V} = \text{grad } \phi$ das Gradientenfeld eines **Potentials** $\phi \in C^2(D)$, so erfüllt ϕ im Innern von D wegen

$$\text{div}(\text{grad } \phi) = \Delta\phi$$

die **Poissonsche** Gleichung

$$\Delta\phi = \sigma.$$

Beispiel zum Integralsatz von Gauß:
Sei $\lambda \neq 0$ vorgegeben. Wir betrachten in $\mathbb{R}^3 \setminus \{0,0,0\}$ das Vektorfeld

$$\mathbf{V}(x,y,z) = \frac{-e^{-\lambda r}}{r^3}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad , \quad r := \sqrt{x^2+y^2+z^2}.$$

Die Divergenz dieses Vektorfeldes ist, wie Sie sicher leicht nachrechnen können, gegeben durch

$$\text{div } \mathbf{V}(x,y,z) = \frac{\lambda e^{-\lambda r}}{r^2}.$$

Seien nun $0 < \varepsilon < R$ vorgegeben, und sei $D = D_{\varepsilon,R}$ gegeben durch $D = \{(x,y,z) \in \mathbb{R}^3 : \varepsilon^2 \leq x^2+y^2+z^2 \leq R^2\}$. Dann ist die linke Seite von (7.1) des Gaußschen Integralsatzes

$$\begin{aligned} &\int_D \text{div } \mathbf{V}(x,y,z) \ d(x,y,z) = \int_D \frac{\lambda e^{-\lambda r}}{r^2} \ d(x,y,z) \\ &= \int_\varepsilon^R \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} \frac{\lambda e^{-\lambda r}}{r^2} r^2 \cos\theta \ d\theta \ d\phi \ dr = 4\pi\Big(e^{-\lambda\varepsilon} - e^{-\lambda R}\Big). \end{aligned}$$

Die Oberfläche von $D = D_{\varepsilon,R}$ besteht aus den Kugeloberflächen O_R und O_ε mit Radius R bzw. ε. Deren äußere Normalenvektoren sind

$$\mathbf{N}(x,y,z) = \frac{1}{R}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad , \quad \mathbf{N}(x,y,z) = -\frac{1}{\varepsilon}\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Dann ist die rechte Seite von (7.1) des Gaußschen Integralsatzes

$$\begin{aligned} \int_{\partial D} \mathbf{V}\cdot\mathbf{N} \ d\sigma &= \int_{O_R} \mathbf{V}\cdot\mathbf{N} \ d\sigma + \int_{O_\varepsilon} \mathbf{V}\cdot\mathbf{N} \ d\sigma \\ &= \int_{O_R} \frac{-e^{-\lambda R}}{R^2} \ d\sigma + \int_{O_\varepsilon} \frac{e^{-\lambda\varepsilon}}{\varepsilon^2} \ d\sigma = 4\pi e^{-\lambda\varepsilon} - 4\pi e^{-\lambda R}. \end{aligned}$$

Beachte: Beim Grenzübergang $\varepsilon \to 0$ wird $D_{\varepsilon,R}$ zu der Menge $D = \{(x,y,z) : 0 < x^2+y^2+r^2 \leq R^2\}$. Bei der Berechnung von

$$\int_D \text{div } \mathbf{V}(x,y,z) \ d(x,y,z)$$

dürfen wir schlampig sein und sofort $\varepsilon = 0$ setzen. Dies hat keine Auswirkungen. Bei der Berechnung von

$$\int_{\partial D} \mathbf{V}\cdot\mathbf{N} \ d\sigma$$

hätte die Schlamperei, ε sofort gleich Null zu setzen, katastrophale Folgen. Denn dann würden wir die Oberfläche O_ε nicht berücksichtigen und den Wert

$$\lim_{\varepsilon\to 0}\int_{O_\varepsilon} \mathbf{V}\cdot\mathbf{N}\ d\sigma = 4\pi$$

unter den Tisch fallen lassen.

Folgerungen aus dem Integralsatz von Gauß:

1. Sei $\phi \in C^2(\mathcal{U})$. Wir wenden Satz 7.1 an für das Gradientenfeld $\mathbf{V} = \text{grad}\ \phi$. Wegen

$$\text{div}\ \mathbf{V} = \text{div}(\text{grad}\ \phi) = \Delta\phi$$

und

$$\frac{\partial\phi}{\partial\mathbf{N}} = \sum_{k=1}^{3}\frac{\partial\phi}{\partial x_k}N_k = \Big(\text{grad}\ \phi\Big)\cdot\mathbf{N} = \mathbf{V}\cdot\mathbf{N}$$

wird aus (7.1)

$$\int_D \Delta\phi\ d(x,y,z) = \int_{\partial D}\frac{\partial\phi}{\partial\mathbf{N}}\ d\sigma.$$

2. Sei zusätzlich $g \in C^1(\mathcal{U})$ gegeben. Wir wollen (7.1) anwenden für das Vektorfeld $g\mathbf{V} = g\cdot\text{grad}(\phi)$.
Es ist

$$\begin{aligned}
\text{div}(g\mathbf{V}) &= \frac{\partial}{\partial x_1}\Big(g\frac{\partial\phi}{\partial x_1}\Big) + \frac{\partial}{\partial x_2}\Big(g\frac{\partial\phi}{\partial x_2}\Big) + \frac{\partial}{\partial x_3}\Big(g\frac{\partial\phi}{\partial x_3}\Big)\\
&= \frac{\partial g}{\partial x_1}\frac{\partial\phi}{\partial x_1} + \frac{\partial g}{\partial x_2}\frac{\partial\phi}{\partial x_2} + \frac{\partial g}{\partial x_3}\frac{\partial\phi}{\partial x_3} + g\frac{\partial^2\phi}{\partial x_1^2} + g\frac{\partial^2\phi}{\partial x_2^2} + g\frac{\partial^2\phi}{\partial x_3^2}\\
&= (\text{grad}\ g)\cdot(\text{grad}\ \phi)\ +\ g\Delta\phi.
\end{aligned}$$

Wegen

$$(g\mathbf{V})\cdot\mathbf{N} = g(\mathbf{V}\cdot\mathbf{N}) = g\frac{\partial\phi}{\partial\mathbf{N}}$$

folgt daher aus (7.1) die bekannte Formel

$$\int_D g\Delta\phi\ d(x,y,z) + \int_D (\text{grad}\ \phi)\cdot(\text{grad}\ g)\ d(x,y,z) = \int_{\partial D} g\frac{\partial\phi}{\partial\mathbf{N}}\ d\sigma.$$

3. Aus **2.** folgt für $\phi \in C^2(\mathcal{U})$ und $g \in C^2(\mathcal{U})$ die Formel

$$\int_D \Big(g\Delta\phi - \phi\Delta g\Big)d(x,y,z) = \int_{\partial D}\Big(g\frac{\partial\phi}{\partial\mathbf{N}} - \phi\frac{\partial g}{\partial\mathbf{N}}\Big)d\sigma.$$

Ist $D \subset \mathbb{R}^3$ ein auf der $x-y-$Ebene stehender Zylinder der Höhe 1 und Grundfläche $D_0 \subset \mathbb{R}^2$ und $V_3 = 0$, so kann man aus Satz 7.1 mit wenig Mühe den nächsten Satz herleiten.

Satz 7.2. (Integralsatz von Green im $\mathbb{R}^2$.)
Sei $D_0 \subset \mathbb{R}^2$ abgeschlossen. Der Rand ∂D_0 von D_0 bestehe aus endlich vielen stückweise glatten geschlossenen (orientierten) Kurven $K_1, K_2, \ldots, K_r$, so daß das Innere von D_0 stets links von $K_1, K_2, \ldots, K_r$ liegt. Sei $\mathbf{V} = (V_1, V_2) : D_0 \to \mathbb{R}^2$ ein Vektorfeld auf D_0, das noch in einer offenen Umgebung von D_0 stetig differenzierbar ist. Dann gilt

$$\int_{D_0} \Big(\frac{\partial V_1}{\partial x} + \frac{\partial V_2}{\partial y}\Big) d(x,y) = \sum_{j=1}^{r} \int_{K_j} \Big(- V_2 dx + V_1 dy \Big).$$

Beispiel: Berechnung der Fläche einer Ellipse
Die Fläche einer Ellipse mit Mittelpunkt $(0,0)$ und den Halbachsen $a > 0$ und $b > 0$ ist nach Satz 7.2 für $V_1(x,y) = x$, $V_2(x,y) = 0$ und die Randkurve
$x(t) = a\cos t$, $y(t) = b\sin t$, $0 \le t \le 2\pi$, gegeben durch

$$\text{Fläche} = \int_{D_0} 1 \;\; d(x,y) = \int_{\partial D_0} V_1 dy = \int_0^{2\pi} x(t)y'(t)dt = \int_0^{2\pi} ab \;\; \cos^2 t \;\; dt = ab\pi.$$

Anwendung von Satz 7.2: Planimeter
Mechanische Messung einer Fläche des $\mathbb{R}^2$ durch ein **Planimeter**.
Man wähle in Satz 7.2 $V_1(x,y) = \frac{x}{2}$, $V_2(x,y) = \frac{y}{2}$.

Zum Abschluß formulieren wir einen weiteren Intergalsatz.

Satz 7.3. (Integralsatz von Stokes für Kurven auf Flächen im $\mathbb{R}^3$.)
Sei $D \subset \mathbb{R}^3$ offen. Sei K eine geschlossene, stückweise glatte (orientierte) Kurve im $\mathbb{R}^3$ und F ein darin eingespanntes Flächenstück, dessen Normalenvektor $\mathbf{N}$ stückweise steig sei und mit der Orientierung von K eine positive Schraube bildet.
Sei $F \cup K \subset D$ und $\mathbf{V} = (V_1, V_2, V_3) \in C^1(D)$. Dann gilt die Stokessche Formel

$$\int_F \Big(\{rot\ \mathbf{V}\} \cdot \mathbf{N}\Big) d\sigma = \int_K \Big(V_1 dx + V_2 dy + V_3 dz\Big),$$

wobei die rechte Seite die "Energiezunahme" angibt.

— finis —