Zur Einleitung: Lineare Gleichungssysteme

Wir untersuchen zunächst mit Methoden, die Sie vermutlich aus der Schule kennen, explizit einige kleine lineare Gleichungssysteme. Das Gleichungssystem I wird dabei sukzessive in "einfachere" Gleichungssysteme II, III, ... umgeformt. Die Umformungen sind so, daß sich die Lösungsmenge nicht ändert.

1) 1 Gleichung, 1 Unbekannte:

$$\begin{array}{lrl} \mathrm{I} & 3x+5=9 & \quad \mid -5 \\ \mathrm{II} & 3x=4 & \quad \mid :3 \\ \mathrm{II} & x=\frac{4}{3} & \end{array}$$

Lösungsmenge $L_{\mathrm{I}} = \{x \in \mathbb{R} \mid 3x+5=9\} = L_{\mathrm{II}} = \{x \in \mathbb{R} \mid 3x = 4\} = L_{\mathrm{III}} = \left\{\frac{4}{3}\right\}$

2) 1 Gleichung, 2 Unbekannte:

$$\begin{array}{lll} \mathrm{I} & 2x+4y=8 & \quad \mid :4 \mid -\frac{x}{2} \\ \mathrm{II} & y=-\frac{x}{2}+2 & \end{array}$$

Betrachte geordnete Paare (x,y) mit $x \in \mathbb{R}, y \in \mathbb{R}$ und setze $\mathbb{R}^2 = \{(x,y) \mid x \in \mathbb{R} \text{ und } y \in \mathbb{R}\}$

$$L_I = \{(x,y) \in \mathbb{R}^2 \mid 2x+4y=8\} = L_{II} = \{(x, -\tfrac{x}{2}+2) \mid x \in \mathbb{R}\}$$

Es gibt unendlich viele Lösungen!

3) 2 Gleichungen, 2 Unbekannte:

$$\begin{array}{lrcll} \mathrm{I} & 2x+4y & = & 8 & \mid \ \cdot\left(-\frac{3}{2}\right) \\ & 3x-6y & = & 4 & \hookleftarrow \\[1em] \mathrm{II} & 2x+\ \ 4y & = & 8 & \\ & 0x-12y & = & -8 & \mid \ :(-12) \\[1em] \mathrm{III} & 2x+4y & = & 8 & \qquad 2x+\frac{8}{3}=8 \qquad 2x=\frac{16}{3} \qquad x=\frac{8}{3} \\ & y & = & \frac{2}{3} & \\[1em] \mathrm{IV} & x & = & \frac{8}{3} & \\ & y & = & \frac{2}{3} & \end{array}$$

$L_I = \{(x,y) \in \mathbb{R}^2 \mid (x,y) \text{ löst (I)}\} = \left\{\left(\frac{8}{3}, \frac{2}{3}\right)\right\}$, genau eine Lösung.

4) 2 Gleichungen, 2 Unbekannte:

$$\begin{array}{lrcll} \mathrm{I} & 2x+4y & = & 8 & \mid \ \cdot\left(-\frac{3}{2}\right) \\ & 3x+6y & = & 4 & \hookleftarrow \\[1em] \mathrm{II} & 2x+4y & = & 8 & \\ & 0x+0y & = & -8 & \end{array}$$

keine Lösung, $\quad L_I = \{(x,y) \in \mathbb{R}^2 \mid (x,y) \text{ erfüllt (I)}\} = \emptyset$

Geometrische Interpretation: Für eine Gleichung $ax + by = c$ mit $a \neq 0$ oder $b \neq 0$ (das wird im folgenden Abschnitt stets vorausgesetzt) ist die Lösungsmenge eine Gerade in der Ebene $\mathbb{R}^2$. (Für $b \neq 0$ etwa gegeben durch $y = -\frac{a}{b}x + \frac{c}{b}$). Zwei lineare Gleichungen mit zwei Unbekannten entsprechen also zwei Geraden und der Lösungsmenge des Gleichungssystems entspricht die Menge der Punkte, die auf jeder der beiden Geraden liegen. Im allgemeinen gibt es also genau eine Lösung des Gleichungssystems (= genau einen Schnittpunkt der beiden Geraden). Folgende Ausnahmefälle sind möglich:

a Verschiedene parallele Geraden ($\Rightarrow$ keine Lösung des Gleichungssystems)
b die beiden Geraden stimmen überein ($\Rightarrow$ unendlich viele Lösungen des Gleichungssystems, Lösungsraum durch einen reellen Parameter parametrisierbar).

Bemerkung: Zum Lösen haben wir nur die Grundrechenarten benützt. Sind die Koeffizienten (a, b etc.) rational (oder komplex), so liefert dieses Verfahren rationale (oder komplexe) Lösungen.

Gaußsches Eliminationsverfahren (überführt Gleichungssystem in äquivalentes in Stufenform $\rightarrow$ rekursiv auflösbar).

5) 3 Gleichungen mit 4 Unbekannten (genannt x_1, x_2, x_3, x_4).

$$\text{I} \qquad \begin{array}{rcrcrcrcl} 2x_1 & + & 4x_2 & + & 4x_3 & + & 6x_4 & = & 2 \\ 3x_1 & + & 7x_2 & + & 5x_3 & + & 9x_4 & = & 4 \\ x_1 & + & 3x_2 & - & x_3 & + & 5x_4 & = & 1 \end{array} \qquad \begin{array}{l} \mid \cdot \left(-\frac{3}{2}\right) \quad \mid \\ \leftarrow\mid \qquad\quad \mid \cdot \left(-\frac{1}{2}\right) \\ \qquad\qquad \leftarrow\mid \end{array}$$

$$\text{II} \qquad \begin{array}{rcl} 2x_1 + 4x_2 + 4x_3 + 6x_4 & = & 2 \\ x_2 - x_3 & = & 1 \\ x_2 - 3x_3 + 2x_4 & = & 0 \end{array} \qquad \begin{array}{l} \\ \mid \ (-1) \\ \leftarrow\mid \end{array}$$

$$\text{III} \qquad \begin{array}{rcrl} 2x_1 + 4x_2 + 4x_3 + 6x_4 & = & 2 & \text{III}_1 \\ x_2 - x_3 & = & 1 & \text{III}_2 \\ -2x_3 + 2x_4 & = & -1 & \text{III}_3 \end{array} \qquad \begin{array}{c} \text{Gleichungssystem in Stufenform} \\ (\text{“rekursiv ausflösbar”}) \end{array}$$

Lösung: Wähle $x_4 = r \in \mathbb{R}$ beliebig

$$\begin{array}{lclcl} \text{III}_3 \Rightarrow x_3 & = & x_4 + \frac{1}{2} & = & r + \frac{1}{2} \\ \text{III}_2 \Rightarrow x_2 & = & x_3 + 1 & = & r + \frac{3}{2} \\ \text{III}_1 \Rightarrow x_1 & = & -2x_2 - 2x_3 - 3x_4 + 1 & = & -7r - 3 \end{array}$$

$$\begin{array}{rcl} \mathbb{R}^4 & = & \{(x_1, x_2, x_3, x_4) \mid x_1 \in \mathbb{R}, x_2 \in \mathbb{R}, x_3 \in \mathbb{R}, x_4 \in \mathbb{R}\} \\ L_I & = & \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid (x_1, x_3, x_3, x_4) \text{ löst I}\} \\ & = & \{(-7r - 3, r + \frac{3}{2}, r + \frac{1}{2}, r) \mid r \in \mathbb{R}\} \text{ Parameterdarstellung des Lösungsraums} \end{array}$$

Kapitel 1: Zur Sprache der Mathematik

übliche Sprache + Fachausdrücke + mathematische Symbole

Abkürzende logische Symbole

(1.1) "$A \Rightarrow B$" bedeutet "aus Aussage A folgt Aussage B".
Beispiel: $x \in \mathbb{R} \Rightarrow x^2 \geq 0$
Andere sprachliche Formen:
"A impliziert B" oder
"A ist hinreichend für B"
"B ist notwendig für A"

(1.2) "$A \Leftrightarrow B$" bedeutet "$A \Rightarrow B$ und $B \Rightarrow A$".
Beispiel: $x \in \mathbb{R}$ und $3x \in \mathbb{Q} \Leftrightarrow x \in \mathbb{Q}$
Andere sprachliche Formen: "A und B sind äquivalent" oder "A gilt genau dann (dann und nur dann), wenn B gilt".

(1.3) "$\exists$" steht für "Es existiert (mindestens) ein ..."
Beispiel: "$\exists x \in \mathbb{R} : x^2 = 2$ steht für "Es existiert ein $x \in \mathbb{R}$, so daß $x^2 = 2$ gilt" (nämlich $x = \sqrt{2}$ oder $x = -\sqrt{2}$.

(1.4) "$\forall$" steht für "Für alle ..."
Beispiel: "$\forall x \in \mathbb{R} : x^2 \geq 0$" steht für "Für alle reellen Zahlen x gilt: $x^2 \geq 0$"

Schließlich: "oder" ist nicht ausschließend (im Gegensatz zu "entweder ... oder").
Die Aussage "Es gilt $1 + 1 = 2$ oder es gilt $1 - 1 = 0$" ist wahr.
Die Aussage "Entweder es gilt $1 + 1 = 2$ oder es gilt $1 - 1 = 0$" ist falsch.

(1.5) "$A := B$" steht für "A ist durch B definiert" oder "A ist nach Definition gleich B".

Mengen und Abbildungen (Georg Cantor 1845-1918)

Menge = Zusammenfassung von (mathematischen) Objekten (=Elementen)
"$x \in M$" steht für "x ist Element der Menge M",
"$x \notin M$" steht für "x ist nicht Element der Menge M".

Angabe von Mengen:

(1.6) Aufzählend, z.B. $M = \{1, 3, 5, 7\} (= \{1, 1, 3, 5, 7\})$

(1.7) Durch Angabe von Eigenschaften: Sei M eine Menge und $A(x)$ eine Aussage über die Elemente von M. Dann kann man eine Menge N definieren durch

$$N = \{x \mid x \in M \text{ und } A(x) \text{ ist wahr}\} = \{x \in M \mid A(x) \text{ ist wahr}\}$$

z.B. $\mathbb{Q} = \{x \in \mathbb{R} \mid \exists p \in \mathbb{Z}, \exists q \in \mathbb{N}, q \neq 0 : x = \frac{p}{q}\}$ Menge der rationalen Zahlen oder $M = \{x \in \mathbb{N} \mid x$ ist ungerade natürliche Zahl kleiner als $9\}$.

Beispiele von Mengen:

$$\begin{aligned} \mathbb{N} &= \{x \in \mathbb{R} \mid x \text{ natürliche Zahl}\} = \{0, 1, 2, 3, \ldots\} \\ \mathbb{Z} &= \{x \in \mathbb{R} \mid x \text{ ganze Zahl}\} = \{0, 1, -1, 2, -2, \ldots\} \end{aligned}$$

Vielleicht kennen Sie schon $\mathbb{C} = \{x + iy \mid x \in \mathbb{R}, y \in \mathbb{R}\}$
$\emptyset$ die leere Menge (enthält kein Element)

(1.8) Def.: Eine Menge A heißt Teilmenge einer Menge B("$A \subseteq B$"), falls jedes Element von A auch Element von B ist, d.h. falls gilt

$$x \in A \Rightarrow x \in B.$$

Beispiel: $\emptyset \subseteq \mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$

(1.9) Zwei Mengen A und B sind gleich ("$A = B$"), wenn A und B die selben Elemente enthalten, d.h. falls gilt

$$x \in A \Leftrightarrow x \in B$$

Oder: $A = B \Leftrightarrow A \subseteq B$ und $B \subseteq A$.

Beispiel: $\{x \mid x \in \mathbb{N}, x$ ungerade natürliche Zahl kleiner $9\} = \{1, 3, 5, 7\}$

Konstruktionen mit Mengen A und B:

(1.10) Schnittmenge (Durchschnitt) $A \cap B$ von A und B:

$$x \in A \cap B \Leftrightarrow x \in A \text{ und } x \in B$$

Vereinigung $A \cup B$ von A und B:

$$x \in A \cup B \Leftrightarrow x \in A \text{ oder } x \in B$$

Komplement $A \setminus B$ von B in A (Differenzmenge von A und B)

$$x \in A \setminus B \Leftrightarrow x \in A \text{ und } x \notin B$$

(1.11) Rechenregeln (für Vereinigung, Durchschnitt, Komplement)

Sind A, B, C Mengen, so gilt:

(a) $A \cup B = B \cup A,\ A \cap B = B \cap A$ "Kommutativgesetze"

(b) $(A \cup B) \cup C = A \cup (B \cup C)$ "Assoziativgesetze"
$(A \cap B) \cap C = A \cap (B \cap C)$

(c) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ "Distributivgesetze"
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

(d) $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$ "De Morgansche Regeln"
$A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$

Beweis von (c):
Sei $x \in A \cap (B \cup C) \Rightarrow x \in A$ und $(x \in B$ oder $x \in C) \Rightarrow$
$(x \in A$ und $x \in B)$ oder $(x \in A$ und $x \in C) \Rightarrow x \in (A \cap B) \cup (A \cap C)$

Sei $x \in (A \cap B) \cup (A \cap C) \Rightarrow (x \in A$ und $x \in B)$ oder $(x \in A$ und $x \in C)$
$\Rightarrow x \in A$ und $(x \in B$ oder $x \in C) \Rightarrow x \in A \cap (B \cup C)$

(1.12) Kartesische Produkt (benannt nach R. Descartes 1596-1650). Sind A, B Mengen, so ist das <u>kartesische Produkt</u> $A \times B$ <u>von</u> A <u>und</u> B die Menge aller geordneten Paare (a, b) mit $a \in A$ und $b \in B$.

$$A \times B := \{(a, b) \mid a \in A, b \in B\}$$

Bem.: $(a, b) = (a', b') \Leftrightarrow a = a'$ und $b = b'$!

Bem.: $\mathbb{R} \times \mathbb{R} =: \mathbb{R}^2$

Allgemeiner: Sind $A_1, \ldots, A_n$ Mengen, so

$$A_1 \times \ldots \times A_n = \{(a_1, \ldots, a_n) \mid a_1 \in A_1, a_2 \in A_2, \ldots, a_n \in A_n\}$$

Dabei gilt – per definitionem – $(a_1, \ldots, a_n) = (a'_1, \ldots, a'_n)$ genau dann, wenn $a_1 = a'_1$, …, $a_n = a'_n$ gelten.

Bem.: $\underbrace{\mathbb{R} \times \ldots \times \mathbb{R}}_{n\text{-mal}} =: \mathbb{R}^n$. Ein Element $(a_1, \ldots, a_n) \in \mathbb{R}^n$ heißt <u>n-Tupel</u> von reellen Zahlen.

Exkurs: Nochmals lineare Gleichungssysteme

Das allgemeine lineare Gleichungssystem mit m Gleichungen und n Unbekannten:

$$\mathrm{I} \quad \begin{array}{lccccccc} I_1 & a_{11}x_1 & + & a_{12}x_2 & +\ldots+ & a_{1n}x_n & = & b_1 \\ I_2 & a_{21}x_1 & + & a_{22}x_2 & +\ldots+ & a_{2n}x_n & = & b_2 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ I_m & a_{m1}x_1 & + & a_{m2}x_2 & +\ldots+ & a_{mn}x_n & = & b_m \end{array}$$

Gegeben sind die $a_{11}, \ldots, a_{mn}, b_1, \ldots, b_m$, gesucht die $x_1, \ldots, x_n$.

Lösungsmenge $L_I = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid (x_1, \ldots, x_n)$ löst $I\}$

(1.13) Satz. *Sei $a_{11} \neq 0$ und sei I' das Gleichungssystem mit den Zeilen*

$$I'_1 = 1, I'_2 = I_2 - \frac{a_{21}}{a_{11}} I_1, \ldots, I'_m = I_m - \frac{a_{m1}}{a_{11}} I_1.$$

Dann gilt: $L_I = L_{I'}$

Beweis: Sei $\overline{x} = (\overline{x}_1, \ldots, \overline{x}_n) \in L_I$. Dann erfüllt $\overline{x}$ die Gleichung I'_1 $(= I_1)$. Ist $2 \leq i \leq m$, so erfüllt $\overline{x}$ auch die Gleichung I_i, also auch $I'_i = I_i - \frac{a_{i1}}{a_{11}} I_1$. Also $x \in L_{I'}$.

Sei $\overline{x} = (\overline{x}_1, \ldots, \overline{x}_n) \in L_{I'}$. Dann erfüllt $\overline{x}$ die Gleichung I_1 $(= I'_1)$.

Ist $2 \leq i \leq m$, so erfüllt $\overline{x}$ auch die Gleichung I'_i, also auch

$$I_i = I'_i + \frac{a_{i1}}{a_{11}} I_1.$$

Also $x \in L_I$.

(1.14) Definition. Das lineare Gleichungssystem I heißt homogen, falls $b_i = 0$ gilt für alle $i \in \{1, \ldots, m\}$.

Bemerkung: Ist I homogen, so ist $(0, \ldots, 0) \in \mathbb{R}^n$ Lösung von I.

Auf $\mathbb{R}^n$ definieren wir eine Addition durch

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_n + y_n) \quad \text{(komponentenweise Addition)}$$

und eine Multiplikation mit reellen Zahlen $r \in \mathbb{R}$ durch

$$r(x_1, \ldots, x_n) := (rx_1, \ldots, rx_n)$$

(1.15) Satz. *Ist I ein homogenes lineares Gleichungssystem, so hat seine Lösungsmenge L_I folgende Eigenschaften:*

$$\begin{array}{lll} \textit{Ist} \quad x = (x_1, \ldots, x_n) \in L_I \textit{ und } y = (y_1, \ldots, y_n) \in L_I, & & \textit{so gilt } x + y \in L_I. \\ \textit{Ist} \quad x = (x_1, \ldots, x_n) \in L_I \textit{ und } r \in \mathbb{R}, & & \textit{so gilt } rx \in L_I. \end{array}$$

Beweis: $x \in L_I$ und $y \in L_I \Rightarrow$ Für alle $i \in \{1, \ldots, m\}$ gelten: $a_{i1}x_1 + \ldots + a_{in}x_n = 0$ und $a_{i1}y_1 + \ldots + a_{in}y_n = 0$. Summiert man diese Gleichungen, so folgt

$$a_{i1}(x_1 + y_1) + \ldots + a_{in}(x_n + y_n) = 0 + 0 = 0$$

Also erfüllt $x+y = (x_1+y_1, \ldots, x_n+y_n)$ für alle $i \in \{1, \ldots, m\}$ die Gleichung I_i, d.h. $x+y \in L_I$.

Die 2. Behauptung folgt analog durch Multiplikation von I_i mit r.

Bemerkung: Besitzt ein homogenes lineares Gleichungssystem I eine Lösung $x \neq (0, \ldots, 0)$, so besitzt I nach (1.15) unendlich viele Lösungen ($\forall r \in \mathbb{R} : rx \in L_I$).

Zu einem linearen Gleichungssystem I bezeichne I^{hom} das zugehörige homogene lineare Gleichungssystem, das sich nur dadurch von I unterscheidet, daß alle rechten Seiten von I durch 0 ersetzt werden.

(1.16) Satz. *Es existiere eine Lösung $\overline{x} = (\overline{x}_1, \ldots, \overline{x}_n)$ des linearen Gleichungssystems I. Dann gilt*

$$L_I = \{\overline{x} + x \mid x \in L_{I^{\text{hom}}}\}.$$

Ohne die Mengensprechweise drückte man die Aussage von (1.16) früher durch folgenden langen Satz aus: Die allgemeine Lösung eines linearen Gleichungssystems ist eine spezielle Lösung dieses Gleichungssystems plus die allgemeine Lösung des zugehörigen homogenen Gleichungssystems.

Beweis:

a) Sei $y \in L_I$. Wie zeigen: $y - \overline{x} \in L_{I\text{hom}}$ (wobei $y - \overline{x} := y + (-1)\overline{x}$).

$$\{y, \overline{x}\} \subseteq L_I \Rightarrow \forall i \in \{1, \ldots, m\} \quad : \quad \begin{array}{l} a_{i1}y_1 + \ldots + a_{in}y_n = b_i \text{ und} \\ a_{i1}\overline{x}_1 + \ldots + a_{in}\overline{x}_n = b_i \end{array}$$

Subtraktion dieser Gleichungen liefert:

$$a_{i1}(y_1 - \overline{x}_1) + \ldots + a_{in}(y_n - \overline{x}_n) = b_i - b_i = 0$$

Also: $y - \overline{x} \in L_{I\text{hom}}$. Wir nennen $y - \overline{x} =: x \in L_{I\text{hom}}$ und erhalten $y = \overline{x} + x$ mit $x \in L_{I\text{hom}}$.

b) Sei $x \in L_{I\text{hom}}$. Dann gilt für alle $i \in \{1, \ldots, m\}$:

$$a_{i1}x_1 + \ldots + a_{in}x_n = 0$$

Wegen $y \in L_I$ gilt für alle $i \in \{1, \ldots, m\}$:

$$a_{i1}y_1 + \ldots + a_{in}y_n = b_i$$

Addition dieser Gleichungen liefert:

$$a_{i1}(x_1 + y_1) + \ldots + a_{im}(x_m + y_m) = b_i \quad (\text{für alle } i \in \{1, \ldots, m\})$$

also $y + x \in L_I$.

Zurück zur Mengenlehre:

Abbildungen zwischen Mengen

(1.17) Def.: Es seien M, N Mengen. Eine Abbildung $f : M \to N$ von M nach N ist eine Vorschrift, die jedem $x \in M$ genau ein Element $f(x) \in N$ zuordnet.

Beispiele:

a) $M = N = \mathbb{R}$, $f : \mathbb{R} \to \mathbb{R}$, $\forall x \in \mathbb{R} : f(x) := 3x^3 - x + 1$
b) $M = \mathbb{R}$, $N = \mathbb{R}_{>0} = \{x \in \mathbb{R} \mid x > 0\}$, $\forall x \in \mathbb{R} : f(x) := e^x$.
c) $f : \mathbb{Z} \to \{0, 1\}$, $f(x) := \begin{cases} 0 & x \text{ gerade} \\ 1 & x \text{ ungerade} \end{cases}$
d) M Menge. $\mathrm{id}_M : M \to M$, $\forall x \in M : \mathrm{id}_M(x) := x$. "Identität von M".
e) $A \subseteq B$. $i : A \to B$, $\forall x \in A : i(x) := x$ "Inklusionsabbildung"
f) Seien $A_1, \ldots, A_n$ Mengen und $i \in \{1, \ldots, n\}$.
$p_i : A_1 \times \ldots \times A_n \to A_i, p_i(a_1, \ldots, a_n) := a_i$ "Projektion auf die i'te Komponente".
g) $+ : \mathbb{R}^2 \to \mathbb{R}$, $\forall (x, y) \in \mathbb{R}^2 : +(x, y) := x + y$

(1.18) Def.: Eine Abbildung $f : M \to N$ heißt

(a) injektiv, falls für alle $x \in M, y \in M$ gilt: $x \neq y \Rightarrow f(x) \neq f(y)$.
(b) surjektiv, falls für alle $y \in N$ ein $x \in M$ existiert, so daß $f(x) = y$ gilt.
(c) bijektiv, falls f injektiv und surjektiv is.

Von den obigen Beispielen ist a) surjektiv, aber nicht injektiv, während b) bijektiv ist.

(1.19) Def. (Komposition, Hintereinanderausführung) Seien A, B, C, D Mengen mit $B \subseteq C$ und $f : A \to B$, $g : C \to D$ Abbildungen. Dann ist $g \circ f : A \to D$ definiert durch: Für alle $x \in A$ ist $(g \circ f)(x) := g(\underbrace{f(x)}_{\in B})$.

Symbolisch: $\overbrace{A \xrightarrow{f} B \subseteq C \xrightarrow{g} D}^{g \circ f}$

Bem.: $A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D \Rightarrow h \circ (g \circ f) = (h \circ g) \circ f =: h \circ g \circ f$

Schreibweise: Ist $f : A \to B$ Abbildung und $C \subseteq A$, so bezeichne

$$f(C) := \{y \in B \mid \exists x \in C : f(x) = y\} \text{“Bild von } C \text{ unter } f\text{”}$$

Ist $D \subseteq B$, so bezeichne

$$f^{-1}(D) := \{x \in A \mid f(x) \in D\} \text{ “Urbild von } D \text{ unter } f\text{”}$$

Bem:

$f : A \to B$ surjektiv $\Leftrightarrow$ $f(A) = B$.
$f : A \to B$ injektiv $\Leftrightarrow$ Für alle $y \in B$ enthält $f^{-1}(\{y\})$ höchstens 1 Element.

(1.20) Rechenregeln: Sei $f : M \to N$ Abbildung, $A \subseteq M$, $B \subseteq M$. Dann gilt

(a) $f(A \cup B) = f(A) \cup f(B)$
(b) $f(A \cap B) \subseteq f(A) \cap f(B)$
(c) $f(M \setminus A) \supseteq f(M) \setminus f(A)$ $\quad (\Leftrightarrow f(M) \setminus f(A) \subseteq f(M \setminus A))$
Für alle $C \subseteq N$, $D \subseteq N$ gilt:
(d) $f^{-1}(C \cup D) = f^{-1}(C) \cup f^{-1}(D)$
(e) $f^{-1}(C \cap D) = f^{-1}(C) \cap f^{-1}(D)$
(f) $f^{-1}(N \setminus C) = M \setminus f^{-1}(C)$

Beweis von (f):

$$\begin{gathered} x \in f^{-1}(N \setminus C) \Rightarrow x \in M \text{ und } f(x) \in N \setminus C \\ \Rightarrow x \in M \text{ und } f(x) \notin C \Rightarrow x \in M \setminus f^{-1}(C) \\ x \in M \setminus f^{-1}(C) \Rightarrow x \in M \text{ und } x \notin f^{-1}(C) \Rightarrow x \in M \text{ und } f(x) \notin C \\ \Rightarrow f(x) \in N \setminus f^{-1}(C) \end{gathered}$$

Beispiel zu (b):

$$\begin{gathered} f : \mathbb{R} \to \mathbb{R}, f(x) = x^2 \\ A = \{x \in \mathbb{R} \mid x < 0\}, B = \{x \in \mathbb{R} \mid x > 0\} \\ \Rightarrow A \cap B = \emptyset, \text{ aber } f(A) = f(B) = \{x \in \mathbb{R} \mid x > 0\} (=: \mathbb{R}_{>0}) \\ \text{d.h. } f(A) \cap f(B) = \mathbb{R}_{>0} \neq \emptyset. \end{gathered}$$

(1.21) Def.: Ist $f : M \to N$ Abbildung und $A \subseteq M$, so ist die Einschränkung (Restriktion) $f|A (= f|_A) : A \to N$ von f auf A definiert durch: Für alle $x \in A$ ist $(f|A)(x) = f(x)$.

Bemerkung: Bezeichnet $i : A \to M$ die Inklusion, so gilt $f|A = f \circ i$

(1.22) Fakt: Ist $f : M \to N$ bijektive Abbildung, so existiert genau eine Abbildung $f^{-1} : N \to M$, so daß $f^{-1} \circ f = \mathrm{id}_M$ gilt. Für diese Abbildung f^{-1} gilt auch $f \circ f^{-1} = \mathrm{id}_N$.

Vorsicht: Das Symbol f^{-1} wird in unterschiedlichen Bedeutungen verwendet!

Bew.: f bijektiv $\Leftrightarrow \forall y \in N$ ex. genau ein $x \in M : f(x) = y$.

Eindeutigkeit von f^{-1}: Gilt $f(x) = y$, so muß $\begin{array}{rcl} f^{-1}(y) & = & f^{-1}(f(x)) = (f^{-1} \circ f)(x) \\ & = & \mathrm{id}_M(x) = x \text{ gelten.} \end{array}$

Definiere $f^{-1} : N \to M$ durch $f^{-1}(y) = x \Leftrightarrow y = f(x)$

(1.23) Def.: Seien M, N Mengen. Eine Relation auf (M, N) ist eine Teilmenge von $M \times N$. (Ist $M = N$, so spricht man von einer Relation auf M)

Schreibweise: Sei $R \subseteq M \times N$. Statt $(x, y) \in R$ schreibt man auch xRy.

Beispiele:

1) $\Delta_M = \{(x, x) \mid x \in M\}$ die Diagonale in $M \times M$.
2) Sei $f : M \to N$ Abbildung. Dann ist der Graph von f,

$$\text{graph } (f) := \{(x, f(x)) \mid x \in M\}$$

eine Relation auf (M, N). Bem.: $\Delta_M = \text{graph } (\mathrm{id}_M)$
3) Sei $M = N = \mathbb{R}$ und $R := \{(x, y) \in \mathbb{R}^2 \mid x = y^2\}$. Diese Relation ist nicht Graph einer Funktion $f : \mathbb{R} \to \mathbb{R}$.

(1.24) Def. Sei M Menge, $R \subseteq M \times M$ Relation auf M. Dann heißt

(a) R reflexiv $\Leftrightarrow \forall x \in M : xRx \quad (\Leftrightarrow \Delta_M \subseteq R)$
(b) R symmetrisch $\Leftrightarrow (\forall x, y \in M : xRy \Leftrightarrow yRx)$
(c) R transitiv $\Leftrightarrow (\forall x, y, z \in M : xRy$ und $yRz \Rightarrow xRz)$
(d) R Äquivalenzrelation $\Leftrightarrow R$ ist reflexiv, symmetrisch und transitiv

Bei Äquivalenzrelationen R schreibt man oft statt xRy: $x \sim_R y$ oder kurz $x \sim y$ (wenn klar ist, welches R gemeint ist).

Beispiele:

1) Auf $M = \mathbb{R}$ kennen wir die Relation "kleiner gleich", d.h.

$$R = \{(x, y) \in \mathbb{R}^2 \mid x \leq y\}$$

R ist reflexiv und transitiv, aber nicht symmetrisch. (Beweis?)

2) Sei $M = \mathbb{Z}$, $p \geq 2$ feste natürliche Zahl.

$$R := \{(m,n) \in \mathbb{Z}^2 \mid n - m \text{ ist durch } p \text{ teilbar}\}.$$

Wir zeigen: R ist eine Äquivalenzrelation

(a) $\forall m \in \mathbb{Z}$: $m \sim_R m$, da $m - m = 0$ durch p teilbar: $0 = 0 \cdot p$

(b) Es gelte $m \sim_R n \quad \Rightarrow \quad n - m$ durch p teilbar $\Rightarrow m - n$ durch p teilbar
$\Rightarrow \quad n \sim_R m$

(c) Es gelte $k \sim_R m$ und $m \sim_R n \Rightarrow \exists j \in \mathbb{Z}, l \in \mathbb{Z} : m - k = jp, n - m = lp$
$\Rightarrow n - k = (l + j)p$ und $l + j \in \mathbb{Z} \Rightarrow k \sim_R n$.

Langes Beispiel: Eine Zerlegung $\mathcal{M}$ einer Menge M, ist eine Menge $\mathcal{M}$, deren Elemente Teilmengen von M sind, so daß gilt:

(1) $A \in \mathcal{M}$, $B \in \mathcal{M} \Rightarrow A = B$ oder $A \cap B = \emptyset$.

(2) $\bigcup\limits_{A \in \mathcal{M}} A = M$. (wobei $\bigcup\limits_{A \in \mathcal{M}} A := \{x \mid \exists A \in \mathcal{M} : x \in A\}$)

Wir werden sehen, daß eine Äquivalenzrelation auf M im Grunde das gleiche ist, wie eine Zerlegung von M. Zunächst überlegen wir, wie wir aus einer Zerlegung eine Äquivalenzrelation erhalten: Zu einer Zerlegung $\mathcal{M}$ von M definieren wir die Relation

$$R = \{(x,y) \in M \times M \mid \exists A \in \mathcal{M} : \{x,y\} \subseteq A\}$$

Wir zeigen: R ist eine Äquivalenzrelation.

(a) Sei $x \in M$. Wir wollen zeigen, daß $(x,x) \in R$ gilt. Wegen (2) existiert ein $A \in \mathcal{M}$, so daß $x \in A$ gilt. Also gilt $\{x,x\}(= \{x\}) \subseteq A$, d.h. $(x,x) \in R$.

(b) $\{x,y\} \subseteq A \Leftrightarrow \{y,x\} \subseteq A$

(c) Es gelte $(x,y) \in R$ und $(y,z) \in R \Rightarrow \exists A \in \mathcal{M}$, $B \in \mathcal{M}$ mit $\{x,y\} \subseteq A$, $\{y,z\} \subseteq B$
$\Rightarrow y \in A \cap B$, d.h. $A \cap B \neq \emptyset \overset{(1)}{\Rightarrow} A = B \Rightarrow \{x,z\} \in A \Rightarrow (x,z) \in R$.

Bezeichnung: Sei $R \subseteq M \times M$ eine Äquivalenzrelation und $x \in M$. Dann heißt

$[x]_R \quad := \quad \{y \in M \mid (x,y) \in R\}$ die Äquivalenzklasse von x und

$M/R \quad := \quad \{[x]_R \mid x \in M\}$ die Menge der Äquivalenzklassen, die auch "M modulo R" genannt wird.

(1.25) Fakt: Sei $R \subseteq M \times M$ eine Äquivalenzrelation. Dann ist M/R eine Zerlegung von M.

Beweis: Offensichtlich ist jedes Element von M/R eine Teilmenge von M (nämlich eine Äquivalenzklasse). Wir zeigen, daß die Menge M/R die Eigenschaften (1) und (2) hat.

Zu (2): Offensichtlich gilt $\bigcup\limits_{[x]_R \in M/R} [x]_R \subseteq M$. Sei umgekehrt $x \in M$. Die Reflexivität (a) von R besagt, daß $(x,x) \in R$ gilt, also $x \in [x]_R$. Damit $M \subseteq \bigcup\limits_{[x]_R \in M/R} [x]_R$.

Zu (1): Wir zeigen, daß folgende Aussagen $(\alpha), (\beta), (\gamma)$ über Elemente $x \in M, y \in M$ äquivalent sind:

(α) $[x]_R \cap [y]_R \neq \emptyset$
(β) $x \sim y$
(γ) $[x]_R = [y]_R$

1. Schritt: $(\alpha) \Rightarrow (\beta)$: Sei $z \in [x]_R \cap [y]_R$, d.h. es gelten $x \sim z$ und $y \sim z$. Die Symmetrie und Transitivität von $\sim$ implizieren dann $x \sim y$.
2. Schritt: $(\beta) \Rightarrow (\gamma)$: Wegen der Symmetrie der Voraussetzung $x \sim y$ genügt es, $[y]_R \subseteq [x]_R$ zu zeigen. Sei $w \in [y]_R$, d.h. $y \sim w$. Unsere Voraussetzung $x \sim y$ ergibt zusammen mit $y \sim w$ und der Transitivität, daß $x \sim w$ gilt, d.h. $w \in [x]_R$.
3. Schritt: $(\gamma) \Rightarrow (\alpha)$: Wegen $x \in [x]_R$ folgt aus $[x]_R = [y]_R$, daß $[x]_R \cap [y]_R = [x]_R \neq \emptyset$ gilt.

Sind nun A und B Elemente von M/R, etwa $A = [x]_R$ und $B = [y]_R$ für Elemente x, y von M, so gilt wegen $(\alpha) \Leftrightarrow (\beta)$ entweder $A \cap B = \emptyset$ oder $A = B$. Das beweist (1).

Beispiel: Es sei nun wieder $p \in \mathbb{N}, p \geq 2$ und wir betrachten die Äquivalenzrelation auf $\mathbb{Z}$, die durch

$$m \sim n \Leftrightarrow n - m \text{ ist durch } p \text{ teilbar}$$

definiert ist, vgl. Bsp. 2 nach (1.24). Dann sind die Äquivalenzklassen von $\sim$ genau die Teilmengen von $\mathbb{Z}$, die aus den ganzen Zahlen bestehen, die bei Division (mit Rest) durch p, den gleichen Rest ergeben. Sie heißen deshalb auch die "Restklassen mod p". Es gibt genau p verschiedene solche Restklassen mod p, nämlich

$$\begin{array}{lcll}
[0]_p & = & \{np \mid n \in \mathbb{Z}\} & \text{Rest } 0 \text{ bei Division durch } p \\
[1]_p & = & \{np + 1 \mid n \in \mathbb{Z}\} & \text{Rest } 1 \text{ bei Division durch } p \\
 & \vdots & & \\
[p-1]_p & = & \{np + (p-1) \mid n \in \mathbb{Z}\} & \text{Rest } p-1 \text{ bei Division durch } p \\
 & = & \{np - 1 \mid n \in \mathbb{Z}\} &
\end{array}$$

Für diese Äquivalenzrelation ist folgende Bezeichnung gebräuchlich:

$$m \sim n \Leftrightarrow m \equiv n \pmod p$$

Kapitel 2: Gruppen und Körper

(2.1) Def. Eine Gruppe ist eine Menge, genannt G, und eine Abbildung ("innere Verknüpfung") von $G \times G$ nach G, hier bezeichnet als

$$\top : G \times G \to G, (a,b) =: ab,$$

so daß folgende Eigenschaften erfüllt sind:

(G_1) Für alle $a, b, c \in G$ gilt: $(a \top b) \top c = a \top (b \top c)$ (Assoziativgesetz)
Es existiert ein Element $e \in G$, so daß gilt:
(G_2) Für alle $a \in G$ gilt $e \top a = a$ (Existenz eines "Linksneutralen")
(G_3) Für alle $a \in G$ existiert ein $b \in G$, so daß gilt: $b \top a = e$. (Ex. eines "Linksinversen")

Eine Gruppe $(G, \top)$ heißt abelsch (oder kommutativ), falls für alle $a, b \in G$ gilt: $a \top b = b \top a$

Die Anzahl der Elemente von G heißt die Ordnung $|G| \in \mathbb{N} \cup \{\infty\}$ einer Gruppe $(G, \top)$. $(G, \top)$ heißt endliche Gruppe, falls $|G| < \infty$ und sonst unendliche Gruppe.

Bem.: Die abelschen Gruppen sind nach dem norwegischen Mathematiker Niels Henrik Abel (1802-1829) benannt.

Beispiele:

1) Einfachste Gruppe: $G = \{e\}$ (mit $e \top e := e$)
2) Zweiteinfachste Gruppe: $G = \{e, a\}$ mit $a \neq e$, wobei $a \top a := e$ (und $e \top a = a \top e = e$)
3) $G := \mathbb{Z}$, $\top := + \rightsquigarrow (\mathbb{Z}, +)$ "unendlich zyklische Gruppe"
   Ebenso: $(\mathbb{Q}, +), (\mathbb{R}, +)$. Hier stets $e := 0$.
   $(\mathbb{N}, +)$ ist keine Gruppe: Es müßte $e = 0$ gelten, aber es existiert zu keinem $n \in \mathbb{N}$ mit $n > 0$ ein additives Inverses. (d.h. (G_3) ist nicht erfüllt).
4) $(\mathbb{Q}_{>0}, \cdot), (\mathbb{R}_{>0}, \cdot), (\mathbb{Q} \setminus \{0\}, \cdot), (\mathbb{R} \setminus \{0\}, \cdot)$ sind Gruppen.
   Hier stets $e := 1$.
   $(\mathbb{Z} \setminus \{0\}, \cdot)$ ist keine Gruppe.

Alle Gruppen unter 1) - 4) sind abelsch, die unter 3) und 4) sind unendlich.

5) M Menge: Sei $S_M = \{f : M \to M \mid f \text{ bijektiv}\}$. S_M mit der Hintereinanderausführung $\circ$ als innere Verknüpfung ist Gruppe. $S_n := S_{\{1,\ldots,n\}}$

Zur Motivation: Der Begriff "Gruppe" ist eng verknüpft mit dem Bemühen, die Symmetrien von Figuren (in der Ebene oder im Raum) zu verstehen. Was soll es z.B. bedeuten, daß zwei verschiedene Figuren "die gleiche Symmetrie" besitzen?

Für eine gegebene Figur F betrachtet man die Menge G_F der Kongruenzabbildungen ("Isometrien") der Ebene bzw. des Raumes, die F in sich überführen. G_F mit der inneren Verknüpfung "Hintereinanderausführung" ist eine Gruppe (die "Isometriegruppe" von F)

1) F = gleichseitiges Dreieck $\Rightarrow |G_F| = 6$
   nämlich: G_F besteht aus den Drehungen um 120° um den Mittelpunkt und den Spiegelungen an den 3 Höhen.
2) F = regelmäßiges n-Eck $\Rightarrow |G_F| = 2n$
3) F = Kreis $\Rightarrow |G_F| = \infty$
4) Es gibt (außer den aus 1) und 2) abgeleiteten) im Wesentlichen nur 3 endliche Gruppen von Kongruenzabbildungen des Raumes, nämlich die Isometriegruppen von

a) Regulärem Tetraeder mit $|G| = 24$
b) Würfel (=Isometriegruppe des regulären Oktaeders) mit $|G| = 48$
c) Regulärem Dodekaeder (=Isometriegruppe des reg. Ikosaeders) mit $|G| = 120$
Alle diese Gruppen sind nicht abelsch.

Vorsicht: Es gibt "verschiedene" Gruppen mit der gleichen Ordnung.
Nach jahrzehntelanger Anstrengung vieler Mathematiker hat man einen genauen Überblick über <u>alle</u> endlichen Gruppen (Beweise in Gesamtlänge von über 3000 Seiten)

Beispiel: Die (endliche) zyklische Gruppe $(\mathbb{Z}_p, +)$ (wobei $p \in \mathbb{N}, p \geq 2$). Die Menge $\mathbb{Z}_p$ besteht aus den Restklassen $\pmod p$, d.h.

$$\mathbb{Z}_p = \{[0]_p, [1]_p, \ldots, [p-1]_p\} = \{[m]_p \mid m \in \mathbb{Z}\}$$

Seien $a, b \in \mathbb{Z}_p$. Wir wollen $a \top b$, hier geschrieben $a + b$, definieren. Dazu wählen wir ein $m \in a \subseteq \mathbb{Z}, n \in b \subseteq \mathbb{Z}$, und überlegen, daß die Restklasse $[m+n]_p$ nicht von diesen Wahlen abhängt: Jedes andere $\tilde{m} \in a$ kann nämlich als $\tilde{m} = m + lp$ für ein $l \in \mathbb{Z}$ geschrieben werden und ebenso jedes andere $\tilde{n} \in b$ als $\tilde{n} = n + jp$ für ein $j \in \mathbb{Z}$. Also:

$$\tilde{m} + \tilde{n} = m + n + (j + l)p, \text{ d.h. } [\tilde{m} + \tilde{n}]_p = [m + n]_p$$

Wir können also definieren

$$a + b = [m + n]_p$$

und das ist unabhängig von der Wahl von $m \in a$ und $n \in b$.

Es gilt dann mit $m \in a, n \in b$ und $l \in c \in \mathbb{Z}_p$:

$$\begin{array}{lcll} (a+b)+c & = & [(m+n)+l]_p = [m+(n+l)]_p = a+(b+c) & \text{Assoziativgesetz} \\ a+b & = & [(m+n)]_p = [n+m]_p = b+a & (\Rightarrow (\mathbb{Z}_p,+) \text{ ist abelsch}) \\ [0]_p + a & = & [(0+m)]_p = [m]_p = a & ([0]_p \text{ ist neutrales Element}) \\ [-m]_p + a & = & [-m+m]_p = [0]_p & ([-m]_p \text{ ist bzgl. + invers zu } [m]_p) \end{array}$$

$(\mathbb{Z}_p, +)$ heißt zyklisch, weil es ein Element $a \in \mathbb{Z}_p$ gibt, so daß

$$\mathbb{Z}_p = \{[0]_p, a, a+a, \ldots, \underbrace{a + \ldots + a}_{(p-1)-\text{mal}}\}$$

während $\underbrace{a + \ldots + a}_{p-\text{mal}} = [0]_p$ gilt, nämlich $a = [1]_p$.

Die Gruppe $(\mathbb{Z}_p, +)$ verhält sich genau gleich wie die ("ist isomorph zur") Gruppe der Drehungen der Ebene (um einen festen Punkt) um die Winkel $\frac{m}{p} \cdot 360°$, $m \in \{0, \ldots, p-1\}$.

(2.2) Fakt: Sei $(G, \top)$ eine Gruppe. Dann gilt:

(i) Es gibt nur ein Element in G, für das (G_2) gilt (nämlich e).
Für alle $a \in G$ gilt: $a \top e = a$.
(ii) Zu jedem $a \in G$ existiert nur ein $b \in G$, für das $b \top a = e$ gilt.
Für dieses b gilt auch: $a \top b = e$.

Bezeichnung: Das Element $b \in G$ mit $b \top a = a \top b = e$ wird mit a^{-1} (bzw. mit $-a$, falls $\top = +$) bezeichnet.

Beweis: Zeige zunächst:

$(*)$ $\qquad a, b \in G$ und $b \top a = e \Rightarrow a \top b = e \qquad (\Rightarrow$ 2. Behauptung in (ii))

Denn: Nach (G_3) existiert zu $b \in G$ ein $c \in G$ mit $c \top b = e$. Es gilt:

$$a \top b \overset{(G_2)}{=} e \top (a \top b) = (c \top b) \top (a \top b) \overset{!}{=} c \top (\underbrace{(b \top a)}_{=e} \top b) \overset{(G_2)}{=} c \top b = e$$

zu "!": $(c \top b) \top \underbrace{(a \top b)}_{=:d} \overset{(G_1)}{=} c \top (b \top \underbrace{(a \top b)}_{=d}) \overset{(G_1)}{=} c \top ((b \top a) \top b)$ Beweis von (i): Sei $\tilde{e} \in G$ ein Element, so daß (G_2) gilt. Zu zeigen: $e = \tilde{e}$. Es gilt:

$$e \overset{(G_1)\text{für } \tilde{e}}{=} \tilde{e} \top e \overset{(*)}{=} e \top \tilde{e} \overset{(G_2)\text{für } e}{=} \tilde{e}$$

Sei $a \in G$. Wir zeigen $a \top e = a$. Nach (G_3) existiert ein $b \in G$ mit $b \top a = e$ und nach $(*)$ folgt $a \top b = e$. Also

$$a \top e = a \top (b \top a) \overset{(G_1)}{=} (a \top b) \top a = e \top a \overset{(G_2)}{=} a$$

Beweis von (ii): Seien $a, b, \tilde{b} \in G$ und es gelte: $b \top a = e = \tilde{b} \top a$. Zu zeigen: $b = \tilde{b}$

$$\text{Es gilt:} \qquad \tilde{b} \overset{(i)}{=} \tilde{b} \top e \overset{(*)}{=} \tilde{b} \top (a \top b) \overset{(G_1)}{=} (\tilde{b} \top a) \top b = e \top b \overset{(G_2)}{=} b$$

(2.3) Fakt. Sei $(G, \top)$ Gruppe, $a, b \in G$. Dann gilt:

(i) $(a^{-1})^{-1} = a$
(ii) $(a \top b)^{-1} = b^{-1} \top a^{-1}$

Beweis:

(i) Zu zeigen ist: a ist das Inverse von a^{-1}, d.h. $a \top a^{-1} = e$
Nach Definition gilt $a^{-1} \top a = e$, und nach (2.2) (ii) folgt daraus $a \top a^{-1} = e$.
(ii) Zu zeigen ist: $(b^{-1} \top a^{-1}) \top (a \top b) = e$. Wir rechnen:

$$\left(b^{-1} \top a^{-1}\right) \top (a \top b) \overset{(G_1)}{=} b^{-1} \top \left(a^{-1} \top a\right) \top b = b^{-1} \top e \top b \overset{(G_2)}{=} b^{-1} \top b = e$$

(2.4) Def.: Ein Körper ist eine Menge, genannt K, und zwei innere Verknüpfungen $+ : K \times K \to K$, $(a, b) \mapsto a + b$, und $\cdot : K \times K \to K$, $(a, b) \mapsto a \cdot b =: ab$, so daß gilt

(i) $(K, +)$ ist eine abelsche Gruppe mit neutralem Element $e =: 0$
(ii) $(K \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe mit neutralem Element $e =: 1$ (speziell $1 \neq 0$!)
(iii) $\forall a, b, c \in K : (a + b)c = (ac) + (bc)$ (Distributivgesetz)

Schreibweisen: $a + (-b) =: a - b$, $(ab) + c =: ab + c$ “Punkt vor Strich”

$(2.3)(\mathrm{i}) \Rightarrow -(-a) = a$
$(2.3)(\mathrm{ii}) \Rightarrow -(a+b) = (-b) + (-a) = -a - b$

Beispiel: $(\mathbb{R}, +, \cdot), (\mathbb{Q}, +, \cdot), (\mathbb{C}, +, \cdot)$ sind Körper.
$(\mathbb{Z}, +, \cdot)$ ist kein Körper (z.B. besitzt 2 in $\mathbb{Z}$ kein multiplikatives Inverses).

Bem.: Verzichtet man in (2.4)(ii) auf die Bedingung der Existenz von multipl. Inversen, so nennt man $(K, +, \cdot)$ einen kommutativen Ring mit 1. Verzichtet man zusätzlich auf die Kommutativität von $(K \setminus \{0\}, \cdot)$ und ergänzt (iii) durch $a(b + c) = ab + ac$, so nennt man $(K, +, \cdot)$ einen Ring mit 1. (Es wird die Existenz von $1 \in K \setminus \{0\}$ mit $1 \cdot a = a \cdot 1 = a$ für alle $a \in K \setminus \{0\}$ gefordert)

Beispiel: $(\mathbb{Z}, +, \cdot)$ ist kommutativer Ring mit 1.

(2.5)Rechenregeln: Sei $(K, +, \cdot)$ Körper (es genügt: Ring), $a, b, c \in K$. Dann gilt:

(i) $0 \cdot a = 0$
(ii) $a \cdot (-b) = (-a) \cdot b = -(ab)(=: -ab)$
(iii) $a(bc) = (ab)c$

Beweis:

(i) Es gilt: $0 \cdot a + a = 0 \cdot a + 1 \cdot a \overset{(2.4)(\mathrm{iii})}{=} (0+1)a = 1 \cdot a = a \quad | -a \Rightarrow 0 \cdot a = 0$
(ii) $(-a)b + ab \overset{(2.4)(\mathrm{iii})}{=} ((-a) + a)b = 0 \cdot b \overset{(\mathrm{i})}{=} 0$
(iii) Nach (2.4)(ii), falls $\{a, b, c\} \subseteq K \setminus \{0\}$. Sonst: (i) $\Rightarrow a(bc) = 0 = (ab)c$

Schreibweise: Ist $a \in K \setminus \{0\}$, $b \in K$, so schreibt man auch $a^{-1} =: \frac{1}{a}$ und $a^{-1}b = ba^{-1} =: \frac{b}{a}$.

Der Körper der komplexen Zahlen (Cardano 1501-1576, C.F. Gauß 1777-1855)

$$\begin{aligned} K = \mathbb{R}^2, \quad + \quad &: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2, (a,b) + (c,d) := (a+c, b+d) \\ \cdot \quad &: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2, (a,b) \cdot (c,d) := (ac - bd, ad + bc) \end{aligned}$$

Eigenschaften:

1) $(\mathbb{R}^2, +)$ ist abelsche Gruppe mit neutralem Element $e = (0, 0)$.
2) $(1, 0) \cdot (a, b) = (a, b)$, d.h. $(1, 0)$ ist neutrales Element bezüglich $\cdot$.
3) $\cdot$ ist kommutativ: $(a, b) \cdot (a', b') = (aa' - bb', ab' + a'b) = (a', b') \cdot (a, b)$
4) Ist $(a, b) \in \mathbb{R}^2 \setminus \{(0, 0)\}$, so gilt: $\left(\frac{a}{a^2+b^2}, \frac{-b}{a^2,b^2}\right) \cdot (a, b) = (1, 0)$,

d.h. $\left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right) = \frac{1}{a^2+b^2}(a, -b)$ ist multiplikatives Inverses zu (a, b).

5) Für alle $(a, b), (a', b'), (a'', b'')$ in $\mathbb{R}^2$ gilt

$$((a, b) \cdot (a', b')) \cdot (a'', b'') = (a, b) \cdot ((a', b') \cdot (a'', b''))$$

d.h. $\cdot$ ist assoziativ (Übung)

2)–5) $\Rightarrow$ $(\mathbb{R}^2 \setminus \{(0,0)\}, \cdot)$ ist abelsche Gruppe mit neutralem Element $(1,0)$

6) Distributivgesetz:

$$\begin{array}{rcl}
(a,b)\cdot((a',b')+(a'',b'')) & = & (a,b)\cdot(a'+a'',b'+b'') \\
 & \overset{\text{Def. von }\cdot}{=} & (a(a'+a'')-b(b'+b''),a(b'+b'')+b(a'+a'')) \\
(a,b)\cdot(a',b')+(a,b)\cdot(a'',b'') & = & (aa'-bb',ab'+ba')+(aa''-bb'',ab''+a''b) \\
 & = & (a(a'+a'')-b(b'+b''),a(b'+b'')+b(a'+a''))
\end{array}$$

Aus 1)–6) folgt, daß $(\mathbb{R}^2,+,\cdot)$ ein Körper ist.

Bem.: Es gilt

$$\begin{array}{rcll}
(0,1)\cdot(0,1) & = & (-1,0)=-(1,0) & \\
(0,1)\cdot(b,0) & = & (0,b) & \forall b\in\mathbb{R} \\
(a,0)\cdot(b,0) & = & (ab,0) & \forall a,b\in\mathbb{R} \\
(a,0)\cdot(0,b) & = & (0,ab) & \forall a,b\in\mathbb{R}
\end{array}$$

Für alle reellen Zahlen a identifiziert man $(a,0)$ mit a und schreibt abkürzend $(0,1) =: i$, $(0,b)=(0,1)\cdot(b,0) =: ib =: bi$, $(a,b)=(a,0)+(0,b) =: a+ib$

Speziell: $(0,0)=0$, $(1,0)=1$, $(0,1)=i$ mit:

(2.6) $i^2=-1$

Mit diesen Bezeichnungen gilt für $a,b,a',b'\in\mathbb{R}$:

(2.7) $(a+ib)+(a'+ib')=a+a'+i(b+b')$

(2.8) $(a+ib)\cdot(a'+ib')=aa'-bb'+i(ab'+ba')$

$\mathbb{C} := \{a+ib \mid a\in\mathbb{R}, b\in\mathbb{R}\}$. Die Veranschaulichung von komplexen Zahlen in der Koordinatenebene $\mathbb{R}^2$ stammt von C.F. Gauß (Gaußsche Zahlenebene).

Ist $z=a+ib\in\mathbb{C}$ mit $a\in\mathbb{R}$, $b\in\mathbb{R}$, so heißen

$$\begin{array}{c}
a \text{ der Realteil von } z, a =: \mathrm{Re}(z) \\
b \text{ der Imaginärteil von } z, b =: \mathrm{Im}(z)
\end{array}$$

Es gilt $z\neq 0 \Leftrightarrow (a,b)\neq(0,0)$ $\quad(\Leftrightarrow a^2+b^2>0)$

4) $\Rightarrow$ Ist $z=a+ib\neq 0$, so ist das multiplikative Inverse z^{-1} von z:

(2.9) $z^{-1}=(a+ib)^{-1}=\frac{1}{a^2+b^2}(a-ib)$

Die Abbildung $z=a+ib\in\mathbb{C}\to\overline{z}=a-ib\in\mathbb{C}$ heißt Konjugation.

Es gilt: $\overline{\overline{z}}=z$, $\overline{z+w}=\overline{z}+\overline{w}$ und $\overline{z\cdot w}=\overline{z}\cdot\overline{w}$

Ist $z = a + ib \in \mathbb{C}$, $a, b \in \mathbb{R}$, eine komplexe Zahl, so heißt $|z| = +\sqrt{a^2 + b^2} \in \mathbb{R}_{\geq 0}$ der Betrag von z. Offensichtlich gilt $z\overline{z} = |z|^2$.

Damit gilt für $z \neq 0$: $z^{-1} = \frac{\overline{z}}{|z|^2}$. Denn: $z^{-1}z = \frac{z\overline{z}}{|z|^2} = 1$.

Existenz von endlichen Körpern

Für $p \in \mathbb{N}$, $p \geq 2$, definieren wir eine Multiplikation auf

$$\mathbb{Z}_p = \{[0]_p, \ldots, [p-1]_p\} = \{[m]_p \mid m \in \mathbb{Z}\}$$

durch: $a = [m]_p, b = [n]_p \to ab = [mn]_p$ unabhängig von der Wahl von $m \in a$, $n \in b$, vgl. Blatt 3, Aufgabe 3.

Es ist leicht nachzuprüfen, daß $(\mathbb{Z}_p, +, \cdot)$ ein kommutativer Ring mit $1 (= [1]_p)$ ist.

(2.10) Satz. *Ist $p \in \mathbb{N}$ Primzahl, so ist $(\mathbb{Z}_p, +, \cdot)$ ein Körper.*

Beweis: Zu zeigen ist die Existenz eines multiplikativen Inversen für jedes $a \in \mathbb{Z}_p \setminus \{0\}$.
$a \in \mathbb{Z}_p \setminus \{0\} \Rightarrow a = [m]_p$ für ein $m \in \{1, \ldots, p-1\}$. Wir zeigen: Die Elemente $[1]_p \cdot [m]_p, [2]_p \cdot [m]_p, \ldots, [p-1]_p \cdot [m]_p$ in $\mathbb{Z}_p$ sind alle ungleich 0.
Denn $[j]_p[m]_p = [0]_p$ ist äquivalent dazu, daß p die Zahl jm teilt. Da p Primzahl ist, teilt p dann j oder m (betrachte die Primfaktorzerlegungen von j, m und jm!), d.h. $[j]_p = 0$ oder $[m]_p = 0$. Dann sind die Elemente $[1]_p \cdot [m]_p, \ldots, [p-1]_p \cdot [m]_p$ alle verschieden, denn aus $[i]_p[m]_p = [j]_p[m]_p$ folgt $[j-i]_p[m]_p = 0$. Also existiert ein $i \in \{1, \ldots, p-1\}$, so daß $[i]_p[m]_p = [1]_p$ gilt, d.h. $[i]_p$ ist multiplikatives Inverses von $a = [m]_p$.

Speziell: $(\mathbb{Z}_2, +, \cdot)$ Körper mit 2 Elementen, $\mathbb{Z}_2 = \{0, 1\}$.

Bem.: Lineare Gleichungssysteme mit Koeffizienten a_{ij} in einem Körper K und "Unbekannten" x_i in K können mit dem Gaußschen Eliminationsverfahren (1.13) gelöst werden.

Kapitel 3: Vektorräume

(3.1) Def.: Ein Vektorraum über einem Körper K ist eine abelsche Gruppe $(V, +)$ zusammen mit einer Abbildung $K \times V \to V$, $(\alpha, v) \in K \times V \mapsto \alpha v$, so daß für alle $\alpha, \beta \in K$ und alle $v, w \in V$ gilt:

(i) $\alpha(\beta v) = (\alpha\beta)v \quad (=: \alpha\beta v)$
(ii) $(\alpha + \beta)v = (\alpha v) + (\beta v) \quad (=: \alpha v + \beta v)$
(iii) $\alpha(v + w) = (\alpha v) + (\alpha w) \quad (=: \alpha v + \alpha w)$
(iv) $1v = v$ (, wobei 1 das Einselement von K bezeichnet!)

Sprechweisen und Bezeichnungen:
"Punkt vor Strich" $\alpha v + \beta w := (\alpha v) + (\beta w)$, außerdem: $v - w := v + (-w)$
Vektorraum über einem Körper K = K-Vektorraum

Das neutrale Element von $(V, +)$ wird (zunächst) mit $\underline{0}$ bezeichnet, zur Unterscheidung vom neutralen Element 0 von $(K, +)$.

Elemente von V werden "Vektoren" (des Vektorraums V) genannt, Elemente des Körpers werden auch "Skalare" genannt.

Beispiele:

1) $K := \mathbb{R}$ und $V := \mathbb{R}^n$ mit den vor (1.15) eingeführten Operationen
$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_n + y_n)$$
$$r(x_1, \ldots, x_n) := (rx_1, \ldots, rx_n).$$
In diesem Fall $\underline{0} = (0, \ldots, 0)$
2) Allgemeiner: Sei K beliebiger Körper, $n \in \mathbb{N}$, $n \geq 1$. Dann können wir genau wie in Beispiel 1) die Menge
$$K^n = \{(x_1, \ldots, x_n) \mid x_1 \in K, \ldots, x_n \in K\}$$
zu einem Vektorraum über dem Körper K machen.
3) Sei I ein <u>homogenes</u> lineares Gleichungssystem mit m Gleichungen und n Unbekannten $x_1, \ldots, x_n$ und Koeffizienten $a_{11}, \ldots, a_{mn}$ in K. Dann bildet die Lösungsmenge
$$L_I = \{x = (x_1, \ldots, x_n) \in K^n \mid x \text{ löst } I\}$$
mit den wie in 1) definierten Operationen einen K-Vektorraum, vgl. (1.15).
4) Sei $K := \mathbb{R}$, $M \neq \emptyset$ eine Menge und $V = \{f \mid f : M \to \mathbb{R}\}$.

Für $r \in \mathbb{R}$, $f, g \in V$ definieren wir $f + g \in V$ und $rf \in V$ durch

(a) $\forall x \in M : (f + g)(x) := f(x) + g(x)$
(b) $\forall x \in M : (rf)(x) := rf(x)$
Mit diesen Operationen ist V ein $\mathbb{R}$-Vektorraum (wobei $\underline{0} \in V$ die Abbildung ist, die jedes $x \in M$ auf $0 \in \mathbb{R}$ abbildet).

Wir beweisen exemplarisch, daß (3.1)(iii) gilt:
Seien $f, g \in V$, $r \in \mathbb{R}$. Zu zeigen ist:

$$(*) \qquad r(f + g) = rf + rg$$

Auf beiden Seiten von $(*)$ stehen Abbildungen von M nach $\mathbb{R}$. Wir haben also zu zeigen, daß für alle $x \in M$ gilt:
$$(r(f + g))(x) = (rf + rg)(x)$$
Wendet man auf beide Seiten (a) und (b) an (auf der linken Seite zunächst (b) dann (a)), so erhält man für die linke Seite
$$(r(f + g)(x)) \overset{\text{(b)}}{=} r((f + g)(x)) \overset{\text{(a)}}{=} r(f(x) + g(x))$$
und für die rechte Seite
$$(rf + rg)(x) \overset{\text{(a)}}{=} (rf)(x) + (rg)(x) \overset{\text{(b)}}{=} rf(x) + rg(x)$$

Schließlich gilt $r(f(x) + g(x)) = rf(x) + rg(x)$ aufgrund des Distributivgesetzes für die reellen Zahlen.

(3.2) Rechenregeln. Für alle $\alpha \in K$, $v \in V$ gilt

(i) $0v = \underline{0}$, $\alpha\underline{0} = \underline{0}$
(ii) $\alpha v = \underline{0} \Rightarrow \alpha = 0$ oder $v = \underline{0}$
(iii) $(-\alpha)v = \alpha(-v) = -(\alpha v) \quad (=: -\alpha v)$

Bew.:

(i) $0v = 0v + (v - v) = (0v + 1v) - v = (0+1)v - v = 1 \cdot v - v = v - v = \underline{0}$
$\alpha\underline{0} = \alpha\underline{0} + (\alpha\underline{0} - \alpha\underline{0}) = \alpha(\underline{0} + \underline{0}) - \alpha\underline{0} = \alpha\underline{0} - \alpha\underline{0} = \underline{0}$
(ii) Sei $\alpha v = \underline{0}$ und $\alpha \neq 0 \overset{\text{(i)}}{\Rightarrow} \alpha^{-1}(\alpha v) = \underline{0} \Rightarrow (\alpha^{-1}\alpha)v = \underline{0} \Rightarrow 1v = \underline{0} \Rightarrow v = \underline{0}$
(iii) $(-\alpha)v + \alpha v = ((-\alpha) + \alpha)v = 0v \overset{\text{(i)}}{=} \underline{0} \Rightarrow (-\alpha)v = -(\alpha v)$
$\alpha(-v) + \alpha v = \alpha((-v) + v) = \alpha\underline{0} \overset{\text{(i)}}{=} \underline{0} \Rightarrow \alpha(-v) = -(\alpha v)$

(3.3) Def.: Sei V K-Vektorraum und $U \subseteq V$. U heißt Unter(vektor)raum von V, falls

(i) $U \neq \emptyset$
(ii) $v_1, v_2 \in U \Rightarrow v_1 + v_2 \in U$
(iii) $\alpha \in K$, $v \in U \Rightarrow \alpha v \in U$

(3.4) Folgerung: Sei V K-Vektorraum, $U \subseteq V$ Unterraum. Dann ist U (mit den von V auf U eingeschränkten Strukturen) ein K-Vektorraum.

Bew.: Sei U Unterraum.

(i) $\Rightarrow +$ ist innere Verknüpfung auf $U (\Rightarrow (G_1))$.
(ii) $\Rightarrow$ die Multiplikation mit Skalaren ist eine Abbildung $K \times U \to U$.

Wir zeigen: $(U, +)$ ist abelsche Gruppe: $v \in U \overset{\text{(ii)}}{\Rightarrow} (-1)v \in U \overset{\text{(3.2)(iii)}}{\Rightarrow} -v \in U$

$U \neq \emptyset \Rightarrow \exists v \in U \Rightarrow v$ und $-v \in U \overset{\text{(ii)}}{\Rightarrow} v + (-v) = \underline{0} \in V$
Daraus folgen $(G_2), (G_3)$ für $(U, +)$, (G_1) für $(U, +)$ folgt aus (G_1) für $(V, +)$. Die anderen Eigenschaften sind offensichtlich erfüllt.

Beispiele:

1) $\{0\}$ und V sind Unterräume von V
2) Ist $m < n$, so ist $K^m \times \{\underbrace{(0, \ldots, 0}_{(n-m)-\text{mal}})\} = \{(x_1, \ldots, x_n) \in K^n \mid x_{m+1} = \ldots = x_n = 0\}$ Unterraum von K^n
3) Ist I homogenes lineares Gleichungssystem für n Unbekannte mit Koeffizienten a_{ij} in K, so ist L_I Unterraum von K^n, vgl. (1.15).
4) $V = \{f \mid f : M \to \mathbb{R}\}$, vgl. Bsp. 4) nach (3.1). Zu festem $a \in M$ sei $U = \{f \in V \mid f(a) = 0\}$. Dann ist U Unterraum von V.

(3.5) Fakt. Ist V ein K-Vektorraum und $\mathcal{U}$ eine Menge von Unterräumen von K, so ist $W := \bigcap\limits_{U \in \mathcal{U}} U$ Unterraum von V.

Bew.: $\forall U \in \mathcal{U} : \underline{0} \in U \Rightarrow \underline{0} \in W$
$v_1, v_2 \in W \Leftrightarrow \forall U \in \mathcal{U} : v_1, v_2 \in U \overset{(3.3)(\text{ii})}{\Rightarrow} \forall U \in \mathcal{U} : v_1 + v_2 \in U \Leftrightarrow v_1 + v_2 \in W$
Ebenso: $\alpha \in K$, $v \in W \Rightarrow \alpha v \in W$.

(3.6) Def.: Sei V K-Vektorraum, $M \subseteq V$. $\mathcal{U}_M := \{U \mid U$ Unterraum von V mit $M \subseteq U\}$

Dann heißt $\operatorname{span}(M) := \bigcap\limits_{U \in \mathcal{U}_M} U$ der von M aufgespannte (oder erzeugte) Unterraum. M heißt Erzeugendensystem von V, falls $\operatorname{span}(M) = V$ gilt. V heißt endlich erzeugt, falls es eine endliche Menge $M \subseteq V$ gibt mit $\operatorname{span}(M) = V$.

Bem.: (3.5) $\Rightarrow \operatorname{span}(M)$ ist Unterraum von V, der (bzgl. der Inklusion) kleinste, M enthaltende Unterraum.

Bsp.:

1) $M = \emptyset$ oder $M = \{\underline{0}\} \Rightarrow \operatorname{span}(M) = \{\underline{0}\}$
2) Ist M Unterraum von V, so $\operatorname{span}(M) = M$

(3.7) Satz. *Sei V K-Vektorraum, $\emptyset \neq M \subseteq V$. Dann gilt:*

$$\operatorname{span}(M) = \{v \in V \mid \exists k \in \mathbb{N}, \alpha_1, \ldots, \alpha_k \in K, v_1, \ldots, v_k \in M : v = \alpha_1 v_1 + \ldots + \alpha_k v_k\}$$

Bez.: $\alpha_1 v_1 + \ldots + \alpha_k v_k$ heißt eine Linearkombination der Vektoren $v_1, \ldots, v_k$.
abgekürzt: $\alpha_1 v_1 + \alpha_s v_2 + \ldots + \alpha_k v_k = \sum\limits_{i=1}^{k} \alpha_i v_i$

Bew.: Wir bezeichnen die Menge auf der rechten Seite des Gleichheitszeichens in (3.7) durch U_0.
1) Zeige U_0 ist Unterraum von V. Wegen $M \neq \emptyset$ gilt $U_0 \neq \emptyset$. Seien $v, w \in U_0$, $v = \alpha_1 v_1 + \ldots + \alpha_k v_k$, $w = \beta_1 w_1 + \ldots + \beta_j w_j$ mit $\alpha_1, \ldots, \beta_j \in K, v_1, \ldots, w_j \in M$. Dann gilt $v + w = \alpha_1 v_1 + \ldots + \alpha_k v_k + \beta_1 w_1 + \ldots + \beta_j w_j \in U_0$.

Sei $\alpha \in K, v = \sum\limits_{i=1}^{k} \alpha_i v_i \in U_0$ mit $v_i \in M$. Dann gilt

$$\alpha v = \alpha(\alpha_1 v_1 + \ldots + \alpha_k v_k) = \alpha(\alpha_1 v_1) + \ldots + \alpha(\alpha_k v_k) = (\alpha\alpha_1) v_1 + \ldots + (\alpha\alpha_k) v_k \in U_0.$$

(Kürzer: $\alpha \sum\limits_{i=1}^{k} \alpha_i v_i = \sum\limits_{i=1}^{k} \alpha\alpha_i v_i$).
Also ist U_0 ein M enthaltender Unterraum. Nach Definition (3.6) folgt $\operatorname{span}(M) \subseteq U_0$.
2) Zeige $U_0 \subseteq \operatorname{span}(M)$. Seien $v_1, \ldots, v_k \in M$, $\alpha_1, \ldots, \alpha_k \in K$. Da $\operatorname{span}(M)$ Unterraum ist, folgt nach (3.3)(iii): $\alpha_1 v_1 \in \operatorname{span}(M), \ldots, \alpha_k v_k \in \operatorname{span}(M)$, und nach (3.3)(ii) $\alpha_1 v_1 + \alpha_2 v_2 \in$

$\mathrm{span}(M)$, $(\alpha_1 v_1 + \alpha_2 v_2) + \alpha_3 v_3 \in \mathrm{span}(M)$, u.s.w. bis

$$\alpha_1 v_1 + \ldots + \alpha_k v_k \in \mathrm{span}(M).$$

Also $U_0 \subseteq \mathrm{span}(M)$.

Bsp.: Betrachte K^n und $e_1 := (1, 0, \ldots, 0) \in K^n$, $e_2 := (0, 1, 0, \ldots, 0) \in K^n$, ..., $e_n = (0, \ldots, 0, 1) \in K^n$. Dann ist $M = \{e_1, \ldots, e_n\}$ Erzeugendensystem für K^n. Es genügt, $K^n \subseteq \mathrm{span}(M)$ zu zeigen. Ist $x = (x_1, \ldots, x_n) \in K^n$, so gilt $x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n \in \mathrm{span}(M)$.

(3.8) Def.: Die Vektoren $v_1, \ldots, v_k$ heißen <u>linear unabhängig</u>, falls gilt: Ist $\alpha_1 \in K, \ldots, \alpha_k \in K$ und $\alpha_1 v_1 + \ldots + \alpha_k v_k = \underline{0}$, so folgt $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$. Sonst heißen die Vektoren $v_1, \ldots, v_k$ linear abhängig. Eine Teilmenge M von V heißt linear unabhängig, falls gilt: Ist $k \in \mathbb{N}$ und sind $v_1, \ldots, v_k$ verschiedene Vektoren in M, so sind die $v_1, \ldots, v_k$ linear unabhängig. Sonst heißt M linear abhängig.

Bem.: $v_1, \ldots, v_k$ linear abhängig $\Leftrightarrow \exists \alpha_1, \ldots, \alpha_k \in K$ <u>nicht</u> <u>alle</u> $= 0$:

$$\alpha_1 v_1 + \ldots + \alpha_k v_k = \underline{0}$$

Bsp.:

1) $\underline{0} \in M \Rightarrow M$ linear abhängig. Denn: $1 \cdot \underline{0} = \underline{0}$.
2) $M = \{v\}$ linear unabhängig $\Leftrightarrow v \neq 0$. Denn: $\alpha v = \underline{0}, v \neq \underline{0} \overset{(3.2)(ii)}{\Rightarrow} \alpha = 0$.
3) Sei $M \subseteq M' \subseteq V$. Dann: M linear abhängig $\Rightarrow$ M' linear abhängig
   M' linear unabhängig $\Rightarrow$ M linear unabhängig
4) $e_1, \ldots, e_n$ sind linear unabhängig in K^n. Denn:
   $x_1 e_1 + \ldots + x_n e_n = (x_1, \ldots, x_n) = \underline{0} \Leftrightarrow x_1 = \ldots = x_n = 0$.
5) $V = \{f \mid f : \mathbb{R} \to \mathbb{R}\}$. Die Funktion $x \to \sin x$ und $x \to \cos x$ sind linear unabhängig. Seien $r, s \in \mathbb{R}$ und $r \sin x + s \cos x = 0$ <u>für alle</u> $x \in \mathbb{R}$. Dann gilt das speziell für $x = 0$ und $x = \frac{\pi}{2}$, d.h.

$$\begin{aligned} r \underbrace{\sin 0}_{=0} + s \underbrace{\cos 0}_{=1} = 0 &\quad \Rightarrow \quad s = 0 \\ r \underbrace{\sin\left(\frac{\pi}{2}\right)}_{=1} + s \underbrace{\cos\left(\frac{\pi}{2}\right)}_{=0} = 0 &\quad \Rightarrow \quad r = 0. \end{aligned}$$

(3.9) Fakt. Die Vektoren $v_1, \ldots, v_k$ seien linear unabhängig. Dann gilt:

$$\alpha_1 v_1 + \ldots + \alpha_k v_k = \beta_1 v_1 + \ldots + \beta_k v_k \Rightarrow \alpha_1 = \beta_1, \ldots, \alpha_k = \beta_k$$

$$\begin{aligned} \text{Bew.: Voraussetzung} &\Rightarrow (\alpha_1 - \beta_1) v_1 + (\alpha_2 - \beta_2) v_2 + \ldots + (\alpha_k - \beta_k) v_k = \underline{0} \\ &\overset{\text{l. unabh.}}{\Rightarrow} \alpha_1 - \beta_1 = 0, \ldots, \alpha_k - \beta_k = 0 \Rightarrow \alpha_1 = \beta_1, \ldots, \alpha_k = \beta_k. \end{aligned}$$

(3.10) Def.: Eine Teilmenge M von V heißt <u>Basis von</u> V, falls M linear unabhängig und ein Erzeugendensystem von V ($\Leftrightarrow \mathrm{span}(M) = V$) ist.

Bsp.: $\{e_1, \dots, e_n\} \subseteq K^n$ ist Basis von K^n, die "Standardbasis von K^n".

Eine Beispielrechnung zur linearen Unabhängigkeit von Vektoren:

Wir fragen, ob die Vektoren $(1,0,1),(0,1,1),(1,1,-1)$ im $\mathbb{R}$-Vektorraum $\mathbb{R}^3$ linear unabhängig sind. Dazu nehmen wir an, daß für die reellen Zahlen r, s und t gilt:

$$r(1,0,1) + s(0,1,1) + t(1,1,-1) = (0,0,0)$$

Die Vektoren sind linear unabhängig, wenn wir aus dieser Gleichung folgern können, daß $r = s = t = 0$ gilt, vgl. Def. (3.8). Die Gleichung ist äquivalent zum homogenen linearen Gleichungssystem

$$\begin{array}{rcll} r + t & = & 0 & \mid \; - \\ s + t & = & 0 & \mid \\ r + s - t & = & 0 & \hookleftarrow\!\mid \end{array}$$

Wir lösen es nach dem Gaußschen Eliminationsverfahren:

$$\begin{array}{rcll} r + t & = & 0 & \\ s + t & = & 0 & \mid \; - \\ s - 2t & = & 0 & \hookleftarrow\!\mid \end{array}$$

$$\begin{array}{rcl} r + t & = & 0 \\ s + t & = & 0 \\ -3t & = & 0 \end{array}$$

Aus der letzten Gleichung folgt $t = 0$, damit aus der vorletzten $s = 0$ und aus der ersten $r = 0$.

Vorsicht: Bezeichnet man im Körper $(\mathbb{Z}_3, +, \cdot)$ wie in Körpern üblich das 0-Element mit 0, das 1-Element mit 1 und das additive Inverse von 1 mit -1, so kann man die Vektoren $(1,0,1),(0,1,1),(1,1,-1)$ auch als Element des $\mathbb{Z}_3$-Vektorraums $(\mathbb{Z}_3)^3$ auffassen. In diesem sind sie linear abhängig, denn

$$(1,0,1) + (0,1,1) - (1,1,-1) = (0,0,0) \text{ in } (\mathbb{Z}_3)^3!$$

(3.11) Lemma. (i) $v \in \operatorname{span}(M) \;\Rightarrow\; \operatorname{span}(M \cup \{v\}) = \operatorname{span}(M)$
(ii) $v \in \operatorname{span}(M) \setminus M \;\Rightarrow\; M \cup \{v\}$ linear abhängig

Beweis von (i): $\operatorname{span}(M) \subseteq \operatorname{span}(M \cup \{v\})$ ist klar.

Sei $w \in \operatorname{span}(M \cup \{v\})$, $w = \alpha v + \sum_{j=1}^{l} \beta_j w_j$ mit $w_j \in M$.

$v \in \operatorname{span}(M) \Rightarrow v = \sum_{i=1}^{k} \alpha_i v_i \Rightarrow w = \sum_{i=1}^{k} (\alpha\alpha_i) v_i + \sum_{j=1}^{l} \beta_j w_j \in \operatorname{span}(M)$.

Beweis von (ii): (3.7) $\Rightarrow$ Es existiert $k \in \mathbb{N}$, $\alpha_1, \dots, \alpha_k \in K$, $v_1, \dots, v_k \in M$: $v \overset{(*)}{=} \sum_{i=1}^{k} \alpha_i v_i$.

Wir können (mittels der Vektorraumaxiome) etwaige Summanden $\alpha_i v_i$ und $\alpha_j v_j$ mit $v_i = v_j$ zu $(\alpha_i + \alpha_j)v_i$ zusammenfassen und dann annehmen, daß alle $v_1, \ldots, v_k$ verschieden sind. $v \notin M \Rightarrow v, v_1, \ldots, v_k$ sind alle verschieden.

$$(*) \Rightarrow 1 \cdot v - \sum_{i=1}^{k} \alpha_i v_i = 0 \Rightarrow M \cup \{v\} \text{ linear abhängig.}$$

(3.12) Satz. *Folgende Aussage über eine Teilmenge M von V sind äquivalent:*

(i) *M ist Basis*
(ii) *M ist linear unabhängig und jede echte Obermenge $M' \subseteq V$ von M ist linear abhängig. ("M ist eine maximale linear unabhängige Teilmenge von V.")*
(iii) *Es gilt* $\operatorname{span}(M) = V$ *und für jede echte Teilmenge M'' von M gilt:* $\operatorname{span}(M'') \neq V$. *("$M$ ist ein minimales Erzeugendensystem von V.")*

Bew.: $V = \{0\} \Leftrightarrow M = \emptyset$ ist Basis von V. Dann ist (3.12) klar. Es genügt also, den Fall $M \neq \emptyset$ zu betrachten.

1) (i) $\Rightarrow$ (ii): Es ist zu zeigen, daß jede echte Obermenge $M' \subseteq V$ von M linear abhängig ist. Sei $v \in M' \backslash M$. Wegen $\operatorname{span}(M) = V$ (M Basis!) gilt $v \in \operatorname{span}(M) \backslash M$. Nach (3.11)(ii) ist $M \cup \{v\}$ und damit auch $M' \supseteq M \cup \{v\}$ linear abhängig.
2) (ii) $\Rightarrow$ (iii). Wir zeigen zunächst, daß $\operatorname{span}(M) = V$ gilt. Sei $v \in V \setminus M$. Wegen (ii) ist $M \cup \{v\}$ linear abhängig. Es existieren also verschiedene $v_1, \ldots, v_k \in M$ und $\alpha, \alpha_1, \ldots, \alpha_k \in K$, nicht alle $= 0$, mit

$$\alpha v + \sum_{i=1}^{k} \alpha_i v_i = \underline{0}$$

Da M linear unabhängig ist ((ii)!), gilt $\alpha \neq 0$. Also

$$v = \sum_{i=1}^{k} \left(\frac{-\alpha_i}{\alpha} \right) v_i \in \operatorname{span}(M)$$

Also $V \setminus M \subseteq \operatorname{span}(M)$, und damit: $V = \operatorname{span}(M)$.
Sei schließlich M'' eine echte Teilmenge von M, $v \in M \backslash M''$. Wäre $v \in \operatorname{span}(M'')$, so würde aus (3.11)(ii) folgen, daß $M'' \cup \{v\} \subseteq M$ linear abhängig ist, im Widerspruch dazu, daß M nach Voraussetzung ((ii)!) linear unabhängig ist.
Also $v \notin \operatorname{span}(M'')$ und deshalb $\operatorname{span}(M'') \neq V$.
3) (iii) $\Rightarrow$ (i). Zu zeigen ist, daß M linear unabhängig ist. Seien $v_1, \ldots, v_k$ verschiedene Elemente in M, $\alpha_1, \ldots, \alpha_k \in K$ und

$$\sum_{i=1}^{k} \alpha_i v_i = \underline{0}$$

wäre eines der α_i ungleich 0, z.B. $\alpha_1 \neq 0$, so folgt:

$$v_1 = \sum_{i=1}^{k} \left(\frac{-\alpha_i}{\alpha_1} \right) v_i \in \operatorname{span}\{v_2, \ldots, v_k\} \subseteq \operatorname{span}(M \setminus \{v_1\})$$

Wegen (3.11)(i) folgt $\operatorname{span}(M) = \operatorname{span}(M \setminus \{v_1\})$, also $\operatorname{span}(M \setminus \{v_1\}) = V$, im Widerspruch zu (iii).

(3.13) Satz. *Jeder Vektorraum besitzt eine Basis.*

Bem.: Wir zeigen sogar: Ist $M_0 \subseteq V$ linear unabhängig, so existiert eine M_0 enthaltende Basis von V.

Bew.: 1) Wir betrachten zunächst den Spezialfall, daß es ein endliches Erzeugendensystem, genannt M_1, von V gibt. Dann existiert das kleinste $n \in \mathbb{N}$, so daß es eine n-elementige Teilmenge von M_1 gibt, die V erzeugt. Sei M ein solches Erzeugendensystem. Dann folgt aus (3.12), daß M eine Basis von V ist. Zusätzlich gilt $M \subseteq M_1$.
2) Für den allgemeinen Fall benötigen wir folgendes

(3.14) Lemma von Zorn. Sei $\mathcal{M} \neq \emptyset$ eine Menge, deren Elemente Mengen sind. Gilt für jede Kette $\mathcal{K} \subseteq \mathcal{M}$, daß $\bigcup\limits_{M \in \mathcal{K}} M \in \mathcal{M}$ gilt, so existiert ein maximales Element in $\mathcal{M}$, d.h. ein $M \in \mathcal{M}$, so daß aus $M \subseteq M' \in \mathcal{M}$ folgt $M = M'$.

Dabei heißt eine Teilmenge $\mathcal{K}$ von $\mathcal{M}$ Kette, falls $\mathcal{K} \neq \emptyset$ und falls für je zwei Elemente $M_1, M_2 \in \mathcal{K}$ gilt:

$$(*) \qquad M_1 \subseteq M_2 \text{ oder } M_2 \subseteq M_1$$

Wir nehmen das Lemma von Zorn als Axiom der Mengenlehre hin und definieren in Abhängigkeit von einer gegebenen, linear unabhängigen Menge $M_0 \subseteq V$:

$$\mathcal{M} := \{M \mid M_0 \subseteq M \subseteq V, M \text{ linear unabhängig}\}$$

Wegen $M_0 \in \mathcal{M}$ gilt $\mathcal{M} \neq \emptyset$. Wir wollen die Voraussetzung von (3.14) nachweisen. Sei also $\mathcal{K}$ eine Kette in $\mathcal{M}$. Dann gilt $M_0 \subseteq \bigcup\limits_{M \in \mathcal{K}} M$ und es bleibt zu zeigen, daß $\bigcup\limits_{M \in \mathcal{K}} M$ linear unabhängig ist.

Seien $v_1, \ldots, v_k$ verschiedene Vektoren in $\bigcup\limits_{M \in \mathcal{K}} M$. Dann existieren $M_1, \ldots, M_k \in \mathcal{K}$ mit $v_1 \in M_1, \ldots, v_k \in M_k$. Wegen $(*)$ existiert eine der Mengen $M_1, \ldots, M_k$, sagen wir M_j, die die anderen enthält. Dann sind $v_1, \ldots, v_k$ verschiedene Elemente in M_j und damit linear unabhängig (, da $M_j \in \mathcal{M}$ linear unabhängig ist). Also $\bigcup\limits_{M \in \mathcal{K}} M \in \mathcal{M}$. Nach (3.2) ist ein maximales Element von $\mathcal{M}$, das nach (3.14) existiert, eine M_0 enthaltende Basis von V.

Wichtig: Es gibt im allgemeinen sehr viele Basen eines Vektorraumes. Sie haben aber alle "gleich viele" Elemente, siehe (3.16). Zum Beispiel bilden zwei Vektoren $v = (a, b)$,

$w = (c, d)$ des $\mathbb{R}^2$ genau dann eine Basis des $\mathbb{R}^2$, wenn $ad - bc \neq 0$ (, was "meistens" der Fall ist!).

Bemerkung: Die "Lösung eines homogenen linearen Gleichungssystems", d.h. das Auffinden einer Parameterdarstellung des Lösungsraums, besteht gerade darin, eine Basis dieses Lösungsraums zu finden.

(3.15) Austauschlemma. Sei $B = \{v_1 \ldots, v_n\}$ eine Basis von V mit n Elementen und $w \in V \setminus \{0\}$. Dann existiert $j \in \{1, \ldots, n\}$, so daß $B' = (B \setminus \{v_j\}) \cup \{w\}$ Basis von V ist.

Bew.: Wegen $\operatorname{span}(B) = V$ existieren $\alpha_1, \ldots, \alpha_n \in K$, so daß

$$(*) \qquad w = \sum_{i=1}^{n} \alpha_i v_i$$

gilt. Da $w \neq 0$ ist, existiert ein j, $1 \leq j \leq n$ mit $\alpha_j \neq 0$. Wir zeigen, daß <u>für jedes solche</u> j gilt: $B' = (B \setminus \{v_j\}) \cup \{w\} = \{v_1, \ldots, v_{j-1}, v_{j-1}, \ldots, v_n, w\}$ ist Basis.

Zunächst gilt $v_j \in \operatorname{span}(B')$, da aus $(*)$ folgt

$$v_j = \frac{1}{\alpha_j}\left(w - \sum_{\substack{i=1\\i\neq j}}^{n} \alpha_i v_i\right) = \frac{1}{\alpha_j} w + \sum_{\substack{i=1\\i\neq j}}^{n} \left(\frac{-\alpha_i}{\alpha_j}\right) v_i$$

und die rechte Seite ist eine Linearkombination von Elementen von B'.

Nach (3.11)(i) folgt

$$\operatorname{span}(B') \overset{(3.11)(\mathrm{i})}{=} \operatorname{span}(B' \cup \{v_j\}) = \operatorname{span}(B \cup \{w\}) \overset{(3.11)(\mathrm{i})}{=} \operatorname{span}(B) = V,$$

d.h. B' ist Erzeugendensystem von V. Um die lineare Unabhängigkeit von B' zu zeigen, nehmen wir an, daß für $\beta \in K$, $\beta_1, \ldots, \beta_{j-1}, \beta_{j+1}, \ldots, \beta_n \in K$ gilt:

$$(**) \qquad \beta w + \sum_{\substack{i=1\\i\neq j}}^{n} \beta_i v_i = \underline{0}$$

Mit $(*)$ folgt

$$\beta \sum_{i=1}^{n} \alpha_i v_i + \sum_{\substack{i=1\\i\neq j}}^{n} \beta_i v_i = \underline{0},$$

also

$$\beta \alpha_j v_j + \sum_{\substack{i=1\\i\neq j}}^{n} (\beta \alpha_i + \beta_i) v_i = \underline{0}.$$

Da $B = \{v_1, \ldots, v_n\}$ linear unabhängig ist, folgt aus der vorangehenden Gleichung $\beta\alpha_j = 0$ und $\beta\alpha_i + \beta_i = 0$ für alle $i \neq j$. Da wir j so gewählt hatten, daß $\alpha_j \neq 0$ ist, folgt $\beta = 0$

und daraus $\beta_i = 0$ für alle $i \neq j$. Damit haben wir gezeigt, daß in $(**)$ alle Koeffizienten β und β_i für $i \neq j$ gleich 0 sind, d.h. B' ist linear unabhängig.

(3.16) Austauschsatz von Steinitz. *Sei $B = \{v_1, \ldots, v_n\}$ eine Basis von V mit n Elementen und seien $w_1, \ldots, w_m$ m linear unabhängige Vektoren in V. Dann gilt*

(i) $m \leq n$
(ii) *Es gibt $(n-m)$ Vektoren in B, bei geeigneter Numerierung $v_{m+1}, \ldots, v_n$, so daß $\{w_1, \ldots, w_m, v_{m+1}, \ldots, v_n\}$ eine Basis von V ist.*

Speziell gilt: Jede Basis von V hat n Elemente.

Bew.: Induktion nach m.
1) $m = 1$. $\underline{0} \neq w_1 \in V \Rightarrow V \neq \{0\} \Rightarrow n \geq 1 = m \Rightarrow$ (i). (ii) folgt aus (3.15).
2) $m > 1$. Induktionsvoraussetzung: (i) und (ii) gelten für $m-1$. Seien $w_1, \ldots, w_m$ linear unabhängig gegeben. Die Induktionsvoraussetzung impliziert: $(m-1) \leq n$ und es existieren $n-m+1$ Vektoren in B, o.E. $v_m \ldots, v_n$, so daß $\{w_1, \ldots, w_{m-1}, v_m \ldots, v_n\}$ eine Basis von V ist.

Zeige (i): $m \leq n$. Sonst gilt nach der vorangehenden Überlegung $m-1 = n$ und die Menge $\{w_1, \ldots, w_{m-1}\}$ ist Basis von V $(n-m+1 = 0!)$, im Widerspruch zur Voraussetzung, daß $\{w_1, \ldots, w_m\}$ linear unabhängig ist. Das beweist $m \leq n$.

Zeige (ii): Wir wollen in der Basis $\{w_1, \ldots, w_{m-1}, v_m, \ldots, v_n\}$ eines der v_i, $m \leq i \leq n$, gegen w_m "austauschen" und zwar so, daß wieder eine Basis entsteht.
Da $\operatorname{span}\{w_1, \ldots, w_{m-1}, v_m, \ldots, v_n\} = V$ gilt, existieren $\alpha_1, \ldots, \alpha_n \in K$, so daß

$$w_m = \alpha_1 w_1 + \ldots + \alpha_{m-1} w_{m-1} + \alpha_m v_m + \ldots + \alpha_n v_n$$

gilt. Da $\{w_1, \ldots, w_m\}$ linear unabhängig ist, existiert ein $j \in \{m, \ldots, n\}$ mit $\alpha_j \neq 0$. Wir können annehmen, daß $j = m$ ist. Wendet man das Austauschlemma (3.15) (vgl. auch den Anfang des Beweises von (3.15)!) auf $w := w_m$ und die Basis $\{w_1, \ldots, w_{m-1}, v_m, \ldots, v_n\}$ an, so folgt, daß $\{w_1, \ldots, w_m, v_{m+1}, \ldots, v_n\}$ eine Basis von V ist. Das beweist (ii).

(3.17) Def.: Ist V endlich erzeugter Vektorraum, so heißt die Anzahl der Elemente einer ($\Rightarrow$ jeder) Basis von V die <u>Dimension</u> <u>von</u> V (abgekürzt: $\dim V$).

Ist V nicht endlich erzeugt, so setzen wir $\dim V := \infty$.

Beispiele:

1) $V = \{\underline{0}\} \Leftrightarrow \dim V = 0$ (mit Basis $M = \emptyset$)
2) Für den K-Vektorraum K^n gilt: $\dim(K^n) = n$
   Begründung: Die Standardbasis $\{e_1, \ldots, e_n\}$ von K^n hat n Elemente.
3) $V = \{f \mid f : \mathbb{R} \to \mathbb{R}\}$, vgl. Bsp. 4) nach (3.1). Für diesen $\mathbb{R}$-Vektorraum gilt $\dim V = \infty$, vgl. Blatt 6, Aufgabe 4(b).

Bem.: V endlich erzeugt $\Leftrightarrow$ V endlich-dimensional.

(3.18) Folgerungen. Sei V Vektorraum, $\dim V = n < \infty$. Dann gilt:

(i) Ist $M \subseteq V$ linear unabhängig, so gilt $\#M \leq n$, und es gilt $\#M = n$ genau dann, wenn M Basis von V ist.
(ii) Ist $M \subseteq V$ Erzeugendensystem von V, so gilt $\#M \geq n$, und es gilt $\#M = n$ genau dann, wenn M Basis von V ist.
(iii) Ist $U \subseteq V$ Unterraum, so gilt $\dim U \leq \dim V$. Jede Basis von U ist in einer Basis von V enthalten. Es gilt genau dann $\dim U = \dim V$, wenn $U = V$ gilt.

Lösung von Blatt 5, Aufgabe 3: Sei U Unterraum des $\mathbb{R}^2$ mit $\{(0,0)\} \neq U \neq \mathbb{R}^2$. Dann existiert $v \in \mathbb{R}^2 \setminus \{(0,0)\}$, so daß $U = \{rv \mid r \in \mathbb{R}\}$.

Bew.: Es gilt $0 < \dim U < 2 = \dim(\mathbb{R}^2)$, also $\dim U = 1$. Sei $\{v\}$ Basis von $U \Rightarrow$ Beh.

Beweis von (3.18):

(i) (3.16)(i) $\Rightarrow \#M \leq n$. Speziell: Ist $\#M = n$, so ist M (im Sinne von (3.12)!) eine maximale linear unabhängige Teilmenge, also nach (3.12) eine Basis.
(ii) Analog zu (i).
(iii) Ist B' Basis von U, so ist B' als Teilmenge von V linear unabhängig, also $\dim U = \#B' \leq n = \dim V$ nach (i). Aus (3.16)(ii) folgt, daß B' in einer Basis von V enthalten ist. Gilt $\dim U = \dim V$, so ist nach (i) jede Basis B' von U auch Basis von V, d.h. $U = \operatorname{span}(B') = V$.

(3.19) Def.: Seien U_1, U_2 Unterräume eines Vektorraums V. Dann heißt der Unterraum

$$U_1 + U_2 := \operatorname{span}(U_1 \cup U_2)$$

die <u>Summe von</u> U_1 <u>und</u> U_2. Gilt $U_1 \cap U_2 = \{\underline{0}\}$, so nennt man $\operatorname{span}(U_1 \cup U_2)$ die <u>direkte Summe von</u> U_1 <u>und</u> U_2 und schreibt

$$\operatorname{span}(U_1 \cup U_2) =: U_1 \oplus U_2$$

Bsp.: Für $0 < m < n$ gilt: $K^n = (K^m \times \{\underline{0}\}) \oplus (\{\underline{0}\} \times K^{n-m})$

(3.20) Fakt. (i) $U_1 + U_2 = \{u_1 + u_2 \mid u_1 \in U_1, u_2 \in U_2\}$
(ii) Gilt $U_1 \cap U_2 = \{0\}$, so existieren für jedes $v \in U_1 \oplus U_2$ genau zwei Vektoren $u_1 \in U_1$, $u_2 \in U_2$, so daß $v = u_1 + u_2$ gilt.

Bsp.: 1) $V = \mathbb{R}^2$, U_1, U_2 Unterräume der Dimension 1, $U_1 \neq U_2$. Dann: $U_1 \cap U_2 = \{(0,0)\}$, $U_1 \oplus U_2 = \mathbb{R}^2$.
2) $V = \mathbb{R}^3$, $U_1 \neq U_2$ Unterräume der Dimension 2. Dann: $\dim(U_1 \cap U_2) = 1$, $U_1 + U_2 = \mathbb{R}^3$.

Beweis: Die Menge $\{u_1 + u_2 \mid u_1 \in U_1, u_2 \in U_2\}$ ist Unterraum (!), der U_1 und U_2 enthält, also $\operatorname{span}(U_1 \cup U_2) \subseteq \{u_1 + u_2 \mid u_1 \in U_1, u_2 \in U_2\}$, vgl. Def. (3.6).
Die Inklusion $\{u_1 + u_2 \mid u_1 \in U_1, u_2 \in U_2\} \subseteq \operatorname{span}(U_1 \cup U_2)$ ist klar!

Gilt $v = u_1 + u_2 = u_1' + u_2'$ mit $\{u_1, u_1'\} \subseteq U_1$, $\{u_2, u_2'\} \subseteq U_2$, so folgt

$$u_1 - u_1' = u_2' - u_2 \in U_1 \cap U_2 = \{\underline{0}\}, \text{ also } u_1 = u_1', u_2 = u_2'.$$

(3.21) Lemma.

(i) Ist U Unterraum von V, so existiert ein Unterraum W von V, so daß $U \cap W = \{\underline{0}\}$ und $U \oplus W = V$ gilt.
(W heißt ein zu U komplementärer Unterraum)

(ii) Sind U_1, U_2 Unterräume von V mit $U_1 \cap U_2 = \{0\}$, so gilt

$$\dim(U_1 \oplus U_2) = \dim U_1 + \dim U_2.$$

Bem.: Im allgemeinen gibt es sehr viele zu U komplementäre Unterräume. Beispiel: Es sei $U = \mathbb{R} \times \{0\} \subseteq \mathbb{R}^2$ und es sei $v = (v_1, v_2) \in \mathbb{R}^2$ und $v_2 \neq 0$. Dann ist $W = \operatorname{span}\{v\}$ zu U komplementär.

Bew.: (i) Ergänze eine Basis B' von U nach (3.18)(iii) zu einer Basis B von V und setze $W := \operatorname{span}(B \setminus B')$.
(ii) Für $i = 1, 2$ seien B_1, B_2 Basen von U_1, U_2. Aus $U_1 \cap U_2 = \{0\}$ folgt $B_1 \cap B_2 = \emptyset$. Wir zeigen, daß $B_1 \cup B_2$ Basis von $U_1 \oplus U_2$ ist ($\Rightarrow \dim(U_1 \oplus U_2) = \#(B_1 \cup B_2) = \#B_1 + \#B_2 = \dim U_1 + \dim U_2$). Wegen (3.20)(i) ist $B_1 \cup B_2$ Erzeugendensystem von $U_1 \oplus U_2$. Um zu zeigen, daß $B_1 \cup B_2$ linear unabhängig ist, sei

$$\sum_{i=1}^{k} \alpha_i v_i + \sum_{j=1}^{l} \beta_j w_j = \underline{0},$$

wobei $\alpha_1, \ldots, \beta_l \in K$, $v_1, \ldots, v_k$ verschiedene Elemente von B_1, $w_1, \ldots, w_l$ verschiedene Elemente von B_2 sind. Dann folgt

$$\sum_{i=1}^{k} \alpha_i v_i = \sum_{j=1}^{l} (-\beta_j) w_j \in U_1 \cap U_2.$$

Wegen $U_1 \cap U_2 = \{0\}$ folgt

$$\sum_{i=1}^{k} \alpha_i v_i = \underline{0} = \sum_{j=1}^{l} (-\beta_j) w_j.$$

Da B_1 und B_2 beide linear unabhängig sind, folgt aus den vorangehenden Gleichungen $\alpha_1 = \ldots = \alpha_k = 0$ und $\beta_1 = \ldots = \beta_l = 0$. Das beweist die lineare Unabhängigkeit von $B_1 \cup B_2$.

Bem.: (3.21)(ii) gilt – richtig interpretiert – auch für den Fall, daß U_1 oder U_2 unendlich-dimensional sind. Für den Rest des Kapitels werden wir aber $\underline{\dim V < \infty}$ voraussetzen.

(3.22) Dimensionssatz. *Seien U_1, U_2 Unterräume eines Vektorraums V. Dann gilt*

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2)$$

Bew.: $U_1 \cap U_2 =: U_{12}$. Nach (3.21)(i) existieren Unterräume W_1(von $U_1 \Rightarrow$ von V) und W_2, so daß $U_1 = U_{12} \oplus W_1$ und $U_2 = U_{12} \oplus W_2$ gilt. Dann folgt

$$(*) \qquad U_1 + U_2 = (U_{12} \oplus W_1) \oplus W_2$$

Denn einerseits läßt sich jedes Element $v \in U_1 + U_2$ als

$$v = u_1 + u_2 = u_{12} + w_1 + u'_{12} + w_2 = (u_{12} + u'_{12}) + w_1 + w_2$$

mit $u_1 \in U_1$, $u_2 \in U_2$, $u_{12}, u'_{12} \in U_{12}$, $w_1 \in W_1$ und $w_2 \in W_2$ schreiben und andererseits gilt $(U_{12} \oplus W_1) \cap W_2 = U_1 \cap (U_2 \cap W_2) = U_{12} \cap W_2 = \{\underline{0}\}$.

Nach (3.21)(ii) folgt aus $(*)$: $\dim(U_1 + U_2) = \dim U_{12} + \dim W_1 + \dim W_2$, und aus $U_1 = U_{12} \oplus W_1$, $U_2 = U_{12} \oplus W_2$:

$$\begin{aligned}\dim U_1 &= \dim U_{12} + \dim W_1 \\ \dim U_2 &= \dim U_{12} + \dim W_2\end{aligned}$$

Subtrahiert man diese beiden Gleichungen von der vorangehenden, so folgt

$$\dim(U_1 + U_2) - \dim U_1 - \dim U_2 = -\dim U_{12}$$

oder

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim U_{12},$$

wie behauptet.

Bez.: Ein Unterraum $H \subseteq V$ heißt Hyperebene, falls $\dim H = \dim V - 1$ gilt.

(3.23) Folgerung. Ist $U \subseteq V$ Unterraum, $H \subseteq V$ Hyperebene, so gilt entweder

$$\dim(U \cap H) = \dim U - 1$$

$$\text{oder} \quad U \subseteq H (\Leftrightarrow U = U \cap H \Leftrightarrow \dim(U \cap H) = \dim U).$$

Bew.: Aus (3.22) folgt

$$(*) \qquad \dim(U \cap H) = \dim U + (\dim V - 1) - \dim(U + H).$$

Gilt $\dim(U + H) < \dim V$, so folgt aus (3.18)(iii) angewendet auf $H \subseteq U + H$, daß $H = U + H$, d.h. $U \subseteq H$ gilt. Gilt $\dim(U + H) = \dim V$, so folgt aus $(*)$: $\dim(U \cap H) = \dim U - 1$.

(3.24) Lemma. Ist K ein Körper und sind $a_1, \ldots, a_n \in K$ nicht alle $= 0$, so ist der Lösungsraum $L_I \subseteq K^n$ der Gleichung

$$I \qquad a_1 x_1 + \ldots + a_n x_n = 0$$

eine Hyperebene in K^n.

Bew.: Sei etwa $a_j \neq 0$. Dann folgt durch Auflösen von I nach x_j:

$$L_I = \{x = (x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_j = \sum_{\substack{i=1 \\ i \neq j}}^{n} \left(-\frac{a_i}{a_j}\right) x_i\}.$$

Wir definieren für $i \neq j$ die Vektoren $v_i = e_i - \frac{a_i}{a_j} e_j \in K^n$. Dann gilt

$$L_I = \mathrm{span}\{v_i \mid i \in \{1, \ldots, n\} \setminus \{j\}\}.$$

Da die Vektoren $v_i, i \neq j$, linear unabhängig sind (!), folgt $\dim L_I = n - 1$.

(3.25) Satz. Sei K ein Körper und

$$I \quad \begin{array}{ccccc} a_{11} & x_1 + \ldots + & a_{1n} & x_n & = 0 \\ \vdots & & \vdots & & \vdots \\ a_{k1} & x_1 + \ldots + & a_{kn} & x_n & = 0 \end{array}$$

ein homogenes lineares Gleichungssystem mit $a_{ij} \in K$ für $1 \leq i \leq k$, $1 \leq j \leq n$. Dann gilt $\dim L_I \geq n - k$.

Beweis: Induktion nach k.
$k = 1$: Nach (3.24) ist der Lösungsraum $L_I \subseteq K^n$ einer einzigen linearen Gleichung $(n-1)$-dimensional, es sei denn, alle Koeffizienten der Gleichung sind 0 (und in diesem Fall gilt $L_I = K^n$). In jedem Fall gilt $\dim L_I \geq n - 1$.

$k > 1$: Es sei I' das Gleichungssystem, das aus den ersten $(k-1)$ Gleichungen von I besteht, und I_k sei die k'te (=letzte) Gleichung von I. Dann gilt

$$L_I = L_{I'} \cap L_{I_k}$$

Die Induktionsvoraussetzung besagt: $\dim L_{I'} \geq n - k + 1$. Sind alle Koeffizienten von I_k gleich 0, so gilt $L_{I_k} = K^n$, also $L_I = L_{I'}$ und damit $\dim L_I = \dim L_{I'} \geq n - k + 1 > n - k$.
Sind nicht alle Koeffizienten von I_k gleich 0, so ist L_{I_k} nach (3.24) eine Hyperebene.
Dann folgt aus (3.23), angewendet auf $U = L_{I'}$, $H = L_{I_k}$:
$\dim L_I = \dim(L_{I'} \cap L_{I_k}) \geq \dim(L_{I'}) - 1 \geq n - k$.

Bez.: Eine Teilmenge A von V heißt <u>k-dimensionaler</u> <u>affiner</u> <u>Unterraum</u>, falls es ein $v \in V$ und einen k-dimensionalen Unterraum $U \subseteq V$ gibt, so daß $A = \{v + u \mid u \in U\} =: v + U$ gilt. Ist $k = 1$, so nennt man A eine affine Gerade, ist $k = 2$, eine affine Ebene und ist $k = \dim V - 1$, eine affine Hyperebene.

Bem.: Ein affiner Unterraum A von V ist genau dann ein Unter(vektor)raum von V, wenn $\underline{0} \in A$ gilt. Sind v_0, $v_1 \in V$ und sind U_0, U_1 Unterräume von V, so gilt $v_0 + U_0 = v_1 + U_1$ genau dann, wenn $U_0 = U_1$ und $v_1 - v_0 \in U_0$ gilt.

(3.26) Folgerung. Sei K ein Körper und I ein inhomogenes lineares Gleichungssystem mit k Gleichungen für n Unbekannte. Dann gilt entweder $L_I = \emptyset$ oder L_I ist ein affiner Unterraum von K^n der Dimension $\geq n - k$. Insbesondere besitzt I höchstens dann genau eine Lösung (d.h. $\# L_I = 1$), wenn $k \geq n$ gilt.

Bew.: Die Behauptung folgt durch Kombination von (1.6) mit (3.25).

Bem.: Die Aussagen von (3.25) und (3.26) kann man auch direkt mit dem Gaußschen Eliminationsverfahren einsehen.

Exkurs: Fehlerkorrigierende Codes

Lit.: MacWilliams, F.J., Sloane, N.J.A.: The Theory of Error-Correcting Codes, North-Holland, Amsterdam 1978.
van Lint, J.H.: Introduction to Coding Theory (3. Aufl.), Springer 1999.

Problem: bei der Übermittlung von Daten entstehen Fehler. Gesucht ist eine Methode, die die (meisten) Fehler automatisch korrigiert.

Daten $\stackrel{\text{Kodierung}}{\longrightarrow}$ endl. Folge der Länge n von $0,1$ $\stackrel{\text{Übermittlung}}{\longrightarrow}$ fehlerhafter Folge von $0,1$ $\stackrel{\text{automat. Fehlerkorrektur}}{\longrightarrow}$ ursprüngl. Folge von $0,1$ $\stackrel{\text{Dekodierung}}{\longrightarrow}$ ursprüngl. Daten.

Das kann nur funktionieren, wenn bei der Kodierung nur spezielle Folgen von $0,1$ ("die Wörter des Codes") verwendet werden.

Mathematische Formulierung:

$$\{0,1\}^n = \{(x_1,\ldots,x_n) \mid x_1 \in \{0,1\},\ldots,x_n \in \{0,1\}\}$$

Folgen von $0,1$ der Länge n, z.B. $(0,1,0,1,0,1,0) \in \{0,1\}^7$

Hamming-Abstand d von Elementen $x, y \in \{0,1\}^n$

$$d(x,y) = \#\{i \mid 1 \leq i \leq n \text{ und } x_i \neq y_i\} \quad (\Rightarrow d(x,y) \in \{0,\ldots,n\})$$

Dabei wird mit $\#M$ die Anzahl der Elemente einer Menge M bezeichnet.

Bsp.: $(n=4)$ $x = (0,0,1,1)$, $y = (0,1,1,0) \Rightarrow d(x,y) = 2$.

Eigenschaften von d: $\forall x, y, z \in \{0,1\}^n$ gilt:

(i) $d(x,y) \geq 0$ und $d(x,y) = 0 \Leftrightarrow x = y$
(ii) $d(x,y) = d(y,x)$ ("Symmetrie")
(iii) $d(x,y) \leq d(x,z) + d(z,y)$ ("Dreiecksungleichung")

(i)-(iii) $\Leftrightarrow$ "d ist Abstandsfunktion"

Zu $x \in \{0,1\}^n$, $t \in \mathbb{N}$ sei

$$B(x,t) = \{y \in \{0,1\}^n \mid d(x,y) \leq t\}$$

der (d-)Ball um x vom Radius t.

Bem.: $B(x,1)$ enthält genau $(n+1)$ Elemente.

(E1) Def.:

(i) Ein Code der Länge n ist eine Teilmenge $C \neq \emptyset$ von $\{0,1\}^n$.
(ii) Ein Code C heißt t-fehlerkorrigierend $\Leftrightarrow$
Ist $x \in C, y \in C, x \neq y$, so gilt $B(x,t) \cap B(y,t) = \emptyset$.

Idee: Sei C t-fehlerkorrigierender Code. Als kodierte Nachrichten sind nur "Codewörter" $x \in C$ erlaubt.

$$x \in C \overset{\text{Übermittlung}}{\longrightarrow} \overline{x} \in \{0,1\}^n \overset{\text{automat. Korrektur}}{\longrightarrow} \overline{\overline{x}} \in C \text{ so, daß } d(\overline{x}, \overline{\overline{x}}) \text{ minimal.}$$

Dann: Ist $d(x, \overline{x}) \leq t$, so gilt $\overline{\overline{x}} = x$, da es nur ein $x \in C$ mit $d(\overline{x}, x) \leq t$ gibt.
D.h.: Sind bei der Übermittlung höchstens t Fehler aufgetreten, so ordnet die automatische Fehlerkorrektur der fehlerhaften Nachricht $\overline{x}$ das ursprüngliche Codewort x zu.

Bem.: 1) $C \subseteq \{0,1\}^n$ ist t-fehlerkorrigierend $\Leftrightarrow$ Ist $x \in C$, $y \in C$ und $x \neq y$, so gilt $d(x, y) \geq 2t + 1$ (Dreiecksungleichung!)
2) $C \subseteq \{0,1\}^n$ 1-fehlerkorriegierend $\Rightarrow \#C \cdot (n+1) \leq 2^n (= \#\{0,1\}^n)$

Problem: Zeitaufwand Fehlerkorrektur $\approx \#C \cdot n \approx 2^n$

Lineare Codes

$\{0,1\}^n$ ist Vektorraum über dem Körper $\{0,1\} = \mathbb{Z}_2$

(E2) Def. $C \subseteq \{0,1\}^n$ heißt linearer Code $\Leftrightarrow C$ ist Unterraum von $\{0,1\}^n$
($\Leftrightarrow \underline{0} \in C$ und $\forall x, y \in C : x + y \in C$)

Sei $p \in \mathbb{N}$ und $n := 2^p - 1$ (z.B. $p = 5, n = 31 \Rightarrow 2^n = 3^{31} = 2 \cdot (2^{10})^3 = 2 \cdot (1024)^3 \approx 2 \cdot 10^9$)

Seien $v_1, \ldots, v_n$ die von $\underline{0} \in \{0,1\}^p$ verschiedenen Elemente von $\{0,1\}^p$ und $F : \{0,1\}^n \rightarrow \{0,1\}^p$ definiert durch

$$F(x_1, \ldots, x_n) = x_1 v_1 + x_2 v_2 + \ldots + x_n v_n \in \{0,1\}^p$$

Dann gilt für alle $x, y \in \{0,1\}^n : F(x + y) = F(x) + F(y)$.

Daraus folgt: $C := \{x \in \{0,1\}^n \mid F(x) = \underline{0}\}$ ist Unterraum (=linearer Code).

C heißt der n'te Hammingcode.

(E3) Satz. *Der n'te Hammingcode ist ein 1-fehlerkorrigierender Code der Länge n*

$$\textit{mit } \#C = \frac{2^n}{n+1} (\textit{"C ist perfekt"})$$

Bew.: Seien $x, y \in C$, $x \neq y$. Zu zeigen: $d(x, y) \geq 3$

$$x, y \in C \Rightarrow \quad (*) \quad (x_1 - y_1)v_1 + \ldots + (x_n - y_n)v_n = \underline{0}.$$

Wegen $x \neq y$ existiert ein $i \in \{1, \ldots, n\} : x_i \neq y_i$. Dann $(x_i - y_i)v_i = v_i \neq \underline{0}$.
Wegen $(*)$ existiert ein $j \neq i$ mit $x_j \neq y_j$, also $(x_j - y_j)v_j = v_j$. Wenn alle anderen Komponenten von x und y übereinstimmten, so wäre $(*)$ die Gleichung

$$v_i + v_j = \underline{0}$$

Da in $\mathbb{Z}_2$ $1+1=0$ gilt, gilt $v+v=\underline{0}$ für alle $v \in \{0,1\}^n$. D.h. aus $v_i+v_j=\underline{0}$, folgt $v_i=v_j$, im Widerspruch dazu, daß die v_i alle verschieden sind. Also gibt es mindestens eine weitere Komponente, in der x und y nicht übereinstimmen, d.h. $d(x,y) \geq 3$, wie behauptet.

Ist $x \in \{0,1\}^n$ und $F(x) \neq \underline{0}$, so gilt $F(x)=v_i$ für ein $i \in \{1,\ldots,n\}$.
Daraus folgt $F(x-e_i)=F(x)-F(e_i)=\underline{0}$, d.h. $x-e_i \in C$ und $d(x,x-e_i)=1$.
Also $\bigcup\limits_{x \in C} B(x,1)=\{0,1\}^n$ und damit $\#C \cdot (n+1)=2^n$.

Algorithmus zur Korrektur <u>eines</u> Fehlers:

Sei $\overline{x} \in \{0,1\}^n$ die empfangene Nachricht. Berechne $F(\overline{x}) \in \{0,1\}^p$.

$$\begin{array}{lcl} F(\overline{x})=\underline{0} & \Rightarrow & \overline{\overline{x}} := \overline{x} \\ F(\overline{x})=v_i & \Rightarrow & \overline{\overline{x}} := \overline{x}+e_i \end{array}$$

Zeitaufwand $\approx n^2p$ $(=31^2 \cdot 5 \approx 5 \cdot 10^3) \ll 2^n \approx 2 \cdot 10^9)$

Kapitel 4: Lineare Abbildungen und Matrizen.

Wie überall, so sind auch in der Mathematik nicht nur die einzelnen Objekte wichtig, sondern auch die Beziehungen zwischen ihnen, die hier meist durch Abbildungen beschrieben werden, die (in einem zu definierenden Sinn) die Struktur der Objekte erhalten, sogenannte Homomorphismen. Im folgenden seien V und W Vektorräume über <u>demselben</u> Körper K.

(4.1) Definition. Eine Abbildung $L: V \to W$ heißt <u>linear</u> (oder <u>Vektorraumhomomorphismus</u>), falls für alle $v, v_1, v_2 \in V$ und alle $\alpha \in K$ gilt:

$$\begin{array}{llcll} \text{(i)} & L(v_1+v_2) & = & L(v_1)+L(v_2) & (L \text{ ist "additiv"}) \\ \text{(ii)} & L(\alpha v) & = & \alpha L(v) & (L \text{ ist "homogen"}) \end{array}$$

Bez.: Statt "lineare Abbildung" ist auch "linearer Operator" üblich. Im Fall $W=K$ spricht man oft von einem "linearen Funktional" oder einer "Linearform" $L: V \to K$.

$$\mathrm{Hom}(V,W) := \{L \mid L: V \to W \text{ linear}\}.$$

Eine lineare Abbildung $L: V \to V$ heißt <u>Endomorphismus</u>, $\mathrm{End}(V)=\{L \mid L: V \to V \text{ linear}\}$.

Ein Homomorphismus $L: V \to W$ heißt <u>Isomorphismus</u>, falls L bijektiv ist. Zwei K-Vektorräume V und W heißen <u>isomorph</u>, falls es einen Isomorphismus $L: V \to W$ gibt.

Beispiele:

0) $\mathrm{id}_V \in \mathrm{End}(V)$.
Ist $L: V \to W$ die "0-Abbildung" definiert durch

$$\forall v \in V: Lv=\underline{0} \in W,$$

so gilt $L \in \mathrm{Hom}(V,W)$. Ist $U \subseteq V$ Unterraum, so ist die Inklusion $i: U \to V$, $\forall u \in U$: $i(u) := u$, eine lineare Abbildung.

1) Für $\alpha \in K$ sei $S_\alpha : V \to V$, $S_\alpha(v) := \alpha v$, die "Streckung" um α. Dann gilt: $S_\alpha \in \mathrm{End}(V)$.
2) $V :=$ der $\mathbb{R}$-Vektorraum $\mathbb{R}$. $L \in \mathrm{End}(\mathbb{R}) \Leftrightarrow \exists m \in \mathbb{R} \forall x \in \mathbb{R} : L(x) = mx$. Nämlich: $m := L(1)$. Dann gilt: $L(x) = L(x \cdot 1) \overset{\text{(ii)}}{=} x \cdot L(1) = m \cdot x$.
3) Zu reellen Zahlen $a < b$ ist $C^0([a,b],\mathbb{R}) = \{f \mid f : [a,b] \to \mathbb{R} \text{ stetig}\}$ ein $\mathbb{R}$-Vektorraum, vgl. Bsp. 4 nach (3.1).

$$L : C^0([a,b],\mathbb{R}) \to \mathbb{R}, \quad L(f) := \int_a^b f(x)dx$$

L ist linearer Operator von $C^0([a,b],\mathbb{R})$ nach $\mathbb{R}$ ("Lineares Funktional").
4) $C^1(\mathbb{R},\mathbb{R}) = \{f \mid f : \mathbb{R} \to \mathbb{R} \text{ ist differenzierbar und } f' : \mathbb{R} \to \mathbb{R} \text{ ist stetig}\}$
$D : C^1(\mathbb{R},\mathbb{R}) \to C^0(\mathbb{R},\mathbb{R}), D(f) := f'$. D ist linearer Operator.
5) Die im Exkurs über Codes definierte Abbildung $F : (\mathbb{Z}_2)^n \to (\mathbb{Z}_2)^p$ ist linear.

Wichtig ist folgender Zusammenhang mit der Analysis: Ganz grob gesprochen ist die Analysis die Kunst, nichtlineare Funktionen (in einer kleinen Umgebung eines festen Punktes x_0) durch lineare Funktionen zu approximieren und aus Eigenschaften der approximierenden linearen Funktion (, mit der man gut explizit rechnen kann) auf Eigenschaften der ursprünglichen nichtlinearen Funktion zu schließen.

Im Fall von (nichtlinearen) Funktionen $f : \mathbb{R} \to \mathbb{R}$ und $x_0 \in \mathbb{R}$ ist die approximierende lineare Funktion gegeben durch $h \in \mathbb{R} \to f'(x_0)h \in \mathbb{R}$, und es gilt für den durch $f(x_0+h) = f(x_0) + f'(x_0)h + R(h)$ definierten "Approximationsfehler" $R(h) : \lim\limits_{h\to 0} \frac{R(h)}{h} = 0$. Betrachtet man (z.B. in der Analysis II) Funktionen $f : \mathbb{R}^n \to \mathbb{R}^m$ und $x_0 \in \mathbb{R}^n$, so wird in der analogen Definition der Differenzierbarkeit aus der linearen Funktion $h \in \mathbb{R} \to f'(x_0)h \in \mathbb{R}$ eine lineare Abbildung $Df(x_0) \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$.
Noch allgemeiner betrachtet man in der Theorie der Differentialgleichungen (und in der Physik und in anderen Wissenschaften) nichtlineare Differentialoperatoren zwischen unendlich-dimensionalen Funktionenräumen, die durch lineare Operatoren zwischen solchen Funktionenräumen approximiert werden. Oft verwendet man in der Physik direkt diese linearisierten Operatoren (z.B. Wellengleichung, Wärmeleitungsgleichung).

Bez.: Ist $L \in \mathrm{Hom}(V,W)$, so heißt

$$\ker L := \{v \in V \mid L(v) = \underline{0} \in W\} = L^{-1}(\{\underline{0}\})$$

der Kern von L und

$$\mathrm{im}\, L := \{w \in W \mid \exists v \in V : L(v) = w\} = L(V)$$

das Bild von L.

(4.2) Fakt. Sei $L \in \mathrm{Hom}(V,W)$, U_1 Unterraum von V, U_2 Unterraum von W. Dann gilt:

(i) $L(U_1)$ ist Unterraum von W.

(ii) $L^{-1}(U_2)$ ist Unterraum von V.

Speziell: $\ker L$ ist Unterraum von V, $\operatorname{im} L$ ist Unterraum von W, und es gilt $L(\underline{0}) = \underline{0}$.

Bew.:

(i) Sind $w_1, w_2 \in L(U_1)$, so existieren $v_1, v_2 \in U_1$, so daß $L(v_1) = w_1$ und $L(v_2) = w_2$ gilt. Da U_1 Unterraum ist, gilt $v_1 + v_2 \in U_1$, also $L(v_1 + v_2) \in L(U_1)$. Die Additivität von L impliziert: $w_1 + w_2 = L(v_1) + L(v_2) = L(v_1 + v_2) \in L(U_1)$.
Analog folgt: Ist $w \in L(U_1), \alpha \in K$, so gilt $\alpha w \in L(U_1)$. Schließlich impliziert $U_1 \neq \emptyset$, daß $L(U_1) \neq \emptyset$ gilt.

(ii) Ist $v \in L^{-1}(U_2), \alpha \in K$, so gilt $\alpha L(v) \in U_2$, da U_2 ein Unterraum ist. Die Homogenität von L impliziert: $L(\alpha v) = \alpha L(v) \in U_2$, also $\alpha v \in L^{-1}(U_2)$. Analog folgt: Sind $v_1, v_2 \in L^{-1}(U_2)$, so gilt $v_1 + v_2 \in L^{-1}(U_2)$. Schließlich zeigen wir, daß $L(\underline{0}) = \underline{0}$ gilt. Daraus folgt dann $\underline{0} \in L^{-1}(U_2)$, speziell $L^{-1}(U_2) \neq \emptyset$. $L(\underline{0}) = L(0\underline{0}) = 0L(\underline{0}) = \underline{0}$, vgl. (3.2)(i).

(4.3) Fakt. Sei $L \in \operatorname{Hom}(V, W)$. Folgende Aussagen sind äquivalent.

(i) L ist injektiv.
(ii) $\ker L = \{\underline{0}\}$.
(iii) Ist $M \subseteq V$ linear unabhängig, so ist $L(M) \subseteq W$ linear unabhängig.

Bew.: (i) $\Rightarrow$ (ii): Ist L injektiv, so ist $\underline{0} \in V$ das einzige Element von V, das durch L auf $\underline{0} \in W$ abgebildet wird, also $\ker L = \{\underline{0}\}$.
(ii) $\Rightarrow$ (iii): Es gelte $\ker L = \{\underline{0}\}$ und $M \subseteq V$ sei linear unabhängig. Seien $w_1, \ldots, w_k$ verschiedene Elemente von $L(M)$, $\alpha_1, \ldots, \alpha_k \in K$, und es gelte $\sum\limits_{i=1}^{k} \alpha_i w_i = \underline{0}$. Wir wollen zeigen, daß $\alpha_1 = \ldots = \alpha_k = 0$ gilt. Zu jedem $w_i \in L(M)$ existiert ein $v_i \in M$ mit $L(v_i) = w_i$. Da die $w_1, \ldots, w_k$ verschieden sind, sind auch die $v_1, \ldots, v_k$ verschieden. Durch Induktion folgt aus der Linearität von L, daß

$$L\left(\sum_{i=1}^{k} \alpha_i v_i\right) = \sum_{i=1}^{k} \alpha_i L(v_i)$$

gilt. Aus $\sum\limits_{i=1}^{k} \alpha_i L(v_i) = \sum\limits_{i=1}^{k} \alpha_i w_i = \underline{0}$ folgt also

$$\sum_{i=1}^{k} \alpha_i v_i \in \ker L = \{0\},$$

d.h. $\sum\limits_{i=1}^{k} \alpha_i v_i = \underline{0}$. Da die $v_1, \ldots, v_k$ verschiedene Elemente der linear unabhängigen Menge M sind, folgt $\alpha_1 = \ldots = \alpha_k = 0$.

(iii) $\Rightarrow$ (i): Seien $v_1 \neq v_2 \in V$. Dann gilt $v_2 - v_1 \neq \underline{0}$, d.h. die Menge $M = \{v_2 - v_1\} \subseteq V$ ist linear unabhängig. Nach (iii) ist dann auch $L(M) = \{L(v_2 - v_1)\} \subseteq W$ linear unabhängig, d.h. $L(v_2 - v_1) \neq \underline{0}$. Nun gilt (vgl. (3.2)(iii)):

$$\underline{0} \neq L(v_2 - v_1) = L(v_2 + (-1)v_1) = L(v_2) + L((-1)v_1) = L(v_2) - L(v_1).$$

Also $L(v_1) \neq L(v_2)$, d.h. L ist injektiv.

Bem.: Die linke Seite eines linearen Gleichungssystems I mit k Gleichungen und n Unbekannten definiert wie folgt eine lineare Abbildung $L \in \mathrm{Hom}(K^n, K^k)$:

$$L(x_1, \dots, x_n) := (a_{11}x_1 + \ldots + a_{1n}x_n, \dots, a_{k1}x_1 + \ldots + a_{kn}x_n) \in K^k.$$

Das Problem, für eine gegebene rechte Seite $b = (b_1, \dots, b_k) \in K^k$ von I die Lösungsmenge L_I zu finden, ist also gerade das Problem die Urbildmenge $L^{-1}(\{b\})$ zu finden,

$$L_I = L^{-1}(\{b\}).$$

Es gilt also:
I ist besitzt für die rechte Seite $b = (b_1, \dots, b_k)$ eine Lösung $\Leftrightarrow$ $b \in \operatorname{im} L$.
I besitzt für jede rechte Seite höchstens eine Lösung $\Leftrightarrow$ L injektiv $\Leftrightarrow$ $\ker L = \{\underline{0}\}$.
I besitzt für jede rechte Seite genau eine Lösung $\Leftrightarrow$ L Isomorphismus.

(4.4) Lemma. Sei $L \in \mathrm{Hom}(V, W)$, $M \subseteq V$. Dann gilt:

$$L(\mathrm{span}(M)) = \mathrm{span}\, L(M).$$

Bew.: Übungsaufgabe.

(4.5) Folgerung. Sei $L \in \mathrm{Hom}(V, W)$. Dann sind äquivalent:

(i) L ist Isomorphismus ($\Leftrightarrow$ L bijektiv).
(ii) Für jede Basis B von V gilt: $L|B$ ist injektiv und $L(B)$ ist Basis von W.
(iii) Es existiert eine Basis B von V, so daß $L|B$ injektiv und $L(B)$ Basis von W ist.

Bew.:

(i) $\Rightarrow$ (ii). Sei $B \subseteq V$ Basis $\overset{(4.3)}{\Rightarrow}$ $L(B)$ ist linear unabhängig.
$\mathrm{span}\, L(B) \overset{(4.4)}{=} L(\mathrm{span}\, B) = L(V) \overset{L \text{ surjektiv}}{=} W \Rightarrow L(B)$ ist Erzeugendensystem von W. L injektiv $\Rightarrow$ $L|B$ injektiv.
(ii) $\Rightarrow$ (iii): klar, da nach (3.13) eine Basis von V existiert.
(iii) $\Rightarrow$ (i). Zeige: L ist injektiv. Sei $v \in \ker L$. Dann existiert $k \in \mathbb{N}$ und verschiedene $v_1, \dots, v_k \in B$ und $\alpha_1, \dots, \alpha_k \in K$: $v = \sum\limits_{i=1}^{k} \alpha_i v_i$. Daraus folgt

$$(*) \qquad \underline{0} = L(v) = \sum_{i=1}^{k} \alpha_i L(v_i).$$

Da $L|B$ injektiv ist, sind die $L(v_1), \ldots, L(v_k)$ verschiedene Elemente der Basis $L(B)$. Also folgt aus $(*)$: $\alpha_1 = \ldots = \alpha_k = 0$ und daraus $v = 0$. Also $\ker L = \{\underline{0}\}$, d.h. L ist injektiv. $L(V) = L \operatorname{span} B) \overset{(4.4)}{=} \operatorname{span} L(B) = W \Rightarrow L$ ist surjektiv.

Bez.: Eine geordnete Basis $\mathcal{G}$ eines n-dimensionalen Vektorraums V ist ein n-Tupel von Vektoren $\mathcal{G} = (v_1, \ldots, v_n)$, so daß $\{v_1, \ldots, v_n\}$ eine Basis von V ist.

(4.6) Satz. *Sei* $\dim V = n < \infty$, $\mathcal{G} = (v_1, \ldots, v_n)$ *geordnete Basis von* V *und* $w_1, \ldots, w_n$ *beliebige Elemente von* W. *Dann existiert genau ein* $L \in \operatorname{Hom}(V, W)$, *so daß* $Lv_1 = w_1 \ldots, Lv_n = w_n$ *gilt.*

Beweis: Vorbemerkung: Wegen $\operatorname{span}\{v_1, \ldots, v_n\} = V$ existieren zu jedem $v \in V$ Skalare $\alpha_1, \ldots, \alpha_n \in K$, so daß $v = \sum\limits_{i=1}^{n} \alpha_i v_i$ gilt, und da die $v_1, \ldots, v_n$ linear unabhängig sind, sind die $\alpha_1, \ldots, \alpha_n$ eindeutig durch v bestimmt, vgl. (3.9).

(i) Eindeutigkeit von L: Es seien L, $L' \in \operatorname{Hom}(V, W)$ und $Lv_i = L'v_i$ für alle $i \in \{1, \ldots, n\}$. Sei $v = \sum\limits_{i=1}^{n} \alpha_i v_i \in V$. Dann gilt

$$L(v) = L\left(\sum_{i=1}^{n} \alpha_i v_i\right) = \sum_{i=1}^{n} \alpha_i L(v_i) = \sum_{i=1}^{n} \alpha_i L'(v_i) = L'(\sum \alpha_i v_i) = L'(v).$$

(ii) Existenz von L: Für $v = \sum\limits_{i=1}^{n} \alpha_i v_i \in V$ definieren wir

$$L(v) := \sum_{i=1}^{n} \alpha_i w_i \in W.$$

Ist $v = \sum\limits_{i=1}^{n} \alpha_i v_i$, $v' = \sum\limits^{n} \alpha_i' v_i$, so $v + v' = \sum\limits_{i=1}^{n} (\alpha_i + \alpha_i') v_i$ und damit:
$L(v + v') = \sum\limits_{i=1}^{n} (\alpha_i + \alpha_i') w_i = \sum\limits_{i=1}^{n} \alpha_i w_i + \sum\limits_{i=1}^{n} \alpha_i' w_i = L(v) + L(v')$, d.h. L ist additiv. Analog folgt, daß L homogen ist, also $L \in \operatorname{Hom}(V, W)$. Offensichtlich gilt $L(v_i) = w_i$ für alle $i \in \{1, \ldots, n\}$.

Bem.: Eine (4.6) entsprechende Aussage gilt auch im Fall $\dim V = \infty$.

(4.7) Folgerung. Seien V, W K-Vektorräume, $\dim V = n < \infty$. Dann gilt:

(i) V ist isomorph zu K^n.
(ii) V ist isomorph zu $W \Leftrightarrow \dim V = \dim W$.

Bew.:

(i) Sei $(v_1, \dots, v_n)$ geordnete Basis von V. Nach (4.6) existiert ein $L \in \mathrm{Hom}(V, K^n)$ mit $L(v_1) = e_1, \dots, L(v_n) = e_n$. Nach (4.5) ist L Isomorphismus.

(ii) Analog folgt aus $\dim V = \dim W$, daß V und W isomorph sind. Ist umgekehrt $L : V \to W$ Isomorphismus, so folgt aus (4.5), daß $\dim V = \dim W$ gilt.

(4.8) Dimensionssatz für lineare Abbildungen. *Ist* $L \in \mathrm{Hom}(V, W)$ *und* $\dim V < \infty$*, so gilt* $\dim(\ker L) < \infty$*,* $\dim(\mathrm{im}\, L) < \infty$ *und*

$$\dim V = \dim(\ker L) + \dim(\mathrm{im}\, L).$$

Speziell folgt: Ist $\dim V = \dim W$*, so gilt:* L *injektiv* $\Leftrightarrow$ L *surjektiv* $\Leftrightarrow$ L *bijektiv.*

Bew.: Aus (3.21) folgt die Existenz eines zu $\ker L$ komplementären Unterraums U von V, d.h. $\ker L \oplus U = V$ und

$$\dim(\ker L) + \dim U = \dim V.$$

Wir zeigen, daß $L|U : U \to \mathrm{im}\, L$ ein Isomorphismus ist und damit $\dim U = \dim(\mathrm{im}\, L)$, woraus mit der vorangehenden Gleichung die Behauptung folgt.

$L|U$ ist injektiv, da $\ker(L|U) = \ker L \cap U = \{\underline{0}\}$, vgl. (4.3). Zu jedem $w \in \mathrm{im}\, L$ existiert ein $v \in V$ mit $L(v) = w$. Wegen $\ker L \oplus U = V$ existieren $v_1 \in \ker L$ und $u \in U$, so daß $v = v_1 + u$ gilt. Dann folgt $w = L(v) = L(v_1) + L(u)$, d.h. $w \in L(U)$. Das zeigt, daß $L|U : U \to \mathrm{im}\, L$ surjektiv ist.

Anwendung:

Sei I die linke Seite eines linearen Gleichungssystems mit n Unbekannten und k Gleichungen und Koeffizienten $a_{ij} \in K$. Sei $L : K^n \to K^k$ die zugehörige lineare Abbildung, vgl. Bem. nach (4.3). Dann gilt: $L_{I^{\mathrm{hom}}} = \ker L$ und

$$\{b = (b_1, \dots, b_k) \in K^k \mid \exists \text{ Lösung von } I \text{ mit der rechten Seite } b\} = \mathrm{im}\, L.$$

Aus (4.8) folgt: $\dim L_{I^{\mathrm{hom}}} = n - \dim(\mathrm{im}\, L) \geq n - k$. Das wurde schon in (3.25) gezeigt.

$$\begin{aligned} I \text{ ist für beliebige rechte Seiten } (b_1, \dots, b_k) \text{ lösbar} \quad &\Leftrightarrow \quad \dim(\mathrm{im}\, L) = k \\ &\Leftrightarrow \quad \dim(L_{I^{\mathrm{hom}}}) = n - k \end{aligned}$$

d.h. ist $n \geq k$ und ist I für beliebige rechte Seite lösbar, so ist die Lösungsmenge für jede rechte Seite ein $(n-k)$-dimensionaler affiner Unterraum. Ist $k = n$, so gilt: I ist für beliebige rechte Seiten $(b_1, \dots, b_n)$ lösbar $\Leftrightarrow L_{I^{\mathrm{hom}}} = \{0\} \Leftrightarrow I$ besitzt für jede rechte Seite $(b_1, \dots, b_n)$ genau eine Lösung.

Ist $L \in \mathrm{Hom}(V, W)$, $\alpha \in K$, so definieren wir:

$$\alpha L : V \to W \text{ durch: } \forall v \in V \text{ ist } (\alpha L)(v) := \alpha L(v) \in W.$$

Es gilt dann: $\alpha L \in \mathrm{Hom}(V, W)$, z.B. zeigen wir, daß αL additiv ist: $(\alpha L)(v_1 + v_2) = \alpha L(v_1 + v_2) = \alpha(L(v_1) + L(v_2)) = \alpha L(v_1) + \alpha L(v_2) = (\alpha L)(v_1) + (\alpha L)(v_2)$. Sind $L_1, L_2 \in \mathrm{Hom}(V, W)$, so definieren wir:

$$L_1 + L_2 : V \to W \text{ durch: } \forall v \in V \text{ ist } (L_1 + L_2)(v) := L_1(v) + L_2(v) \in W.$$

Es gilt dann: $L_1 + L_2 \in \mathrm{Hom}(V, W)$, z.B. zeigen wir, daß $L_1 + L_2$ homogen ist: $(L_1 + L_2)(\alpha v) = L_1(\alpha v) + L_2(\alpha v) = \alpha L_1(v) + \alpha L_2(v) = \alpha(L_1(v) + L_2(v)) = \alpha(L_1(v) + L_2(v)) = \alpha(L_1 + L_2)(v)$.

(4.9) Satz. *Mit dieser Addition und dieser Multiplikation mit Körperelementen $\alpha \in K$ ist $\mathrm{Hom}(V, W)$ ein K-Vektorraum.*

Bew.: Es muß nachgeprüft werden, daß $(\mathrm{Hom}(V, W), +)$ eine abelsche Gruppe ist und daß die Bedingungen (3.1)(i)-(iv) erfüllt sind. Das ist ebenso langwierig wie langweilig, aber man sollte es (einmal) getan haben. Wir führen exemplarisch einige der notwendigen Schritte durch: Das neutrale Element der Addition in $\mathrm{Hom}(V, W)$ ist die Abbildung $L \in \mathrm{Hom}(V, W)$, die jedes $v \in V$ auf $\underline{0} \in W$ abbildet, genannt die 0-Abbildung. Das additive Inverse zu $L \in \mathrm{Hom}(V, W)$ ist $-L \in \mathrm{Hom}(V, W)$, definiert durch $(-L)(v) := -L(v)$.

Um (3.1)(iii) zu zeigen, seien $\alpha \in K$, $L_1, L_2 \in \mathrm{Hom}(V, W)$. Für beliebiges $v \in V$ gilt:

$$\begin{aligned}(\alpha(L_1 + L_2))(v) &= \alpha((L_1 + L_2)(v)) = \alpha(L_1(v) + L_2(v)) = \alpha L_1(v) + \alpha L_2(v)\\ &= (\alpha L_1)(v) + (\alpha L_2)(v) = (\alpha L_1 + \alpha L_2)(v).\end{aligned}$$

Also gilt: $\alpha(L_1 + L_2) = \alpha L_1 + \alpha L_2$.

(4.1) Def. Eine $m \times n$-Matrix A über dem Körper K ist ein rechteckiges Schema, das aus $m \cdot n$ Körperelementen $a_{ij} \in K$ besteht:

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$$

A heißt quadratisch $\Leftrightarrow m = n$.

Die Vektoren $(a_{11}, \dots, a_{1n}) \in K^n, \dots, (a_{m1}, \dots, a_{mn}) \in K^n$ heißen die Zeilenvektoren von A. Die $(m \times 1)$-Matrizen $\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \dots, \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$ heißen die Spaltenvektoren von A.

Mit $K^{m \times n}$ wird die Menge der $(m \times n)$-Matrizen über K bezeichnet.

Bsp.: Die linke Seite eines linearen Gleichungssystems mit m Gleichungen und n Unbekannten (in K) ist durch eine Matrix $A \in K^{m \times n}$ gegeben.

(4.11) Def. Sei $L \in \mathrm{Hom}(V, W)$, $\dim V = n$, $\mathcal{G} = (v_1, \dots, v_n)$ geordnete Basis von V, $\dim W = m$, $\mathcal{G}' = (w_1, \dots, w_m)$ geordnete Basis von W. Dann heißt die Matrix $A = (a_{ij}) \in K^{m \times n}$, die durch

$$L(v_j) =: \sum_{i=1}^{m} a_{ij} w_i \qquad \text{für } 1 \le j \le n,$$

definiert ist, die Matrix von L bezüglich $\mathcal{G}$ und $\mathcal{G}'$,

$$A := \mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}'}(L).$$

(4.12) Folgerung. Die Abbildung $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}'} : \mathrm{Hom}(V,W) \to K^{m\times n}$ ist bijektiv.

Bew.: Injektivität: Seien $L_1, L_2 \in \mathrm{Hom}(V,W)$ und $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}'}(L_1) = \mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}'}(L_2)$. Dann gilt: $L_1(v_j) = \sum\limits_{i=1}^{m} a_{ij} w_i = L_2(v_j)$ für $j \in \{1,\ldots,n\}$. Aus (4.6) folgt dann: $L_1 = L_2$.

Surjektivität: Sei $A = (a_{ij}) \in K^{m\times n}$ gegeben. Wir definieren für $j \in \{1,\ldots,n\}$:

$$\overline{w}_j := \sum_{i=1}^{m} a_{ij} w_i \in W.$$

Nach (4.6) existiert (genau) ein $L \in \mathrm{Hom}(V,W) : Lv_j = \overline{w}_j$. Dann gilt: $M_{\mathcal{G}}^{\mathcal{G}'}(L) = A$.

Bem.: Da eine Matrix $A \in K^{m\times n}$ im Grunde nur ein etwas anders notiertes Element von $K^{m\cdot n}$ ist, läßt sich leicht eine "natürliche" K-Vektorraumstruktur auf $K^{m\times n}$ definieren: $\alpha \in K$, $A = (a_{ij}) \in K^{m\times n} \to \alpha A := (\alpha a_{ij}) \in K^{m\times n}$, d.h.

$$\alpha A := \begin{pmatrix} \alpha a_{11} & \ldots & \alpha a_{1n} \\ \vdots & & \vdots \\ \alpha a_{m1} & \ldots & \alpha a_{mn} \end{pmatrix}$$

$A = (a_{ij}) \in K^{m\times n}$, $B = (b_{ij}) \in K^{m\times n} \to A + B := (a_{ij} + b_{ij}) \in K^{m\times n}$, d.h.

$$A + B := \begin{pmatrix} a_{11} + b_{11} & \ldots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \ldots & a_{mn} + b_{mn} \end{pmatrix}$$

(4.13) Fakt. Mit diesen Operationen ist $K^{m\times n}$ ein K-Vektorraum der Dimension $m\cdot n$. Die Matrizen $E_{ij} \in K^{m\times n}$, $1 \leq i \leq m$, $1 \leq j \leq n$

$$E_{ij} = \begin{pmatrix} 0 & \ldots & \ldots & 0 & \ldots & \ldots & 0 \\ \vdots & & & \vdots & & & \vdots \\ \vdots & & & 0 & & & \vdots \\ 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ \vdots & & & 0 & & & \vdots \\ \vdots & & & \vdots & & & \vdots \\ 0 & \ldots & \ldots & 0 & \ldots & \ldots & 0 \end{pmatrix}$$

(wobei die 1 am Schnittpunkt der i'ten Zeile und der j'ten Spalte sitzt) bilden eine Basis, die kanonische Basis, von $K^{m\times n}$.

Bem.: Ist $A = (a_{ij}) \in K^{m\times n}$, so gilt $A = \sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n} a_{ij} E_{ij}$.

(4.14) Folgerung. $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}'} : \mathrm{Hom}(V,W) \to K^{m\times n}$ ist ein Isomorphismus.

Bew.: (4.12) besagt, daß $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}'}$ bijektiv ist. Es bleibt zu zeigen, daß $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}'}$ homogen und additiv ist:

$$\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}'}(L) = (a_{ij}) \in K^{m\times n} \Leftrightarrow \forall j \in \{1,\dots,n\} : L(v_j) = \sum a_{ij} w_i.$$

Dann gilt für $j \in \{1,\dots,n\}$:

$$(\alpha L)(v_j) = \sum_{i=1}^{m} \alpha a_{ij} w_i.$$

Also $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}'}(\alpha L) = \alpha(a_{ij}) = \alpha \,\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}'}(L)$.

Analog folgt: $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}'}(L_1 + L_2) = \mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}'}(L_1) + \mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}'}(L_2)$.

Frage: Gegeben 3 K-Vektorräume V, W, Z mit $\dim V = n$, $\dim W = m$, $\dim Z = k$ und geordneten Basen $\mathcal{G}, \mathcal{G}', \mathcal{G}''$ und $L \in \mathrm{Hom}(V, W)$, $J \in \mathrm{Hom}(W, Z)$. Wie hängt $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}''}(J \circ L)$ mit $\mathrm{Mat}_{\mathcal{G}'}^{\mathcal{G}''}(J)$ und $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}'}(L)$ zusammen?

(4.15) Satz. *Ist* $A = (a_{hl}) = \mathrm{Mat}_{\mathcal{G}'}^{\mathcal{G}''}(J) \in K^{k\times m}$, $B = (b_{ij}) = \mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}'}(L) \in K^{m\times n}$, *so ist* $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}''}(J \circ L)$ *durch die Matrix* $C = (c_{hj}) \in K^{k\times n}$ *gegeben mit* $c_{hj} = \sum_{i=1}^{m} a_{hi} b_{ij}$ *für* $1 \le h \le k$, $1 \le j \le n$.

Vorbemerkung: Assoziativ-, Kommutativgesetz von $+$, Distributivgesetz $\Rightarrow$

$$\sum_{i=1}^{m} b_i \left(\sum_{j=1}^{k} a_{ji} v_j \right) = \sum_{i=1}^{m} \left(\sum_{j=1}^{k} b_i a_{ji} v_j \right) = \sum_{j=1}^{k} \left(\sum_{i=1}^{m} b_i a_{ji} \right) v_j.$$

Bsp.: $b_1(a_{11}v_1 + a_{21}v_2) + b_2(a_{12}v_1 + a_{22}v_2) = (b_1 a_{11} + b_2 a_{12})v_1 + (b_1 a_{21} + b_2 a_{22})v_2$.

Bew. von (4.15): Sei $\mathcal{G} = (v_1, \dots, v_n)$, $\mathcal{G}' = (w_1, \dots, w_m)$, $\mathcal{G}'' = (z_1, \dots, z_k)$ und

$$L(v_j) = \sum_{i=1}^{m} b_{ij} w_i, \; J(w_i) = \sum_{h=1}^{k} a_{hi} z_h.$$

Dann gilt:
$$\begin{aligned} (J \circ L)(v_j) &= J(L(v_j)) = J\left(\sum_{i=1}^{m} b_{ij} w_i \right) = \sum_{i=1}^{m} b_{ij} J(w_i) \\ &= \sum_{i=1}^{m} b_{ij} \left(\sum_{h=1}^{k} a_{hi} z_h \right) \overset{\text{Vorbem.}}{=} \sum_{h=1}^{k} \left(\sum_{i=1}^{m} a_{hi} b_{ij} \right) z_h. \end{aligned}$$

Definition. Ist $A = (a_{hl}) \in K^{k\times m}$, $B = (b_{ij}) \in K^{m\times n}$, so heißt die Matrix $C = (c_{hj}) \in K^{k\times n}$ mit

$$c_{hj} := \sum_{j=1}^{m} a_{hi} b_{ij} \qquad \text{für } 1 \le h \le k, 1 \le j \le n$$

das Produkt der Matrizen A und B, $C =: AB$. Dann gilt mit den Bezeichnungen von (4.15):

$$\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}''}(J \circ L) = \mathrm{Mat}_{\mathcal{G}'}^{\mathcal{G}''}(J)\, \mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}'}(L).$$

Vorsicht: Damit AB definiert ist, muß gelten

$$\# \text{ Spalten von } A = \# \text{ Zeilen von } B$$

Explizit:

$$\underbrace{\begin{pmatrix} 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}}_{\in K^{2\times 3}} \underbrace{\begin{pmatrix} 3 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 2 & 1 & 1 & 0 \end{pmatrix}}_{\in K^{3\times 4}} = \underbrace{\begin{pmatrix} 8 & 3 & 3 & 0 \\ 4 & 2 & 1 & 1 \end{pmatrix}}_{\in K^{2\times 4}}$$

$$\underbrace{\begin{pmatrix} 3 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}}_{\in K^{2\times 3}} \underbrace{\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}}_{\in K^{3\times 1}} = \underbrace{\begin{pmatrix} 6 \\ 2 \end{pmatrix}}_{\in K^{2\times 1}}$$

(4.16) Folgerung.

(i) Die Matrizenmultiplikation ist assoziativ: $A(BC) = (AB)C$

(ii) $A(B_1 + B_2) = AB_1 + AB_2$, $(A_1 + A_2)B = A_1B + A_2B$ (Distributivgesetze)

Bew.:

(i) Nachrechnen. Oder: $A = \mathrm{Mat}^{\mathcal{G}_4}_{\mathcal{G}_3}(L_3), B = \mathrm{Mat}^{\mathcal{G}_3}_{\mathcal{G}_2}(L_2), C = \mathrm{Mat}^{\mathcal{G}_2}_{\mathcal{G}_1}(L_1)$ und $(L_3 \circ L_2) \circ L_1 = L_3 \circ (L_2 \circ L_1) \overset{(4.15)}{\Rightarrow}$ Beh.

(ii) Nachrechnen. Oder: $J \circ (L_1 + L_2) = (J \circ L_1) + (J \circ L_2)$ (J linear!) $\overset{(4.14)+(4.15)}{\Rightarrow} A(B_1 + B_2) = AB_1 + AB_2$

(4.17) Abhängigkeit von der Wahl der Basen. Seien $\mathcal{G}, \overline{\mathcal{G}}$ geordnete Basen von V, $\mathcal{G}', \overline{\mathcal{G}}'$ geordnete Basen von V' und $L \in \mathrm{Hom}(V, V')$. Dann gilt:

$$\mathrm{Mat}^{\overline{\mathcal{G}}'}_{\overline{\mathcal{G}}}(L) = \mathrm{Mat}^{\overline{\mathcal{G}}'}_{\mathcal{G}'}(\mathrm{id}_{V'}) \cdot \mathrm{Mat}^{\mathcal{G}'}_{\mathcal{G}}(L) \cdot \mathrm{Mat}^{\mathcal{G}}_{\overline{\mathcal{G}}}(\mathrm{id}_V)$$

Bew.: Wende (4.15) auf $L = \mathrm{id}_{V'} \circ L \circ \mathrm{id}_V$ bzgl. der Basen $\overline{\mathcal{G}}, \mathcal{G}, \mathcal{G}', \overline{\mathcal{G}}'$ an.

Bem.: $\mathcal{G} = (v_1, \ldots, v_n)$, $\overline{\mathcal{G}} = (\overline{v}_1, \ldots, \overline{v}_n)$. Dann

$$\mathrm{Mat}^{\overline{\mathcal{G}}}_{\mathcal{G}}(\mathrm{id}_V) = (a_{ij})_{1\le i,j\le n} \Longleftrightarrow v_j = \sum_{i=1}^{n} a_{ij}\overline{v}_i \qquad \text{für } 1 \le j \le n.$$

1. Spezialfall: $V' = V$

Ein Homomorphismus $L : V \to V$ heißt <u>Endomorphismus</u>, $\mathrm{End}(V) := \mathrm{Hom}(V, V)$. Sei $\dim V = n$. Für jede geordnete Basis $\mathcal{G}$ von V gilt:

$$\mathrm{Mat}^{\mathcal{G}}_{\mathcal{G}}(\mathrm{id}_V) = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} =: E_n \in K^{n\times n}$$

E_n heißt die n-dimensionale Einheitsmatrix.

$(\mathrm{End}(V), +, \circ)$ ist Ring (mit 1). Neutrales Element von $\circ$ ist id_V. $(K^{n\times n}, +, \cdot)$ ist Ring (mit 1). Neutrales Element von $\cdot$ ist E_n.

$\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}} : \mathrm{End}(V) \to K^{n\times n}$ ist Isomorphismus von Ringen, vgl. (4.14) und (4.15).

Ein bijektiver Endomorphismus heißt Automorphismus, $\mathrm{Aut}(V) := \{L \mid L \in \mathrm{End}(V) \text{ bijektiv}\}$.
$(\mathrm{Aut}(V), \circ)$ ist Gruppe (Untergruppe von $(S_V, \circ)$, vgl. Bsp. 5) nach (2.1)).
Denn:

(i) $\mathrm{id}_V \in \mathrm{Aut}(V)$ (neutrales Element bezüglich $\circ$)
(ii) $L \in \mathrm{Aut}(V) \Rightarrow L^{-1} \in \mathrm{Aut}(V)$ (L^{-1} = Umkehrabbildung von L)
L^{-1} ist bijektiv. Linearität von L^{-1}:

$$\left.\begin{array}{lcl} L(L^{-1}(v+w)) & = & v+w \\ L(L^{-1}(v)+L^{-1}(w)) & = & L(L^{-1}(v)) + L(L^{-1}(w)) = v+w \end{array}\right\} \Rightarrow \begin{array}{l} L^{-1}(v+w) \\ = L^{-1}(v) + L^{-1}(w) \end{array}$$

ebenso $L^{-1}(\alpha v) = \alpha L^{-1}(v)$
(iii) $L_1, L_2 \in \mathrm{Aut}(V) \Rightarrow L_1 \circ L_2 \in \mathrm{Aut}(V)$
$(L_1 \circ L_2)(\alpha v) = L_1(L_2(\alpha v)) = L_1(\alpha L_2(v)) = \alpha L_1(L_2(v)) = \alpha((L_1 \circ L_2)(v))$

Analoges Objekt in $K^{n\times n}$:

$$\mathrm{GL}_n(K) := \{A \in K^{n\times n} \mid \exists B \in K^{n\times n} : BA = E_n\}$$

Def.: Seien $(G, \top), (G', \top')$ Gruppen. $H : G \to G'$ heißt Gruppenhomomorphismus $\Leftrightarrow$ $\forall g, h \in G$: $H(g \top h) = H(g) \top' H(h)$
H Gruppenisomorphismus $\Leftrightarrow$ H bijektiver Gruppenhomomorphismus

(4.18) Satz. $(\mathrm{GL}_n(K), \cdot)$ *ist Gruppe und* $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}} : \mathrm{Aut}(V) \to \mathrm{GL}_n(K)$ *ist Gruppenisomorphismus.*

Bew.:

(i) Zeige: $L \in \mathrm{Aut}(V) \Rightarrow \mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L) \in \mathrm{GL}_n(K)$
$E_n = \mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(\mathrm{id}_V) = \mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L^{-1} \circ L) \overset{(4.15)}{=} \underbrace{\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L^{-1})}_{\in K^{n\times n}} \cdot \mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L)$
(ii) Zeige: $\forall A \in \mathrm{GL}_n(K) \exists L \in \mathrm{Aut}(V) : \mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L) = A$

$A \in \mathrm{GL}_n(K) \Rightarrow \exists B \in K^{n\times n} : BA = E_n$. Nach (4.14) existieren $J, L \in \mathrm{End}(V)$: $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(J) = B$, $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L) = A$. Also: $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(J \circ L) = BA = E_n$. Da $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}$ injektiv ist, folgt $J \circ L = \mathrm{id}_V$, speziell L ist injektiv. Nach (4.8) ist L bijektiv, also $L \in \mathrm{Aut}(V)$ und $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L) = A$, außerdem folgt $J \in \mathrm{Aut}(V)$ und damit nach (i) $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(J) = B \in \mathrm{GL}_n(K)$.

(i)+(ii) $\Rightarrow \mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}} : \mathrm{Aut}(V) \to \mathrm{GL}_n(K)$ bijektiv.

(4.15) $\Rightarrow$ $(\mathrm{GL}_n(K),\cdot)$ ist Gruppe und $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}$ Gruppenisomorphismus.

Bez.: Die Elemente von $\mathrm{GL}_n(K)$ heißen reguläre (oder nicht-ausgeartete oder invertierbare) $(n\times n)$-Matrizen. Ist $A \in \mathrm{GL}_n(K)$, $B \in K^{n\times n}$ und $BA = E_n$, so folgt (vgl. Bew. von (4.18)) $B \in \mathrm{GL}_n(K)$ und damit (da $\mathrm{GL}_n(K)$ eine Gruppe ist) $B = A^{-1}$, speziell auch $AB = E_n$. A^{-1} heißt die zu $A \in \mathrm{GL}_n(K)$ inverse Matrix.

(4.19) Folgerung (aus (4.17)): Seien $\mathcal{G}, \overline{\mathcal{G}}$ geordnete Basen von V, $L \in \mathrm{End}(V)$ und $P := \mathrm{Mat}_{\overline{\mathcal{G}}}^{\mathcal{G}}(\mathrm{id}_V)$. Dann gilt $P \in \mathrm{GL}_n(K)$, $P^{-1} = \mathrm{Mat}_{\mathcal{G}}^{\overline{\mathcal{G}}}(\mathrm{id}_V)$ und

$$\mathrm{Mat}_{\overline{\mathcal{G}}}^{\overline{\mathcal{G}}}(L) = P^{-1}\,\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L)P$$

Bew.: $\mathrm{Mat}_{\mathcal{G}}^{\overline{\mathcal{G}}}(\mathrm{id}_V)\,\mathrm{Mat}_{\overline{\mathcal{G}}}^{\mathcal{G}}(\mathrm{id}_V) = \mathrm{Mat}_{\overline{\mathcal{G}}}^{\overline{\mathcal{G}}}(\mathrm{id}_V) = E_n \Rightarrow P \in \mathrm{GL}_n(K)$ und $P^{-1} = \mathrm{Mat}_{\mathcal{G}}^{\overline{\mathcal{G}}}(\mathrm{id}_V)$. Die letzte Gleichung folgt aus (4.17).

Bez.: Zwei Matrizen A, $\tilde{A} \in K^{n\times n}$ heißen ähnlich $\Leftrightarrow \exists P \in \mathrm{GL}_n(K) : P^{-1}AP = \tilde{A}$. "Ähnlich" ist eine Äquivalenzrelation. Nichttriviales Problem: Man finde zu gegebenem $A \in K^{n\times n}$ eine "möglichst einfache" ähnliche Matrix (oder äquivalent: Man finde zu $L \in \mathrm{End}(V)$ eine Basis $\mathcal{G}$, so daß $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L)$ möglichst einfach ist). Der im allgemeinen nicht erreichbare Idealfall ist:

$$\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L) = \text{ Diagonalmatrix } = \begin{pmatrix} d_1 & & & \\ & \ddots & 0 & \\ & 0 & \ddots & \\ & & & d_n \end{pmatrix}$$

2. Spezialfall:
$$\begin{aligned} V &= K^n \text{ mit Basis } \mathcal{G} = (e_1,\ldots,e_n) \\ V' &= K^m \text{ mit Basis } \mathcal{G}' = (e_1,\ldots,e_m) \end{aligned}$$

Bez.: Im Fall dieser natürlichen Basen schreiben wir für $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}'}$ kurz $\mathrm{Mat} : \mathrm{Hom}(K^n, K^m) \to K^{m\times n}$. Wir identifizieren K^n (bzw. K^m) mit den "einspaltigen" Matrizen $K^{n\times 1}$ (bzw. $K^{m\times 1}$) durch den Vektorraumisomorphismus

$$(x_1,\ldots,x_n) \in K^n \to \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^{n\times 1}.$$

(4.20) Satz. *Mit diesen Identifikationen gilt: Ist* $L \in \mathrm{Hom}(K^n, K^m)$ *und* $A = (a_{ij}) = \mathrm{Mat}(L)$*, so gilt für alle* $x = (x_1,\ldots,x_n) \in K^n$*:*

$$L(x) = A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Bew.: $\mathrm{Mat}(L) = (a_{ij})_{\substack{1\le i\le m \\ 1\le j\le n}} \Leftrightarrow L(e_j) = \sum\limits_{j=1}^{m} a_{ij}e_i$ für $1 \le j \le n$.

$$\begin{aligned}\text{Also:}\quad L(x) &= L\left(\sum_{j=1}^n x_j e_j\right) = \sum_{j=1}^n x_j L(e_j) = \sum_{j=1}^n x_j \left(\sum_{i=1}^m a_{ij} e_i\right) \overset{\text{vgl. Bew. von (4.15)}}{=} \\ &= \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} x_j\right) e_i, \text{ d.h.} \\ L(x) &= \left(\sum_{j=1}^n a_{1j}x_j, \sum_{j=1}^n a_{2j}x_j, \ldots, \sum_{j=1}^n a_{mj}x_j\right) \in K^m. \qquad \text{Andererseits:}\end{aligned}$$

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix} \in K^{m\times 1}.$$

Bem.: $L \in \mathrm{Hom}(K^n, K^m)$, $A = \mathrm{Mat}(L)$. Dann "ist" $L(e_j) \in K^m$ der j'te Spaltenvektor von A

$$L(e_j) = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} 0 \\ \cdot \\ 0 \\ 1 \\ 0 \\ \cdot \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \in K^{m\times 1}.$$

Speziell: Ist $A \in K^{n\times n}$, so gilt: $A \in \mathrm{GL}_n(K) \Leftrightarrow$ die Spaltenvektoren von A sind linear unabhängig.

Begründung: Die lineare Abbildung $L \in \mathrm{End}(K^n)$ mit $\mathrm{Mat}(L) = A$ bildet die Basis $e_1, \ldots, e_n$ auf die n Spaltenvektoren von A ab. Nach (4.5) ist L genau dann ein Automorphismus, wenn diese Spaltenvektoren eine Basis von $K^{n\times 1}$ bilden.

(4.21) Def.: Der <u>Rang</u> <u>$\mathrm{rg}(A)$</u> von $A \in K^{m\times n}$ ist die maximale Zahl linear unabhängiger Spaltenvektoren von A.

Bem.: 1) $0 \le \mathrm{rg}(A) \le \min\{n, m\}$. Begründung: Da A n Spaltenvektoren hat, gilt $\mathrm{rg}(A) \le n$. Da der Raum $K^{m\times 1} \simeq K^m$ der Spaltenvektoren m-dimensional ist, gilt $\mathrm{rg}(A) \le m$, vgl. (3.16) Austauschsatz von Steinitz.
2) Für $A \in K^{n\times n}$ gilt: $\mathrm{rg}(A) = n \Leftrightarrow A \in \mathrm{GL}_n(K)$.

(4.22) Fakt: Sei $L \in \mathrm{Hom}(K^n, K^m)$ und $A = \mathrm{Mat}(L)$. Dann gilt

$$\begin{aligned}\mathrm{rg}(A) &= \dim(\mathrm{span}\{\text{Spaltenvektoren von } A\}) = \dim(\mathrm{im}\, L) \\ &= n - \dim(\ker L) = n - \dim\{x \in K^{n\times 1} \mid Ax = 0\}.\end{aligned}$$

Bew.: Die 1. Gleichung folgt aus der Tatsache, daß die Dimension eines Vektorraums die maximale Anzahl linear unabhängiger Vektoren dieses Vektorraums ist. Die 2. Gleichung, folgt aus "$L(e_j) = j$'ter Spaltenvektor von A" und $\operatorname{im} L = \operatorname{span}\{L(e_j) \mid 1 \leq j \leq n\}$. Die 3. Gleichung folgt aus (4.8) und die 4. Gleichung folgt aus der dritten.

(4.23) Satz. *Für jedes $A \in K^{m\times n}$ ist $\operatorname{rg}(A)$ auch die maximale Zahl linear unabhängiger Zeilenvektoren von A ("Spaltenrang = Zeilenrang").*

Bew.: Siehe (4.25).

(4.24) Satz. *Sei $A = (a_{ij}) \in K^{m\times n}$ und I das lineare Gleichungssystem*

$$\mathrm{I} \qquad A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

und $(A|b) \in K^{m\times(n+1)}$ die Matrix $\begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix}$

Dann gilt: I besitzt eine Lösung $\Leftrightarrow \operatorname{rg}(A) = \operatorname{rg}(A|b)$

Bew.: $A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \Leftrightarrow \begin{pmatrix} a_{11}x_1 & +\ldots+ & a_{1n}x_n & = b_1 \\ \vdots & & \vdots & \\ a_{m1}x_1 & +\ldots+ & a_{mn}x_n & = b_m \end{pmatrix} \Leftrightarrow$

$$x_1\underbrace{\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}}_{=:A_1} + x_2\underbrace{\begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix}}_{=:A_2} + \ldots + x_n\underbrace{\begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}}_{=:A_n} = \underbrace{\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}}_{=b} \in K^{m\times 1}.$$

Also: I besitzt Lösung $\Leftrightarrow \operatorname{span}\{A_1, \ldots, A_n\} = \operatorname{span}\{A_1, \ldots, A_n, b\}$
$\Leftrightarrow \operatorname{rg}(A) = \operatorname{rg}(A|b)$

Konkrete Rechenverfahren:
(4.25) **Lösen eines linearen Gleichungssystems mit dem Gaußschen Algorithmus in Matrizenschreibweise**. (Zum Lösen von linearen Gleichungssystemen siehe auch (1.13), (1.15), (1.16), (3.25), (3.26) und (4.8) plus Anwendung).

Gegeben das lineare Gleichungssystem

$$\begin{array}{lcll} a_{11}x_1 & +\ldots+ & a_{1n}x_n & = b_1 \\ \vdots & & \vdots & \\ a_{m1}x_1 & +\ldots+ & a_{mn}x_n & = b_m. \end{array}$$

Setze

$$A_1 := \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix} \in K^{m\times(n+1)}.$$

1. Fall $a_{11} \neq 0$:

$$A_1 \to A_2 := \begin{pmatrix} 1 & \frac{a_{12}}{a_{11}} & \dots & \frac{a_{1n}}{a_{11}} & \frac{b_1}{a_{11}} \\ 0 & a_{22} - a_{21}\frac{a_{12}}{a_{11}} & \dots & a_{2n} - a_{21}\frac{a_{1n}}{a_{11}} & b_2 - a_{21}\frac{b_1}{a_{11}} \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & a_{m2} - a_{m1}\frac{a_{12}}{a_{11}} & \dots & a_{mn} - a_{m1}\frac{a_{1n}}{a_{11}} & b_m - a_{m1}\frac{b_1}{a_{11}} \end{pmatrix}$$

2. Fall: $a_{11} = 0$, aber es existiert $i > 1 : a_{i1} \neq 0$. Vertausche 1. und i'te Zeile von A_1. Dann 1. Fall.

3. Fall: $a_{i1} = 0$ für $1 \leq i \leq m$, d.h.

$$A_1 = \begin{pmatrix} 0\, a_{12} & \dots & b_1 \\ \vdots & & \vdots \\ 0\, a_{m2} & \dots & b_m \end{pmatrix} \in K^{m \times (n+1)}. \qquad \text{Betrachte } \tilde{A}_1 = \begin{pmatrix} a_{12} & \dots & b_1 \\ \vdots & & \vdots \\ a_{m2} & \dots & b_m \end{pmatrix} \in K^{m \times n}.$$

Dann 1. Fall (für $n-1$).

Iteriere! $\to$ Endergebnis ist eine Matrix $(\overline{A}|\overline{b}) \in K^{m \times (n+1)}$ in Treppenform, z.B.

$$\overline{A} = \begin{pmatrix} 1 & * & * & \cdot & \cdot & \cdot & * \\ 0 & 0 & 1 & * & \cdot & \cdot & * \\ 0 & 0 & 0 & 1 & * & \cdot & * \\ \cdot & \cdot & \cdot & 0 & & & \\ 0 & 0 & 0 & 0 & & & \end{pmatrix} \qquad \text{(hier ist } j_1 = 1,\ j_2 = 3,\ j_3 = 4\text{)},$$

d.h. es gibt Zahlen $k \leq \min\{m, n\}$ und $1 \leq j_1 < j_2 < \ldots < j_k \leq n$ ($\Rightarrow j_i \geq i$ für $i \in \{1, \ldots, k\}$) mit:

1) $a_{ij_i} = 1$ für $1 \leq i \leq k$ und $a_{ij} = 0$ für $1 \leq i \leq k$ und $j < j_i$.
2) $a_{ij} = 0$ für $i > k$.

Wegen (1.13) gilt für alle $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^{n \times 1}$ und alle $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in K^{m \times 1}$:

$$(I) \qquad Ax = b \Leftrightarrow \overline{A}x = \overline{b} \qquad (\overline{I})$$

Falls $b = 0$, so $\overline{b} = 0$. Deshalb folgt aus (4.22):

$$\operatorname{rg}(A) = n - \dim\{x \mid Ax = 0\} = n - \underbrace{\dim\{x \mid \overline{A}x = 0\}}_{=n-k} = \operatorname{rg}(\overline{A}) = k.$$

$$\begin{aligned} \text{Andererseits: } \operatorname{rg}(A) = k &= \dim(\operatorname{span}\{\text{Zeilenvektoren von } \overline{A}\}) \\ &= \dim(\operatorname{span}\{\text{Zeilenvektoren von } A\}) \end{aligned}$$

Daraus folgt (4.23).

I besitzt Lösung $\Leftrightarrow$ Für $i > k = \mathrm{rg}(A)$ gilt: $\overline{b}_i = 0$. Dann ist $\overline{I}$ rekursiv auflösbar, beginnend mit der k'ten Zeile, und die Lösungsmenge von I (=Lösungsmenge von $\overline{I}$) ist ein $(n-\mathrm{rg}(A))$-dimensionaler affiner Unterraum von K^n.

(4.26) **Berechnung der inversen Matrix mit dem Gaußschen Algorithmus**. *Gegeben* $A \in K^{n\times n}$. *Frage: Gilt* $A \in \mathrm{GL}_n(K)$. *Wenn ja, berechne* A^{-1}.

Betrachte $(A \mid E_n) = \left(\begin{array}{c|ccc} & 1 & & 0 \\ A & & \ddots & \\ & 0 & & 1 \end{array}\right) \in K^{n\times(2n)}$ und wende (4.25) analog auf $(A \mid E_n)$ an. Man erhält $(\overline{A} \mid \overline{B}) \in K^{n\times(2n)}$, und es gilt $A \in \mathrm{GL}_n(K) \Leftrightarrow \overline{A} \in \mathrm{GL}_n(K) \Leftrightarrow \overline{a}_{ii} = 1$ für $1 \le i \le n$. Ausserdem gilt $\overline{a}_{ij} = 0$ für $i > j \ldots$ (d.h. $\overline{A} = \begin{pmatrix} 1 & * & \ldots & * \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \ldots & 0 & 1 \end{pmatrix}$). Durch geeignete Additionen der letzten, der vorletzten, …, der 2. Zeile verwandelt man $(\overline{A} \mid \overline{B})$ in (En | $\overline{\overline{B}}$), und es gilt für $j \in \{1, \ldots, n\}$ und alle $x \in K^{n\times 1}$:

$$Ax = e_j \Leftrightarrow \mathrm{En}\, x (= x) = j\text{'te Zeile von } \overline{\overline{B}} = \overline{\overline{B}} e_j.$$

Also: $A\overline{\overline{B}}e_j = Ax = e_j$ für $j \in \{1, \ldots, n\}$ und damit $A\overline{\overline{B}} = \mathrm{En}$. Also gilt:

$$\overline{\overline{B}} \text{ ist zu } A \text{ inverse Matrix.}$$

Zwei explizite Beispiele zu den Rechenverfahren.

Zu (4.25):

$$\mathrm{I} \qquad \begin{array}{rcrcrcrcr} x_1 & + & x_2 & - & x_3 & - & x_4 & = & -2 \\ 3x_1 & + & 3x2 & - & x_3 & - & 2x_4 & = & 6 \\ -x_1 & - & x_2 & + & 3x_3 & + & 2x_4 & = & 14 \end{array}$$

$$(A|b) = \left(\begin{array}{rrrr|r} 1 & 1 & -1 & -1 & -2 \\ 3 & 3 & -1 & -2 & 6 \\ -1 & -1 & 3 & 2 & 14 \end{array}\right) \qquad \begin{array}{rr} -3\ | & +\ | \\ \hookleftarrow| & | \\ & \hookleftarrow| \end{array}$$

$$\left(\begin{array}{rrrr|r} 1 & 1 & -1 & -1 & -2 \\ 0 & 0 & 2 & 1 & 12 \\ 0 & 0 & 2 & 1 & 12 \end{array}\right) \qquad \begin{array}{rl} & \\ -\ | & : \frac{1}{2} \\ \hookleftarrow| & \end{array}$$

$$\left(\begin{array}{rrrr|r} 1 & 1 & -1 & -1 & -2 \\ 0 & 0 & 1 & \frac{1}{2} & 6 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \qquad \begin{array}{r} \leftarrow| \\ +| \\ \end{array}$$

$$\left(\begin{array}{rrrr|r} 1 & 1 & 0 & -\frac{1}{2} & 4 \\ 0 & 0 & 1 & \frac{1}{2} & 6 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Es folgt: $\mathrm{rg}(A) = 2$, I ist lösbar, der Lösungsraum L_I ist ein 2-dimensionaler affiner Unterraum.

Explizite Lösung:

$$\begin{array}{rcl} & & x_4 =: \lambda \\ x_3 + \frac{1}{2}x_4 = 6 & \Rightarrow & x_3 = 6 - \frac{1}{2}\lambda \\ & & x_2 =: \mu \\ x_1 + x_2 \quad - \frac{1}{2}x_4 = 4 & \Rightarrow & x_1 = \frac{1}{2}\lambda - \mu + 4 \end{array}$$

$$\begin{aligned} L_I &= \{(\tfrac{1}{2}\lambda - \mu + 4, \mu, 6 - \tfrac{1}{2}\lambda, \lambda) \mid (\lambda, \mu) \in \mathbb{R}^2\} \\ &= \{(4,0,6,0) + \lambda(\tfrac{1}{2}, 0, -\tfrac{1}{2}, 1) + \mu(-1,1,0,0) \mid (\lambda,\mu) \in \mathbb{R}^2\} \\ &= (4,0,6,0) + \mathrm{span}\{(\tfrac{1}{2}, 0, -\tfrac{1}{2}, 1), (-1,1,0,0)\} \subseteq \mathbb{R}^4 \end{aligned}$$

Zu (4.26)

Berechnung von A^{-1} für $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 1 \end{pmatrix} \in \mathbb{Q}^{3\times 3}$

$$\left(\begin{array}{rrr|rrr} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 & 1 & 0 \\ 1 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \qquad \begin{array}{rr} -\,| & -\,| \\ \leftarrow| & | \\ & \leftarrow| \end{array}$$

$$\left(\begin{array}{rrr|rrr} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \end{array}\right) \qquad \begin{array}{rr} & \\ +| & \cdot(-1) \\ \leftarrow| & \end{array}$$

$$\left(\begin{array}{rrr|rrr} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 & 0 \\ 0 & 0 & 1 & -2 & 1 & 1 \end{array}\right) \qquad \begin{array}{rr} & \leftarrow| \\ \leftarrow| & | \\ +| & -| \end{array}$$

$$\left(\begin{array}{rrr|rrr} 1 & 1 & 0 & 3 & -1 & -1 \\ 0 & 1 & 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & -2 & 1 & 1 \end{array}\right) \qquad \begin{array}{r} \leftarrow| \\ -| \\ \end{array}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 4 & -1 & -2 \\ 0 & 1 & 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & -2 & 1 & 1 \end{array}\right) \qquad \Rightarrow \qquad A^{-1} = \begin{pmatrix} 4 & -1 & -2 \\ -1 & 0 & 1 \\ -2 & 1 & 1 \end{pmatrix} \in \mathrm{GL}_3(\mathbb{Q}) \subseteq \mathbb{Q}^{3\times 3}.$$

Speziell folgt: $A \in \mathrm{GL}_3(\mathbb{Q})$.

5. Die Determinante

Motivation: Ist V ein $\mathbb{R}$-Vektorraum der Dimension n und sind $v_1, \ldots, v_n$ linear unabhängige Vektoren in V, so heißt

$$P(v_1, \ldots, v_n) = \left\{ \sum_{i=1}^{n} t_i v_i \mid \forall i \in \{1, \ldots, n\} : 0 \leq t_i \leq 1 \right\}$$

das von $v_1, \ldots, v_n$ aufgespannte Parallelotop.

Bsp.: $n = 2$: Parallotop = Parallelogramm.

Wie es bisher für uns keinen mathematischen Sinn hat von der "Länge eines Vektors $v \in V$" zu sprechen (dazu benötigt man eine zusätzliche "mathematische Struktur" - eine Norm oder ein Skalarprodukt auf V), so benötigt man auch eine zusätzliche Struktur auf V, um "das" Volumen eines Parallelotops definieren zu können. Diese zusätzliche Struktur auf dem $\mathbb{R}$-Vektorraum V heißt eine Determinantenform und ist eine Funktion $D : V^n = \underbrace{V \times \ldots \times V}_{n\text{-mal}} \to \mathbb{R}$, die einer geordneten Basis $(v_1, \ldots, v_n)$ von V eine reelle Zahl $(\neq 0)$ zuordnet, deren Betrag wir als Definition für das Volumen von $P(v_1, \ldots, v_n)$ nehmen wollen. (Es stellt sich heraus, daß die Funktion D, die positive und negative Werte annimmt, das mathematisch natürliche Objekt ist. Das Vorzeichen von $D(v_1, \ldots, v_n)$ sagt etwas über die "Orientierung" der geordneten Basis $(v_1, \ldots, v_n)$ aus, d.h. im Fall $V = \mathbb{R}^3$ (und in physikalischer Sprechweise) ob (v_1, v_2, v_3) ein Rechts- oder ein Linkssystem ist). Damit man so zu einem vernünftigen Volumenbegriff kommt, wird D besondere Eigenschaften haben müssen. Es ist eine erfreuliche Tatsache, daß D durch folgende natürliche Forderungen nahezu eindeutig (d.h. bis auf Multiplikation mit einer reellen Zahl $\neq 0$) bestimmt ist.

(V_1) $(v_1, \ldots, v_n)$ Basis $\Rightarrow D(v_1, \ldots, v_n) \neq 0$.

(V_2) Für alle $i \in \{1, \ldots, n\}, (v_1, \ldots, v_n) \in V^n$ und $r \in \mathbb{R}$ gilt:

$$D(v_1, \ldots, v_{i-1}, rv_i, v_{i+1}, \ldots, v_n) = rD(v_1, \ldots, v_n)$$

(V_3) Für alle $1 \leq i < j \leq n$ und alle $(v_1, \ldots, v_n) \in V^n$ gilt

$$D(v_1, \ldots, v_{i-1}, v_i + v_j, v_{i+1}, \ldots, v_n) = D(v_1, \ldots, v_n) = D(v_1, \ldots, v_{j-1}, v_i + v_j, \ldots, v_n)$$

Dabei hat (V_2) die anschauliche Interpretation, daß das Volumen von $P(v_1, \ldots, v_{i-1}, rv_i, v_{i+1}, \ldots, v_n)$ das $|r|$-fache des Volumens von $P(v_1, \ldots, v_n)$ sein soll.

Die anschauliche Bedeutung von (V_3) ist die "Scherungsinvarianz" des Volumens.

(5.1) Def. Sei V ein K-Vektorraum, $k \in \mathbb{N}$. Eine Abbildung $f : V^k \to K$ heißt

(i) k-linear (k-Linearform), falls f in jedem Argument linear ist, d.h. falls für jedes $i \in \{1, \ldots, k\}$ und alle $v_1, \ldots, v_{i-1}, v, w, v_{i+1}, \ldots, v_k \in V$, $\alpha, \beta \in K$ gilt:

$$\begin{aligned} f(v_1, \ldots, v_{i-1}, \alpha v + \beta w, v_{i+1}, \ldots, v_k) &= \alpha f(v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_k) \\ &+ \beta f(v_1, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_k) \end{aligned}$$

(ii) alternierend (oder schiefsymmetrisch) falls für alle $1 \le i < j \le k$ und alle $v_1, \ldots, v_k \in V$ gilt: $v_i = v_j \Rightarrow f(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) = 0$

(5.2) Lemma. Sei V K-Vektorraum, $\dim V = n$ und $f : V^n \to K$. f erfüllt genau dann die Bedingungen (V_2) (für K statt $\mathbb{R}$) und (V_3), wenn f alternierend und n-linear ist.

Bew.: siehe R. Walter, Einf. i.d. Lin. Alg. (Vieweg-Verlag 1982) S. 146, Satz D).

(5.3) Lemma. Für jede alternierende k-Linearform $f : V^k \to K$ gilt:

(i) $\forall 1 \le i < j \le k$: $f(\ldots, \overset{i}{v}, \ldots, \overset{j}{w}, \ldots) = -f(\ldots, \overset{i}{w}, \ldots, \overset{j}{v}, \ldots)$

(ii) $v_1, \ldots, v_n$ linear abhängig $\Rightarrow f(v_1, \ldots, v_n) = 0$

(iii) $\forall 1 \le i < j \le k$: $f(\ldots, \overset{i}{v}, \ldots, w \overset{j}{+} \lambda v, \ldots) = f(\ldots, \overset{i}{v}, \ldots, \overset{j}{w}, \ldots)$

Bew.:

(i)

$$\begin{aligned} 0 &= f(\ldots, v + w, \ldots, v + w, \ldots) \\ &= f(\ldots, v, \ldots, v + w, \ldots) + f(\ldots, w, \ldots, v + w, \ldots) \\ &= \underbrace{f(\ldots, v, \ldots, v, \ldots)}_{=0} + f(\ldots, v, \ldots, w, \ldots) \\ &\quad + f(\ldots, w, \ldots, v, \ldots) + \underbrace{f(\ldots, w, \ldots, w, \ldots)}_{=0} \end{aligned}$$

(ii) Sei etwa $v_i = \sum\limits_{\substack{j=1 \\ j \ne i}}^{k} \alpha_j v_j$. Dann:

$$f(v_1, \ldots, v_n) = \sum_{\substack{j=1 \\ j \ne i}}^{n} \alpha_j \underbrace{f(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_k)}_{=0} = 0$$

(iii) klar.

Wir werden uns als nächstes damit beschäftigen, herauszufinden, wie sich der Wert einer alternierenden Form ändert, wenn man die Argumente ganz beliebig vertauscht ("permutiert"). Zunächst einiges zu diesen Permutationen:

Die Menge $S_n = \{\sigma \mid \sigma : \{1, \dots, n\} \to \{1, \dots, n\}, \sigma \text{ bijektiv}\}$ mit der Komposition $\circ$ von Abbildungen als Verknüpfung bildet eine Gruppe, die symmetrische Gruppe. Ein $\sigma \in S_n$ heißt eine Permutation von $\{1, \dots, n\}$.

Sind $a_1, \dots, a_k$ Körperelemente, so führen wir in Analogie zum Summensymbol Σ ein:

$$\prod_{i=1}^{k} a_i := a_1 \cdot a_2 \cdot \ldots \cdot a_k.$$

(5.4) Def.: Ist $\sigma \in S_n$, so heißt

$$\operatorname{sgn}(\sigma) := \prod_{1 \le i < j \le n} \frac{\sigma(j) - \sigma(i)}{j - i}$$

das Signum von σ.

Bsp.: 1) $\sigma = \mathrm{id}_{\{1,\dots,n\}} \Rightarrow \operatorname{sgn}(\sigma) = 1$.
2) Ist $\sigma \in S_2$ mit $\sigma(1) := 2$, $\sigma(2) := 1$, so $\operatorname{sgn}(\sigma) = \frac{1-2}{2-1} = -1$.
3) $\sigma \in S_n$ heißt <u>Transposition</u> $\Leftrightarrow \exists 1 \le i < j \le n$: $\sigma(i) = j$, $\sigma(j) = i$ und $\sigma(k) = k$, falls $k \in \{1, \dots, n\} \setminus \{i, j\}$.

(5.5) Fakt.

(i) $\operatorname{sgn} : S_n \to \{\pm 1\}$
(ii) $\operatorname{sgn}(\sigma) = (-1)^{\text{Fehlstandszahl von } \sigma}$, wobei die Fehlstandszahl von σ die Anzahl der Paare (i, j) ist mit $1 \le i < j \le n$ und $\sigma(i) > \sigma(j)$.
(iii) $\sigma \in S_n$ Transposition $\Rightarrow \operatorname{sgn}(\sigma) = -1$ (σ hat $2(j-i)-1$ Fehlstände!).

Bez.: Eine Permutation σ heißt gerade, falls $\operatorname{sgn}(\sigma) = 1$, sonst ungerade.

(5.6) Satz. *Für alle* $\sigma, \tau \in S_n$ *gilt* $\operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}(\sigma) \operatorname{sgn}(\tau)$, *d.h.* σ *ist Homomorphismus von* $(S_n, \circ)$ *in die Gruppe(!)* $(\{\pm 1\}, \cdot)$.

Bew.: Vorbemerkung: Wegen $\frac{\sigma(j)-\sigma(i)}{j-i} = \frac{\sigma(i)-\sigma(j)}{i-j}$ kann man in der Definition von $\operatorname{sgn}(\sigma)$ statt über alle Paare (i, j) mit $1 \le i < j \le n$ zu multiplizieren über irgendeine Menge von Paaren (i, j) multiplizieren, die die Eigenschaft hat, daß die Menge der zugehörigen Paarmengen $\{i, j\}$ die Menge aller 2-elementigen Teilmengen von $\{1, \dots, n\}$ ist.

$$\operatorname{sgn}(\sigma \circ \tau) = \prod_{1 \le i < j \le n} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{j - i} = \underbrace{\prod_{1 \le i < j \le n} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)}}_{\substack{= \operatorname{sgn}(\sigma) \\ \text{(vgl. Vorbemerkung)}}} \underbrace{\prod_{1 \le i < j \le n} \frac{\tau(j) - \tau(i)}{j - i}}_{=:\operatorname{sgn}(\tau)}$$

(5.7) Folgerung.

(i) $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)$.

(ii) Sei $\tau \in S_n$ ungerade. Dann bildet die Abbildung $\sigma \to \tau \circ \sigma$ die Menge der geraden Permutationen bijektiv auf die Menge der ungeraden Permutationen ab.

Bew.:

(i) $1 = \operatorname{sgn}(\mathrm{id}) = \operatorname{sgn}(\sigma \circ \sigma^{-1}) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\sigma^{-1}) \Rightarrow \operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$.

(ii) $\operatorname{sgn}(\sigma) = 1 \Rightarrow \operatorname{sgn}(\tau \circ \sigma) = -1$. Die Abbildung $\sigma \to \tau \circ \sigma$ ist injektiv, da aus $\tau \circ \sigma_1 = \tau \circ \sigma_2$ folgt $\tau^{-1} \circ (\tau \circ \sigma_1) = \sigma_1 = \tau^{-1} \circ (\tau \circ \sigma_2) = \sigma_2$. Ist $\tilde{\sigma} \in S_n$ ungerade, so gilt $\tau \circ (\tau^{-1} \circ \tilde{\sigma}) = \tilde{\sigma}$ und $\operatorname{sgn}(\tau^{-1} \circ \tilde{\sigma}) = \operatorname{sgn}(\tau^{-1})\operatorname{sgn}(\tilde{\sigma}) \overset{(i)}{=} \operatorname{sgn}(\tau)\operatorname{sgn}(\tilde{\sigma}) = (-1)(-1) = 1$.

(5.8) Fakt. Ist $n \geq 2$ und $\sigma \in S_n$, so existiert $k \in \{1, \ldots, n\}$ und Transpositionen $\tau_1, \ldots, \tau_k$, so daß $\sigma = \tau_1 \circ \ldots \circ \tau_k$ gilt. Es gilt dann $\operatorname{sgn}(\sigma) = (-1)^k$.

Bew.: Offensichtlich (oder per Induktion). $\operatorname{sgn}(\sigma) = (-1)^k$ folgt aus (5.6) und (5.5) (iii).

(5.9) Lemma. Sei V K-Vektorraum und $f : V^k \to K$ k-linear und alternierend. Dann gilt für alle $(v_1, \ldots, v_k) \in V^k$ und alle $\sigma \in S_k$:

$$f(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) = \operatorname{sgn}(\sigma) f(v_1, \ldots, v_k).$$

Bew.: Folgt aus (5.3) (i) und (5.8).

(5.10) Satz. *Sei V K-Vektorraum, $\dim V = n$, $(v_1, \ldots, v_n)$ Basis V und $\alpha \in K$. Dann gibt es genau eine alternierende n-Linearform $f : V^n \to K$ mit $f(v_1, \ldots, v_n) = \alpha$ und für dieses f gilt: Ist $(w_1, \ldots, w_n) \in V^n$, $w_j = \sum\limits_{i=1}^{n} a_{ij} v_i$, so*

$$(*) \qquad f(w_1, \ldots, w_n) = \alpha \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdot a_{\sigma(2)2} \cdot \ldots \cdot a_{\sigma(n)n}.$$

Speziell: Ist $f(v_1, \ldots, v_n) = 0$ für eine Basis $(v_1, \ldots, v_n)$, so ist $f \equiv 0$.

Bew.: (i) Eindeutigkeit: Sei f alternierende n-Form, $f(v_1, \ldots, v_n) = \alpha$

$$\begin{aligned}
f(w_1, \ldots, w_n) &= f\left(\sum_{i_1=1}^{n} a_{i_1 1} v_{i_1}, \sum_{i_2=1}^{n} a_{i_2 2} v_{i_2}, \ldots, \sum_{i_n=1}^{n} a_{i_n n} v_{i_n}\right) \\
&\overset{f\ n\text{-linear}}{=} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \ldots \sum_{i_n=1}^{n} a_{i_1 1} \cdot a_{i_2 2} \cdot \ldots \cdot a_{i_n n} \underbrace{f(v_{i_1}, v_{i_2}, \ldots, v_{i_n})}_{=0,\text{ falls } k \to i_k \text{ nicht injektiv}} \\
&\overset{f\ \text{alternierend}}{=} \sum_{\sigma \in S_n} a_{\sigma(1)1} \cdot a_{\sigma(2)2} \cdot \ldots \cdot a_{\sigma(n)n} f(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) \\
&\overset{(5.9)}{=} \alpha \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n}
\end{aligned}$$

Vorbem. zu (ii): $\sigma \in S_n \Rightarrow \prod\limits_{i=1}^{n} b_i = \prod\limits_{i=1}^{n} b_{\sigma(i)}$ $\quad$ ($\cdot$ kommutativ!)

(ii) Wir zeigen: $\sum\limits_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} = \sum\limits_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n}$.

Denn $\sum\limits_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} \overset{\sigma \in S_n \to \sigma^{-1} \in S_n \text{ bijektiv}}{=} \sum\limits_{\sigma \in S_n} \operatorname{sgn}(\sigma^{-1}) \underbrace{a_{1\sigma^{-1}(1)} \cdot \ldots \cdot a_{n\sigma^{-1}(n)}}_{= \prod_{i=1}^n b_i \text{ für } b_i := a_{i\sigma^{-1}(i)}}$

$$\overset{(5.7)(i)}{=} \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \underbrace{a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n}}_{= \prod_{i=1}^n b_{\sigma(i)}}$$

(iii) Existenz: Definiere f durch $(*)$. Man sieht leicht, daß f n-linear ist. Außerdem gilt wegen $v_j = \sum\limits_{i=1}^n \delta_{ij} v_i$ (mit $\delta_{ij} = 1$ für $i = j$, $\delta_{ij} = 0$ für $i \neq j$):

$$f(v_1, \ldots, v_n) = \alpha \sum \operatorname{sgn}(\sigma) \delta_{\sigma(1)1} \cdot \ldots \cdot \delta_{\sigma(n)n} = \alpha \operatorname{sgn}(\mathrm{id}) = \alpha.$$

Bleibt zu zeigen, daß $f(w_1, \ldots, w_n) = 0$ gilt, falls für ein Paar (i, j) mit $1 \leq i < j \leq n$ gilt: $w_i = w_j$. Ist $w_k = \sum\limits_{l=1}^n a_{lk} v_l$, $1 \leq k \leq n$, so folgt aus $w_i = w_j$:

$(**)$ $$a_{li} = a_{lj} \text{ für } 1 \leq l \leq n.$$

Sei τ die Transposition, die i und j vertauscht.

$$\begin{aligned} f(w_1, \ldots, w_n) &\overset{(ii)}{=} \alpha \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} \overset{(5.7)(ii)}{=} \\ &= \alpha \sum_{\sigma \in S_n \text{gerade}} a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)} - \alpha \sum_{\sigma \in S_n \text{gerade}} a_{1\tau\circ\sigma(1)} \cdot \ldots \cdot a_{n\tau\circ\sigma(n)} = 0, \end{aligned}$$

da aus $(**)$ folgt: $a_{l\sigma(l)} = a_{l\tau\circ\sigma(l)}$ für alle $l \in \{1, \ldots, n\}$ und alle $\sigma \in S_n$.

(5.11) Def.: Sei V K-Vektorraum, $\dim V = n$. Eine alternierende n-Linearform $D : V^n \to K$, $D \neq 0$, heißt <u>Determinantenform</u> <u>auf</u> V.

Erinnerung: $D \neq 0$ bedeutet: Es existiert $(v_1, \ldots, v_n) \in V^n$: $D(v_1, \ldots, v_n) \neq 0$.

Bsp.: $D_0 : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, $D_0((x_1, x_2), (y_1, y_2)) = x_1 y_2 - x_2 y_1$.

$$D_0(e_1, e_2) = 1$$

(5.12) Folgerung. Sei D Determinantenform auf dem n-dimensionalen K-Vektorraum V und $(v_1, \ldots, v_n) \in V^n$. Dann gilt

(i) $(v_1, \ldots, v_n)$ Basis von $V \Leftrightarrow D(v_1, \ldots, v_n) \neq 0$.
(ii) Ist $f : V^n \to K$ alternierende n-Linearform, so existiert $\alpha \in K : f = \alpha D$.

Bem.: 1) (ii) $\Leftrightarrow \{f \mid f \text{ alternierende } n\text{-Linearform}\}$ ist 1-dimensionaler K-Vektorraum. Die Determinantenformen sind die Elemente $\neq 0$ dieses Vektorraumes.

2) $f = \alpha D$, $(v_1, \ldots, v_n)$ Basis von $V \Rightarrow \alpha = \frac{f(v_1, \ldots, v_n)}{D(v_1, \ldots, v_n)}$

Bew.: (i) Sei $(v_1, \ldots, v_n)$ Basis von V. Dann folgt aus (5.10) ("Speziell:") und $D \neq 0$, daß $D(v_1, \ldots, v_n) \neq 0$ gilt. Ist $D(v_1, \ldots, v_n) \neq 0$, so folgt aus (5.3)(ii), daß die $v_1, \ldots, v_n$ linear unabhängig sind, also - wegen $\dim V = n$ - eine Basis von V.

(ii) Sei $(v_1, \ldots, v_n)$ eine Basis von V und $\alpha := \frac{f(v_1,\ldots,v_n)}{D(v_1,\ldots,v_n)}$. Dann sind f und αD alternierende n-Linearformen, die auf der Basis $(v_1, \ldots, v_n)$ den gleichen Wert annehmen. Also folgt aus (5.10) (Eindeutigkeit): $f = \alpha D$.

Einschub: Die Determinante quadratischer Matrizen.

In K^n haben wir die Standardbasis $(e_1, \ldots, e_n)$, die - nach (5.10) - eine "Standarddeterminantenform" $D_0 : K^n \times \ldots \times K^n \to K$ durch

$$D_0(e_1, \ldots, e_n) = 1$$

eindeutig bestimmt. Ist $A = (a_{ij})_{1 \leq i,j \leq n} \in K^n$, so seien $A_j := \sum\limits_{i=1}^{n} a_{ij} e_i$ die "Spaltenvektoren" von A.

(5.13) Def. Die Determinante $\det : K^{n \times n} \to K$ ist definiert durch

$$\det A := D_0(A_1, \ldots, A_n)$$

Bezeichnungsweise: $\det A = \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix}$

(5.14) Folgerung (Gottfried Wilhelm Leibniz 1646-1716). Ist $A = (a_{ij})_{1 \leq i,j \leq n} \in K^{n \times n}$, so

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

Bem.: Beweis (5.10)(ii) $\Rightarrow$ $\det A = \sum\limits_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}$.

Bew.: $\det(A) = D_0(A_1, \ldots, A_n)$, $A_j = \sum\limits_{i=1}^{n} a_{ij} e_i$, $D_0(e_1, \ldots, e_n) = 1 \overset{(5.10)}{\Rightarrow}$ Beh.

Speziell: $n = 1: \quad A = (a_{11}) \to \quad \det A = a_{11}$
$n = 2: \quad A = \binom{a_{11} a_{12}}{a_{21} a_{22}} \to \quad \det A = a_{11}a_{22} - a_{12}a_{21}$

$n = 3$: Regel von Sarrus (Rechenverfahren)

$$\begin{array}{ccccccccc}
a_{11} & & a_{12} & & a_{13} & & a_{11} & & a_{12} \\
 & \diagdown & & \times & & \times & & \diagup & \\
a_{21} & & a_{22} & & a_{23} & & a_{21} & & a_{22} \\
 & \diagup & & \times & & \times & & \diagdown & \\
a_{31} & & a_{32} & & a_{33} & & a_{31} & & a_{32}
\end{array}$$

$$\begin{aligned} \det A &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &- a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12} \end{aligned}$$

(5.15) Satz (einige Eigenschaften von $\det A$)

(i) $\det A$ *ist n-linear und alternierend in den Spaltenvektoren von A.*
(ii) $\det E_n = 1$.
(iii) $\det A \neq 0 \Leftrightarrow A \in \mathrm{GL}_n(K)$.

Bew.:

(i) folgt direkt aus der Definition.
(ii) $\det E_n = D_0(e_1, \ldots, e_n) = 1$.
(iii) $\det A \neq 0 \overset{(5.12)(i)}{\Leftrightarrow}$ Spaltenvektoren linear unabhängig $\Leftrightarrow A \in \mathrm{GL}_n(K)$.

Folgerungen aus (5.15)(i):

$$\begin{vmatrix} a_{11} & \ldots & \alpha a_{1j} & \ldots & a_{n1} \\ \vdots & & & & \vdots \\ a_{m1} & \ldots & \alpha a_{nj} & \ldots & a_{nn} \end{vmatrix} = \alpha \begin{vmatrix} a_{11} & \ldots & a_{n1} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix}$$

und: $\det(\alpha A) = \alpha^n \det A$.

Vertauscht man 2 Spalten von A, so ändert sich das Vorzeichen der Determinante. Addiert man zu einer Spalte eine Linearkombination der anderen Spalten, so ändert sich $\det A$ nicht, z.B.

$$\begin{vmatrix} a_{11} + \alpha a_{12} & a_{12} & \ldots & a_{1n} \\ a_{21} + \alpha a_{22} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} + \alpha a_{n2} & a_{n2} & \ldots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix}$$

(5.15)(iii): $\det A = 0 \Leftrightarrow$ Spaltenvektoren sind linear abhängig.

(5.16) Def. Sei $A \in K^{m\times n}$. Die transponierte Matrix $A^T \in K^{n\times m}$ zu A hat als Zeilen gerade die Spalten von A (in der gleichen Reihenfolge), d.h.

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} a_{11} & \ldots & a_{m1} \\ a_{12} & \ldots & a_{m2} \\ \vdots & & \vdots \\ a_{1n} & \ldots & a_{mn} \end{pmatrix}$$

oder: $A = (a_{ij})_{\substack{1\leq i\leq m \\ 1\leq j\leq n}} \in K^{m\times n}$, so $A^T = (\tilde{a}_{ji})_{\substack{1\leq j\leq n \\ 1\leq i\leq m}}$ mit $\tilde{a}_{ji} = a_{ij}$.

Bsp.: $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in \mathbb{Q}^{2\times 3} \Rightarrow A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \in \mathbb{Q}^{3\times 2}$.

(5.17) Folgerung (aus Bew. (5.10)(ii)).

(i) $\forall A \in K^{n\times n}: \det A = \det(A^T)$.

(ii) $\det A$ ist auch in den Zeilenvektoren von A n-linear und alternierend. (D.h. (5.15)(i) + Folgerungen gelten auch für Zeilenvektoren).

Bew.:

(i) $(A^T = (\tilde{a}_{ji})_{\substack{1\le j\le n\\ 1\le i\le n}}$ mit $\tilde{a}_{ji} = a_{ij}$.

$$\det(A^T) = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\tilde{a}_{\sigma(1)1}\cdot\ldots\cdot\tilde{a}_{\sigma(n)n} \sum_{\sigma\in S_n}\operatorname{sgn}(\sigma)a_{1\sigma(1)}\cdot\ldots\cdot a_{n\sigma(n)} \overset{\text{Bew.(5.10)(ii)}}{=} \det A.$$

(ii) folgt aus (i).

Konkretes Verfahren zur Berechnung von $\det A$ für größere n: Durch Addition von Vielfachen von Zeilen zu anderen Zeilen und - falls nötig - durch Vertauschung von Zeilen verwandle A in "obere Dreiecksmatrix" $\overline{A} = (\overline{a}_{ij})$ (d.h. $\overline{a}_{ij} = 0$ für $i > j$), analog zu (4.25). Dann gilt, wenn k die Anzahl der Zeilenvertauschungen bezeichnet:

$$\det A = (-1)^k \det\overline{A} = (-1)^k\overline{a}_{11}\cdot\ldots\cdot\overline{a}_{nn}. \text{ Denn:}$$

(5.18) Lemma. Ist $A = (a_{ij}) \in K^{n\times n}$ obere (oder untere) Dreiecksmatrix, so gilt $\det A = a_{11}\cdot\ldots\cdot a_{nn}$.

Bew.: $a_{\sigma(1)1}\cdot\ldots\cdot a_{\sigma(n)n} \neq 0 \Rightarrow \forall i\in\{1,\ldots,n\}: \sigma(i)\le i \Rightarrow \sigma = \mathrm{id} \Rightarrow$ Beh.

Bsp.:

$$\begin{vmatrix} 1 & 2 & 3\\ 4 & 8 & 6\\ 7 & 8 & 10\end{vmatrix}\begin{array}{cc} -4 \,\rfloor & -7\,\rfloor \\ \hookleftarrow & \\ & \hookleftarrow \end{array} = \begin{vmatrix} 1 & 2 & 3\\ 0 & 0 & -6\\ 0 & -6 & -11\end{vmatrix} = -\begin{vmatrix} 1 & 2 & 3\\ 0 & -6 & -11\\ 0 & 0 & -6\end{vmatrix} = -36$$

Zurück zur Theorie:

Bem.: Ist $K = \mathbb{R}$, $\dim V = n$ und $D: V^n \to \mathbb{R}$ Determinantenform, so definiert man das **Volumen** (bzgl. D) $\mathrm{vol}_D(P)$ eines Parallelotops $P = P(v_1,\ldots,v_n) \subseteq V$ durch $\mathrm{vol}_D(P) = |D(v_1,\ldots,v_n)|$. Wegen (5.12)(ii) gilt:

Sind D, $\tilde{D}$ Determinantenformen auf V, so existiert $\alpha\in\mathbb{R}_{>0}$, so daß für alle Parallelotope $P = P(v_1,\ldots,v_n)\subseteq V$ gilt:

$$\mathrm{vol}_{\tilde{D}}(P) = \alpha\,\mathrm{vol}_D(P).$$

(5.19) Satz und Def. *Sei V n-dimensionaler K-Vektorraum und $L\in\mathrm{End}(V)$. Dann ist für jede Basis $(v_1,\ldots,v_n)$ von V und jede Determinantenform D auf V der Quotient $\frac{D(L(v_1),\ldots,L(v_n))}{D(v_1,\ldots,v_n)}$ das gleiche Element von K und wir definieren*

$$\det L := \frac{D(L(v_1),\ldots,L(v_n))}{D(v_1,\ldots,v_n)}.$$

Bem.: Sei $K = \mathbb{R}$. Dann gilt für jedes Parallelotop $P \subseteq V^n$ und jede Determinantenform D: $\mathrm{vol}_D(L(P)) = |\det L| \, \mathrm{vol}_D(P)$.
(L linear $\Rightarrow L(P(v_1, \ldots, v_n)) = P(L(v_1), \ldots, L(v_n))$!)

Bew. von (5.19): Wegen (5.12)(i) gilt $D(v_1, \ldots, v_n) \neq 0$. Aus (5.12)(ii) folgt, daß $\det L$ unabhängig von der Wahl von D ist. Um die Unabhängigkeit von der Wahl der Basis zu beweisen, bemerken wir, daß

$$f : (w_1, \ldots, w_n) \in V^n \to D(L(w_1), \ldots, L(w_n))$$

eine alternierende n-Linearform ist, so daß aus (5.12)(ii) folgt: Es existiert $\alpha \in K$:

$$f = \alpha D.$$

Für jede Basis $v_1, \ldots, v_n)$ von V gilt also:

$$\alpha = \frac{f(v_1, \ldots, v_n)}{D(v_1, \ldots, v_n)} = \frac{D(L(v_1), \ldots, L(v_n))}{D(v_1, \ldots, v_n)} (=: \det L).$$

(5.20) Folgerung. Seien $L, L_1, L_2 \in \mathrm{End}(V), \alpha \in K$. Dann gilt:

(i) $\det(\mathrm{id}_V) = 1$
(ii) $L \in \mathrm{Aut}(V) \Leftrightarrow \det L \neq 0$
(iii) $\det(L_2 \circ L_1) = \det L_2 \det L_1$
(iv) $L \in \mathrm{Aut}(V) \Rightarrow \det(L^{-1}) = (\det L)^{-1}$
(v) $\det(\alpha L) = \alpha^n \det(L)$

Bew.:

(ii) $L \in \mathrm{Aut}(V) \overset{(4.5)}{\Leftrightarrow}$ Ist $(v_1, \ldots, v_n)$ Basis, so auch $(L(v_1), \ldots, L(v_n))$
$\overset{(5.12)(i),\,(5.19)}{\Leftrightarrow} \det L \neq 0.$

(iii) 1. Fall: $L_1 \in \mathrm{Aut}(V)$. Sei $(v_1, \ldots, v_n)$ Basis von V. Dann ist $(L(v_1), \ldots, L(v_n))$ Basis von V, und es gilt:

$$\begin{aligned} \det(L_2 \circ L_1) &= \tfrac{D(L_2(L_1(v_1)), \ldots, L_2(L_1(v_n)))}{D(v_1, \ldots, v_n)} \\ &= \tfrac{D(L_2(L_1 v_1), \ldots, L_2(L_1(v_n)))}{D(L_1(v_1), \ldots, L_1(v_n))} \cdot \det L_1 \\ &= \det L_2 \cdot \det L_1 \end{aligned}$$

2. Fall: $L_1 \notin \mathrm{Aut}(V) (\overset{(ii)}{\Leftrightarrow} \det L_1 = 0)$. Dann ist L_1 nicht injektiv, vgl. (4.8). Also ist auch $L_2 \circ L_1$ nicht injektiv, also $\det(L_2 \circ L_1) = 0$. Daraus folgt:

$$0 = \det L_2 \underbrace{\det L_1}_{=0} = \det(L_2 \circ L_1)$$

(iv) $L \in \mathrm{Aut}(V) \Rightarrow 1 \overset{(i)}{=} \det(\mathrm{id}_V) = \det(L^{-1} \circ L) \overset{(iii)}{=} \det(L^{-1}) \det L \Rightarrow$ Beh.
(v) folgt direkt auf Def. (5.19).

Zusammenhang mit der Determinante von Matrizen:

(5.21) Fakt.

(i) Sei $L \in \operatorname{End}(K^n)$ und $\operatorname{Mat}(L) = A$, d.h. "$L(x) = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$".

$$\text{Dann gilt: } \det L = \det A.$$

(ii) Sei V K-Vektorraum $\mathcal{G} = (v_1, \ldots, v_n) \in V^n$ Basis von V und $L \in \operatorname{End}(V)$.

$$\text{Dann gilt: } \det L = \det(\operatorname{Mat}_{\mathcal{G}}^{\mathcal{G}}(L)).$$

Bew.:

(i) $\operatorname{Mat}(L) = A \Leftrightarrow L(e_j) = A_j \Rightarrow$

$$\det L = \frac{D_0(L(e_1), \ldots, L(e_n))}{D_0(e_1, \ldots, e_n)} = D_0(A_1, \ldots, A_n) = \det A.$$

(ii) Sei $J \in \operatorname{Hom}(K^n, V)$ der Isomorphismus mit $J(e_j) = v_j$ für $j \in \{1, \ldots, n\}$, vgl. (4.5). Dann gilt $\operatorname{Mat}(J^{-1} \circ L \circ J) = \operatorname{Mat}_{\mathcal{G}}^{\mathcal{G}}(L)$, da $J^{-1} \circ L \circ J(e_j) = \sum\limits_{i=1}^{n} a_{ij} \underbrace{e_i}_{=J^{-1}(v_i)} \Leftrightarrow$

$L(v_j) = \sum\limits_{i=1}^{n} a_{ij} v_i$. Wir verwenden auf V die Determinantenform (!) D, die durch

$$D(w_1, \ldots, w_n) = D_0(J^{-1}(w_1), \ldots, J^{-1}(w_n))$$

definiert ist. Dann gilt $D(v_1, \ldots, v_n) = D_0(e_1, \ldots, e_n) = 1$ und damit

$$\begin{aligned} \det L &= D(L(v_1), \ldots, L(v_n)) = D_0(J^{-1} \circ L(v_1), \ldots, J^{-1} \circ L(v_1)) \\ &= D_0(J^{-1} \circ L \circ J(e_1), \ldots, J^{-1} \circ L \circ J(e_n)) \\ &= \det(J^{-1} \circ L \circ J) \overset{\text{(i)}}{=} \det(\operatorname{Mat}(J^{-1} \circ L \circ J)) = \det(\operatorname{Mat}_{\mathcal{G}}^{\mathcal{G}}(L)). \end{aligned}$$

Bem.: Es gilt i.a. nicht $\det L = \det(\operatorname{Mat}_{\mathcal{G}}^{\overline{\mathcal{G}}}(L))$, wenn $\mathcal{G}$ und $\overline{\mathcal{G}}$ verschiedene Basen von V sind.

(5.22) Folgerung (aus (5.20)(ii), (iii), (5.21) und (4.15)).

(i) $A, B \in K^{n \times n} \Rightarrow \det(AB) = \det A \cdot \det B$.
(ii) $A \in \operatorname{GL}_n(K) \Rightarrow \det(A^{-1}) = (\det A)^{-1}$

Speziell besagt (5.22): Die Abbildung $A \to \det A$ ist ein Gruppenhomomorphismus von $(\operatorname{GL}_n(K), \cdot)$ auf $(K \setminus \{0\}, \cdot)$. $\operatorname{SL}_n(K) = \{A \in \operatorname{GL}_n(K) | \det A = 1\}$ ist Untergruppe von $\operatorname{GL}_n(K)$, die spezielle lineare Gruppe.

Eigenwerte und Eigenvektoren (eine Anwendung der Determinante)

(5.23) Def.: Sei V K-Vektorraum, $L \in \mathrm{End}(V)$. $\lambda \in K$ heißt Eigenwert (EW) von L, falls ein $v \in V \setminus \{0\}$ existiert mit $L(v) = \lambda v$. $v \in V$ heißt Eigenvektor (EV) von L, falls $v \neq 0$ gilt und ein $\lambda \in K$ existiert mit $L(v) = \lambda v$. In diesem Fall heißt v ein Eigenvektor zum Eigenwert λ.

Bem.: v EV zum EW λ, $\alpha \in K \setminus \{0\} \Rightarrow \alpha v$ EV zum EW λ.

Bsp.:

1) Ist $L = \lambda\,\mathrm{id}_V$, so ist λ der einzige EW von L und jedes $v \in V \setminus \{0\}$ ist EV zum EW λ.
2) Ist $L : K^n \to K^n$ definiert durch $L(e_i) = \lambda_i e_i$ für alle $i \in \{1, \ldots, n\}$, d.h.

$$\mathrm{Mat}(L) = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix},$$

so sind $\lambda_1, \ldots, \lambda_n$ Eigenwerte von L und $e_1, \ldots, e_n$ Eigenvektoren von L.
3) $L : \mathbb{R}^2 \to \mathbb{R}^2$, $L(x_1, x_2) = (x_1 + x_2, x_2)$ "Scherung", mit

$$\mathrm{Mat}(L) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Dann ist $\lambda = 1$ der einzige EW von L und $x \in \mathbb{R}^2$ ist genau dann EV von L, wenn $x = (x_1, 0)$ für ein $x_1 \neq 0$ gilt.
4) $L : \mathbb{R}^2 \to \mathbb{R}^2$, $L(x_1, x_2) = (ax_1 - bx_2, bx_1 + ax_2)$ mit

$$\mathrm{Mat}(L) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Es folgt aus Aufgabe 1b), Blatt 11, daß L für $b \neq 0$ keinen (reellen) EW hat. Geometrisch ist L eine "Drehstreckung". Identifiziert man $\mathbb{R}^2$ mit $\mathbb{C}$ durch $(x_1, x_2) \leftrightarrow x_1 + ix_2$, vgl. Kap. 2, so kann man L schreiben als: $L(x_1 + ix_2) = (a + ib)(x_1 + ix_2)$. Die Abbildung, die $z \in \mathbb{C}$ die Zahl $(a + ib)z \in \mathbb{C}$ zuordnet, ist also eine Drehstreckung.
5) "Gekoppelte Schwingungen": Gegeben $A \in \mathrm{End}(\mathbb{R}^n)$, gesucht Lösungen $x : \mathbb{R} \to \mathbb{R}^n$ der "Differentialgleichung":

$$(*) \qquad x''(t) = A(x(t)) \text{ für alle } t \in \mathbb{R}.$$

Ist $v \in \mathbb{R}^n$ EV von A zum EW $\lambda \in \mathbb{R}$ und ist $f : \mathbb{R} \to \mathbb{R}$ Lösung von $f''(t)$-$\lambda f(t)$ $= 0$ (diese sind explizit bekannt!), so ist $x(t) := f(t)v$ Lösung von $(*)$:

$$x''(t) = f''(t)v = \lambda f(t)v = f(t)A(v) = A(f(t)v) = A(x(t)).$$

(5.24) Satz. *Sei* $1 \leq \dim V < \infty$, $L \in \mathrm{End}(V)$. *Dann gilt:*

$$\lambda \text{ EW } von\ L \Leftrightarrow \det(L - \lambda\,\mathrm{id}_V) = 0.$$

Bew.: $\det(L-\lambda\,\mathrm{id}_V)=0 \overset{(5.20)(\mathrm{ii})}{\Leftrightarrow} (L-\lambda\,\mathrm{id}_V)\notin \mathrm{Aut}(V) \overset{(4.8)}{\Leftrightarrow}$
$\mathrm{kern}(L-\lambda\,\mathrm{id}_V)\neq\{0\} \Leftrightarrow \exists v\in V\setminus\{0\}:(L-\lambda\,\mathrm{id}_V)(v)=0$
$\Leftrightarrow \exists v\in V\setminus\{0\}: L(v)=\lambda v.$

Bem.: Ist $\mathrm{Mat}^{\mathcal{G}}_{\mathcal{G}}(L)=(a_{ij})\in K^{n\times n}$, so gilt $\mathrm{Mat}^{\mathcal{G}}_{\mathcal{G}}(L-\lambda\,\mathrm{id}_V)=(a_{ij}-\lambda\delta_{ij})_{1\le i,j\le n}$. Also nach (5.21):

$$\begin{aligned}\det(L-\lambda\,\mathrm{id}_V) &= \sum_{\sigma\in S_n}\mathrm{sgn}(\sigma)(a_{1\sigma(1)}-\lambda\delta_{1\sigma(1)})\cdot\ldots\cdot(a_{n\sigma(n)}-\lambda\delta_{n\sigma(n)})\\ &= (-1)^n\lambda^n+\left(\sum_{i=1}^n a_{ii}\right)(-1)^{n-1}\lambda^{n-1}+?\lambda^{n-2}+\ldots+?\lambda+\det L,\end{aligned}$$

mit von den a_{ij} abhängenden Koeffizienten ?, die uns im Moment nicht weiter interessieren. Einen Ausdruck dieser Art nennt man “ein Polynom in (der Variablen) λ”.

Bez.: $P_L(\lambda):=\det(L-\lambda\,\mathrm{id}_V)$ heißt das <u>charakteristische Polynom von</u> L.

(5.24) besagt: λ EW von $L\Leftrightarrow P_L(\lambda)=0 \overset{\text{def.}}{\Leftrightarrow} \lambda$ Nullstelle von P_L. Ob Polynome (vom Grad ≥ 1) <u>stets</u> Nullstellen haben, hängt vom Körper K ab!

$K=\mathbb{Q},\mathbb{R}$ $\rightarrow$ nicht immer, z.B. λ^2+1
$K=\mathbb{C}$ $\rightarrow$ stets (“Fundamentalsatz der Algebra”)

Verfahren zur Bestimmung von EW’en und EV’en eines $L\in\mathrm{End}(V)$:

1) Berechne P_L $(=\det(L-\lambda\,\mathrm{id}_V))$.
2) Bestimme Nullstellen von P_L (= EW’e von L). Das ist oft nur approximativ möglich.
3) λ Nullstelle von P_L $\Rightarrow$ Das lineare Gleichungssystem $L(v)=\lambda v$ besitzt Lösung $v\neq 0$. Bestimme alle Lösungen $\neq 0$(= EVen zum EW λ).

Es ist leicht zu zeigen: Besitzt P_L $n(=\dim V)$ verschiedene Nullstellen, so ist L diagonalisierbar ($\Leftrightarrow$ es existiert Basis $\mathcal{G}$ von V: $\mathrm{Mat}^{\mathcal{G}}_{\mathcal{G}}(L)=\begin{pmatrix}\lambda_1 & & & \\ & \ddots & 0 & \\ & 0 & \ddots & \\ & & & \lambda_n\end{pmatrix}$, nämlich: $\mathcal{G}=(v_1,\ldots,v_n)$, v_i EV von L zum EW λ_i).

Weitere Anwendungen der Determinante:

(5.25) Def.: Eine (geordnete) Basis $(v_1,\ldots,v_n)$ des $\mathbb{R}^n$ heißt positiv orientiert (oder Rechtssystem), falls $D_0(v_1,\ldots,v_n)>0$ (sonst negativ orientiert oder Linkssystem).

Bsp.: $(e_1,\ldots,e_n)$ ist positiv orientiert, $(-e_1,e_2,\ldots,e_n)$ und $(e_2,e_1,e_3,\ldots,e_n)$ sind negativ orientiert.

Bem.: Ist $L \in \operatorname{Aut}(\mathbb{R}^n)$, $\det L > 0$, so gilt:

$$(v_1, \ldots, v_n) \text{ positiv orientierte Basis} \Leftrightarrow (L(v_1), \ldots, L(v_n)) \text{ positiv orientierte Basis.}$$

(5.26) Cramersche Regel (Gabriel Cramer 1704-1752). Sei

$$\text{I} \qquad \begin{array}{cccc} a_{11}x_1 & + \ldots + & a_{1n}x_n = & b_1 \\ \vdots & & & \vdots \\ a_{n1}x_1 & + \ldots + & a_{nn}x_n = & b_1 \end{array} \qquad \text{oder kurz:} \quad A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = b$$

Gleichungssystem mit n Gleichungen für n Unbekannte. Dann gilt: Ist $\det A \neq 0$, so ist

$$x_k = \frac{1}{\det A} \begin{vmatrix} a_{11} & \ldots & a_{1,k-1} & b_1 & a_{1,k+1} & \ldots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \ldots & a_{n,k-1} & b_n & a_{n,k+1} & \ldots & a_{nn} \end{vmatrix} \qquad (\text{ für } k = 1, \ldots, n)$$

die (einzige) Lösung von I.

Bew.: $\det A \neq 0 \overset{(5.20)(\text{ii})}{\Leftrightarrow} x \in \mathbb{R}^n \to Ax \in \mathbb{R}^n$ ist bijektiv. Also existiert genau eine Lösung von I, d.h. genau ein $x \in \mathbb{R}^n$: $Ax = b$. Sind $A_1, \ldots, A_n$ die Spaltenvektoren von A, so bedeutet $Ax = b$ gerade

$$x_1 A_1 + x_2 A_2 + \ldots + x_n A_n = b \quad (\text{vgl. I!})$$

Also gilt:

$$\begin{aligned} D_0(A_1, \ldots, A_{k-1}, b, A_{k+1}, \ldots, A_n) &= D_0(A_1, \ldots, A_{k-1}, \sum_{i=1}^n x_i A_i, A_{k+1}, \ldots, A_n) \\ &\overset{D_0\ n\text{-linear}}{=} \sum_{i=1}^n x_i D_0(A_1, \ldots, A_{k-1}, A_i, A_{k+1}, \ldots, A_n) \\ &\overset{D_0 \text{ alternierend}}{=} x_k \det A. \end{aligned}$$

Eine weitere Erkenntnis, die uns die Determinante schenkt:

Nach (5.15)(iii) wissen wir, daß eine <u>reelle</u> quadratische Matrix $A \in K^{n \times n}$ genau dann in $\mathrm{GL}(n, \mathbb{R})$ liegt, wenn $\det A \neq 0$ gilt. Daraus folgt mit etwas Analysis: Für die "meisten" Matrizen ("für eine offene, dichte Menge" bzw. "für alle bis auf eine Menge vom Maß 0") $A \in \mathbb{R}^{n \times n}$ gilt $A \in \mathrm{GL}_n(R)$. Speziell: "Typischerweise" ist ein System von n linearen Gleichungen für n (reelle) Unbekannte eindeutig lösbar. Oder geometrisch: "Typischerweise" schneiden sich (z.B.) zwei 3-dim. affine Unterräume des $\mathbb{R}^6$ in genau einem Punkt des $\mathbb{R}^6$ (2mal 3 lineare Gleichungen für 6 Unbekannte).

Zum Abschluß des Kapitels noch eine (i.a. nicht besonders praktische) Möglichkeit zum Berechnen von Determinanten - der Laplacesche Entwicklungssatz (Pierre Simon Laplace 1749-1827)).

(5.27) Lemma. Ist $\tilde{A} \in K^{(n-1)\times(n-1)}$ und $A := \begin{pmatrix} 1 & 0\dots 0 \\ 0 & \\ \vdots & \tilde{A} \\ 0 & \end{pmatrix} \in K^{n\times n}$, so gilt

$$\det A = \det \tilde{A}.$$

Bew.: Als Funktion der Spaltenvektoren von $\tilde{A}$ ist $\det A$ $(n-1)$-linear, alternierend, und für $\tilde{A} = E_{n-1}$ gilt $A = E_n$ und $\det E_n = 1$. Mit der Eindeutigkeitsaussage von (5.10) impliziert das: $\det \tilde{A} = \det A$.

Bez.: Zu $A \in K^{n\times n}$ und $1 \leq i, j \leq n$ bezeichne $A_{ij} \in K^{(n-1)\times(n-1)}$ die Matrix, die aus A durch <u>Streichen</u> der i'ten Zeile und der j'ten Spalte entsteht.

Bsp.:

$$A : \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 8 & 7 & 6 & 5 \\ 4 & 3 & 2 & 1 \end{pmatrix} \Rightarrow A_{23} = \begin{pmatrix} 1 & 2 & 4 \\ 8 & 7 & 5 \\ 4 & 3 & 1 \end{pmatrix}.$$

(5.28) Entwicklungssatz. *Für alle* $A \in K^{n\times n}$, $j \in \{1, \dots, n\}$ *gilt:*

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij} \quad \textit{(Entwicklung nach der j'ten Spalte).}$$

Für alle $i \in \{1, \dots, n\}$ *gilt:*

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij} \quad \textit{(Entwicklung nach der i'ten Spalte).}$$

Bsp.:

$$\begin{vmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 0 & 2 \\ 3 & 1 & 1 & 1 \\ 1 & 0 & 0 & 2 \end{vmatrix} \overset{\substack{\text{Entwicklung nach} \\ \text{der 3. Spalte}}}{=} (-1)^{1+3} \begin{vmatrix} 0 & 1 & 2 \\ 3 & 1 & 1 \\ 1 & 0 & 2 \end{vmatrix} + (-1)^{3+3} \begin{vmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 2 \end{vmatrix} = -2$$

$$\overset{\substack{\text{oder Entwicklung nach} \\ \text{der 4. Zeile}}}{=} (-1)^{1+3} \begin{vmatrix} 2 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 1 & 1 \end{vmatrix} + (-1)^{4+4} \cdot 2 \cdot \begin{vmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 3 & 1 & 1 \end{vmatrix} = -2.$$

Bew.: $A_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}, \ldots, A_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix}$ seien die Spaltenvektoren von A.

$$\begin{aligned} \det A &= D_0(A_1, \ldots, A_n) = D_0(A_1, \ldots, A_j = \sum a_{ij} e_i, \ldots, A_n) \\ &= \sum_{i=1}^{n} a_{ij} D_0(A_1, \ldots, A_{j-1}, e_i, A_{j+1} \ldots, A_n) \end{aligned}$$

Es bleibt zu zeigen: $D_0(A_1, \ldots, A_{j-1}, e_i, A_{j+1} \ldots, A_n) = (-1)^{i+j} \det A_{ij}$

$D_0(A_1, \ldots, A_{j-1}, e_i, A_{j+1} \ldots, A_n) = (-1)^{j-1} D_0(e_i, A_1, \ldots, A_{j-1}, A_{j+1}, \ldots, A_n)$

$$= \quad (-1)^{j-1} \left| \begin{array}{ccc|cc} 0 & a_{11} & \ldots & \ldots & a_{1n} \\ \vdots & & & & \vdots \\ 1 & a_{i1} & \ldots & \ldots & a_{in} \\ \vdots & & & & \vdots \\ 0 & a_{n1} & \ldots & \ldots & a_{nn} \end{array} \right| = (-1)^{i+j-2} \left| \begin{array}{ccc|cc} 1 & a_{i1} & \ldots & \ldots & a_{in} \\ 0 & a_{11} & \ldots & \ldots & a_{1n} \\ \vdots & \vdots & & & \vdots \\ 0 & a_{i-1,1} & \ldots & \ldots & a_{i-1,n} \\ 0 & a_{i+1,1} & \ldots & \ldots & a_{i+1,n} \\ \vdots & \vdots & & & \vdots \\ 0 & a_{n1} & \ldots & \ldots & a_{nn} \end{array} \right|$$

$$= \quad (-1)^{i+j} \left| \begin{array}{ccc|cc} 1 & 0 & \ldots & \ldots & 0 \\ 0 & a_{11} & \ldots & \ldots & a_{1n} \\ \vdots & \vdots & & & \vdots \\ 0 & a_{i-1,1} & \ldots & \ldots & a_{i-1,n} \\ 0 & a_{i+1,1} & \ldots & \ldots & a_{i+1,n} \\ \vdots & \vdots & & & \vdots \\ 0 & a_{n1} & \ldots & \ldots & a_{nn} \end{array} \right|$$

$$\overset{(5.27)}{=} \quad (-1)^{i+j} \quad \det A_{ij}$$

(hierbei soll der senkrechte Strich in den Determinanten andeuten, daß die j'te Spalte gestrichen ist!)

6. Euklidische Vektorräume

Im folgenden sei V ein Vektorraum über dem Körper $\mathbb{R}$.

(6.1) Def.: Eine Bilinearform $b : V \times V \to \mathbb{R}$ heißt

(i) symmetrisch $\Leftrightarrow \forall\, v, w \in V$: $b(v, w) = b(w, v)$
(ii) positiv definit $\Leftrightarrow \forall\, v \in V \setminus \{0\}$: $b(v, v) > 0$
(iii) Skalarprodukt $\Leftrightarrow$ b ist symmetrisch und positiv definit.

Bez.: Für ein Skalarprodukt b schreibt man oft: $\langle v, w \rangle := b(v, w)$. Ein reeller Vektorraum V mit einem Skalarprodukt $\langle , \rangle$ heißt ein euklidischer Vektorraum.

Bsp.: 1) $V = \mathbb{R}^n$, $\quad b(x,y) := \langle x, y\rangle := \sum\limits_{i=1}^{n} x_i y_i$

$$b(x,x) = \sum_{i=1}^{n} x_i^2$$

2) $$V = C^0([0,2\pi],\mathbb{R}) = \{f \mid f : [0,2\pi] \to \mathbb{R} \text{ stetig}\}$$

$$b(f_1,f_2) := \langle f_1, f_2\rangle_{L^2} = \int_0^{2\pi} f_1(t) f_2(t)\,dt$$

(6.2) Def.: Sei $V, \langle,\rangle$ euklidischer Vektorraum. Dann heißt $\| v \| := \sqrt{\langle v, v\rangle} \geq 0$ die Norm (oder Länge) von $v \in V$ (bzgl. $\langle,\rangle$) und $\| v - w \|$ der Abstand zwischen v und w.

(6.3) Satz. *Sei $V, \langle,\rangle$ euklidischer Vektorraum. Dann gilt für alle $v, w \in V, \lambda \in \mathbb{R}$:*

(i) $\| v \| \geq 0$ *und:* $\| v \| = 0 \Leftrightarrow v = 0$
(ii) $\| \lambda v \| = |\lambda| \, \| v \|$
(iii) $|\langle v, w\rangle| \leq \| v \| \| w \|$ *(Cauchy-Schwarzsche Ungleichung) mit "=" $\Leftrightarrow$ v und w sind linear abhängig.*
(iv) $\| v + w \| \leq \| v \| + \| w \|$ *(Dreiecksungleichung)*

Bew.:

(i) ist klar
(ii) $\| \lambda v \| = \sqrt{\langle \lambda v, \lambda v\rangle} = \sqrt{\lambda^2 \langle v, v\rangle} = |\lambda| \, \| v \|$
(iii) Es genügt, den Fall $w \neq 0$ zu betrachten. Dann gilt für alle $t \in \mathbb{R}$:

$$\begin{aligned} 0 \leq \langle v + tw, v + tw\rangle &= \| v \|^2 + 2t\langle v, w\rangle + t^2 \| w \|^2 \\ &= \left(t \| w \| + \tfrac{\langle v,w\rangle}{\|w\|}\right)^2 + \| v \|^2 - \tfrac{\langle v,w\rangle^2}{\|w\|^2} \end{aligned}$$

Als Funktion von t nimmt $\langle v + tw, v + tw\rangle$ für $t_0 = -\frac{\langle v,w\rangle}{\|w\|^2}$ sein (nichtnegatives!) Minimum an. Daraus folgt (iii). Gleichheit in (iii) tritt genau dann ein, wenn dieses Minimum gleich Null ist, d.h. falls $\langle v + t_0 w, v + t_0 w\rangle = 0$. Daraus folgt, daß $v + t_0 w = 0$ gilt, d.h. v und w sind linear abhängig. Umgekehrt gilt für linear abhängige v und w offensichtlich Gleichheit in (iii).

(iv)

$$\begin{aligned} \| v + w \|^2 &= \langle v + w, v + w\rangle = \| v \|^2 + 2\langle v, w\rangle + \| v \|^2 \\ &\overset{\text{(iii)}}{\leq} \| v \|^2 + 2 \| v \| \| w \| + \| w \|^2 = (\| v \| + \| w \|)^2. \end{aligned}$$

(6.4) Def. Sei $V, \langle,\rangle$ euklidischer Vektorraum und $v, w \in V \setminus \{0\}$. Dann ist der Winkel $\varphi = \varphi(v, w) \in [0, \pi]$ zwischen v und w definiert durch:

$$\cos\varphi = \frac{\langle v, w\rangle}{\| v \| \| w \|} \quad \text{bzw.} \quad \varphi = \arccos \frac{\langle v, w\rangle}{\| v \| \| w \|} \in [0, \pi]$$

Bem. 1) (5.3)(iii) $\Rightarrow -1 \leq \frac{\langle v,w\rangle}{\|v\|\|w\|} \leq 1$. $\cos : [0,\pi] \to [-1,1]$ ist bijektiv mit Umkehrabbildung arccos.

2) $\varphi(v,w) = \varphi(w,v)$
3) v und w heißen <u>orthogonal</u> (senkrecht) zueinander $\Leftrightarrow \langle v,w\rangle = 0$
4) Die Definition von φ ist so eingerichtet, daß der "Kosinussatz" gilt:

$$\| v - w \|^2 = \| v \|^2 - 2\langle v,w\rangle + \| w \|^2 = \| v \|^2 + \| w \|^2 - 2\cos\varphi \, \| v \| \| w \| \, .$$

(6.5) Def.: Eine Teilmenge $S \subseteq V$ heißt Orthonormalsystem (ONS), falls gilt

(i) Für alle $v \in S$ gilt $\| v \| = 1$ ("*alle $v \in S$ normiert*") und
(ii) $v, w \in S$, $v \neq w \Rightarrow \langle v,w\rangle = 0$ ("*je zwei orthogonal*")
Eine Orthonormalbasis (ONB) ist ein Orthonormalsystem, das V erzeugt.

Bem.: S ONS $\Rightarrow$ S linear unabhängig: $v_1,\dots,v_n \in S$ verschieden, $r_1,\dots,r_n \in \mathbb{R}$ und $\sum\limits_{i=1}^{n} r_i v_i = 0 \Rightarrow 0 = \langle \sum\limits_{i=1}^{n} r_i v_i, \sum\limits_{j=1}^{n} r_j v_j\rangle = \sum\limits_{i,j=1}^{n} r_i r_j \langle v_i, v_j\rangle = \sum\limits_{i=1}^{n} r_i^2 \Rightarrow$ alle $r_i = 0$.

Bsp.:

1) $V = \mathbb{R}^n$, $\langle x,y\rangle := \sum\limits_{i=1}^{n} x_i y_i \Rightarrow \{e_1,\dots,e_n\}$ ist ONB von $(\mathbb{R}^n, \langle,\rangle)$.
2) $V = C^0([0,2\pi],\mathbb{R})$, $\langle,\rangle = \langle,\rangle_{L^2}$.
Sei $f_0 \in V$ definiert durch $f_0(x) := \frac{1}{\sqrt{2\pi}}$ und für $k \geq 1$:

$$f_k(x) := \frac{1}{\sqrt{\pi}}\cos(kx), g_k(x) := \frac{1}{\sqrt{\pi}}\sin(kx).$$

Dann ist $S = \{f_k \mid k \in \mathbb{N}\} \cup \{g_k \mid k \in \mathbb{N} \setminus \{0\}\}$ ein ONS.

(6.6) Satz:

(i) *(Besselsche Ungleichung). Ist S ONS, $v \in V$ und $E \subseteq S$ endlich, so gilt*

$$\sum_{e\in E} \langle v,e\rangle^2 \leq \| v \|^2 \, .$$

(ii) *Ist $B = (v_1,\dots,v_n)$ ONB von V, so gilt*

$$v = \sum_{i=1}^{n} \langle v, v_i\rangle v_i.$$

Bem.: Im Fall von Bsp. 2) heißen $a_k := \langle f, f_k\rangle_{L^2}$ für $k \geq 0$ (d.h. $a_k = \frac{1}{\sqrt{\pi}} \int\limits_0^{2\pi} f(x)\cos(kx)dx$ für $k \geq 1$ und $a_0 = \frac{1}{\sqrt{2\pi}} \int\limits_0^{2\pi} f(x)dx$) und $b_k := \langle f, g_k\rangle_{L^2} = \frac{1}{\sqrt{\pi}} \int\limits_0^{2\pi} f(x)\sin(kx)dx$ für $k \geq 1$ die

Fourierkoeffizienten von $f \in V = C^0([0, 2\pi], \mathbb{R})$. Aus (6.6)(i) folgt:

$$a_0^2 + \sum_{k=1}^{\infty}(a_k^2 + b_k^2) \leq \| f \|_{L^2}^2 = \int_0^{2\pi} f^2(x)dx.$$

In Wahrheit ist diese Ungleichung stets eine Gleichheit (man sagt zu dieser Eigenschaft, daß dieses S ein vollständiges ONS ist), aber das können wir hier nicht beweisen.

Bew.: (i) $0 \leq \| v - \sum_{e \in E} \langle v, e \rangle e \| = \| v \|^2 - 2 \sum_{e \in E} \langle v, e \rangle^2 + \sum_{e \in E} \langle v, e \rangle^2$

(ii) $(v_1, \ldots, v_n)$ Basis $\Rightarrow \exists r_1, \ldots, r_n \in \mathbb{R} : v = \sum_{j=1}^{n} r_j v_j \Rightarrow \langle v, v_i \rangle = \sum_{j=1}^{n} r_j \langle v_j, v_i \rangle = r_i$

(6.7) Satz (Gram-Schmidtsches Orthonormalisierungsverfahren). *Seien $v_1, \ldots, v_k \in V$ linear unabhängig. Dann existiert ein* ONS $\tilde{v}_1, \ldots, \tilde{v}_k \in V$ *mit* $\mathrm{span}\{v_1, \ldots, v_i\} = \mathrm{span}\{\tilde{v}_1, \ldots, \tilde{v}_i\}$ *für alle $1 \leq i \leq k$. Insbesondere: Ist V endlichdimensional, so existiert eine* ONB *von V.*

Bew.: Induktion nach k:

$$k = 1 : \tilde{v}_1 = \frac{1}{\| v_1 \|} v_1 \quad (v_1 \neq 0!)$$

Induktionsschritt: Nach Induktionsvoraussetzung existiert ein ONS $\tilde{v}_1, \ldots, \tilde{v}_{k-1}$ mit

$$\mathrm{span}\{v_1, \ldots, v_i\} = \mathrm{span}\{\tilde{v}_1, \ldots, \tilde{v}_i\} \text{ für alle } 1 \leq i \leq k-1.$$

Setze $w_k = v_k - \sum_{j=1}^{k-1} \langle v_k, \tilde{v}_j \rangle \tilde{v}_j \quad (\Rightarrow \langle w_k, \tilde{v}_i \rangle = 0$ für $1 \leq i \leq k-1)$ und $\tilde{v}_k := \frac{1}{\|w_k\|} w_k$. Dann ist $\tilde{v}_1, \ldots, \tilde{v}_k$ ONS und $\mathrm{span}\{\tilde{v}_1, \ldots, \tilde{v}_k\} = \mathrm{span}\{v_1, \ldots, v_k\}$.

(6.8) Def. Seien $V, \langle , \rangle$ und $\tilde{V}, \tilde{\langle , \rangle}$ euklidische Vektorräume. Ein $L \in \mathrm{Hom}(V, \tilde{V})$ heißt orthogonal(bzgl. $\langle , \rangle$ und $\tilde{\langle , \rangle}$), falls für alle $v, w \in V$ gilt: $\langle v, w \rangle = \tilde{\langle L(v), L(w) \rangle}$. Bem.:

1) L orthogonal $\Rightarrow L$ injektiv $(L(v) = 0 \Rightarrow \| L(v) \|^2 = \| v \|^2 = 0 \Rightarrow v = 0)$
   Speziell: Gilt $\dim V = \dim \tilde{V} < \infty$, so ist jede orthogonale Abbildung bijektiv.
2) Sei $\tilde{V} = V, \dim V < \infty, \tilde{\langle , \rangle} = \langle , \rangle$. Dann bilden die orthogonalen $L \in \mathrm{End}(V)$ eine Untergruppe von $(\mathrm{Aut}(V), \circ)$, die orthogonale Grupe $O(V)$ von V.

(6.9) Folgerung. Ist $V, \langle , \rangle_V$ euklidischer Vektorraum, $\dim V = n$, so existiert eine (bijektive) orthogonale Abbildung $L : (V, \langle , \rangle_V) \to (\mathbb{R}^n, \langle , \rangle)$.

Bew.: Sei $f_1, \ldots, f_n$ ONB von V. Definiere L durch $L(f_i) = e_i \in \mathbb{R}^n$.

$$\begin{array}{rcllcl} \text{Ist } v & = & \sum r_i f_i \text{ und } w = \sum s_j f_s & \text{so folgt} \quad \langle v, w \rangle_V & = & \sum r_i s_i. \\ \text{Daraus folgt } L(v) & = & \sum r_i e_i, L(w) = \sum s_j e_j & \text{und} \quad \langle L(v), L(w) \rangle & = & \sum r_i s_i = \langle v, w \rangle_V. \end{array}$$

(6.10) Satz. *Sei $L \in \mathrm{End}(V)$ und $\mathcal{G} = (e_1, \ldots, e_n)$* ONB *von V und $A = \mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L)$.*

$$\textit{Dann gilt: } L \in O(V) \Leftrightarrow A^T A = E_n \quad (\Leftrightarrow A^T = A^{-1})$$

Bew.: $\delta_{ij} = \langle e_i, e_j \rangle = \langle L(e_i), L(e_j) \rangle = \left\langle \sum_{k=1}^{n} a_{ki} e_k, \sum_{l=1}^{n} a_{lj} e_l \right\rangle$

$= \sum_{k,l=1}^{n} a_{ki} a_{lj} \delta_{kl} = \sum_{k=1}^{n} a_{ki} a_{kj}$

$\Leftrightarrow A^T A = E_n.$

Die Hauptachsentransformation

Im folgenden sei V ein n-dimensionaler $\mathbb{R}$-Vektorraum mit Skalarprodukt $\langle , \rangle$ und $b : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform. Ist $(e_1, \dots, e_n)$ eine Basis von V und sind $x = \sum_{i=1}^{n} x_i e_i$, $y = \sum_{j=1}^{n} y_j e_j$ beliebige Vektoren in V, so folgt aus der Bilinearität von b:

$$b(x, y) = b\left(\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j e_j\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j \, b(e_i, e_j).$$

Man nennt die Matrix $B = (b_{ij}) \in \mathbb{R}^{n \times n}$ mit $b_{ij} = b(e_i, e_j)$ die <u>Matrix</u> <u>von</u> b bzgl. der Basis $(e_1, \dots, e_n)$. Ist b symmetrisch (wie wir vorausgesetzt haben), so gilt $b_{ij} = b(e_i, e_j) = b(e_j, e_i) = b_{ji}$, d.h. $B = B^\top$ (solche Matrizen heißen symmetrisch). Es gilt dann für $x = \sum_{i=1}^{n} x_i e_i$, $y = \sum_{j=1}^{n} y_j e_j$:

$$b(x, y) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i b_{ij} y_j = \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} B \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

wobei der letzte Term als Produkt von Matrizen zu interpretieren ist.

Besonders einfach zu verstehen ist eine Bilinearform b, deren Matrix (bzgl. $e_1, \dots, e_n$) eine Diagonalmatrix $b_{ij} = \lambda_i \delta_{ij}$ ist. Dann gilt:

$$b(x, y) = \sum_{i=1}^{n} x_i \lambda_i y_i.$$

Zum Beispiel ist ein soches b genau dann positiv definit, wenn alle λ_i (für $i = 1, \dots, n$) positiv sind.

(6.11) Satz (Hauptachsentransformation). *Zu jeder symmetrischen Bilinearform $b : V \times V \to \mathbb{R}$ existiert eine* ONB $f_1, \dots, f_n$ *von V und $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, so daß für alle $x = \sum_{i=1}^{n} x_i f_i$, $y = \sum_{j=1}^{n} y_j f_j$ in V gilt*

$$b(x, y) = \sum_{i=1}^{n} x_i \lambda_i y_i.$$

Anders ausgedrückt: Es existiert stets eine ONB *von* V*, bezüglich derer die Matrix von* b *eine Diagonalmatrix ist. (Die* λ_i *sind dann gerade die Diagonalelemente der Matrix.)*

Herkunft der Bezeichnung "Hauptachsentransformation": Ist $b : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ positiv definite, symmetrische Bilinearform, so ist die Menge

$$E = \{x \in \mathbb{R}^2 \mid b(x,x) = 1\}$$

eine Ellipse. Satz (6.11) besagt, daß es eine ONB f_1, f_2 (bzgl. des üblichen Skalarprodukts) gibt, so daß

$$E = \{x = x_1 f_1 + x_2 f_2 \in \mathbb{R}^2 \mid \lambda_1 x_1^2 + \lambda_2 x_2^2 = 1\}$$

gilt, d.h. die Geraden durch $0 \in \mathbb{R}^2$ in Richtung f_1 bzw. f_2 sind die "Hauptachsen" der Ellipse E mit zugehörigen Achsenabschnitten der Länge $(\lambda_1)^{-\frac{1}{2}}$ bzw. $(\lambda_2)^{-\frac{1}{2}}$.

Satz (6.11) spielt in folgendem Zusammenhang in der Analysis II eine Rolle ("Kurvendiskussion für Funktionen $f : \mathbb{R}^n \to \mathbb{R}$"): Ist $f : \mathbb{R}^n \to \mathbb{R}$ zweimal stetig differenzierbar, $x_0 \in \mathbb{R}^n$ und $h \in \mathbb{R}^n$, so gilt (siehe Analysis II)

$$f(x_0 + h) = f(x_0) + \mathrm{Df}(x_0)(h) + \frac{1}{2}\mathrm{D}^2\mathrm{f}(x_0)(h,h) + R_{f,x_0}(h),$$

wobei $\mathrm{Df}(x_0) \;:\; \mathbb{R}^n \to \mathbb{R}$ linear (1. Ableitung von f an der Stelle x_0),
$\mathrm{D}^2\mathrm{f}(x_0) \;:\; \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ symmetrisch, bilinear (2. Ableitung von f an der Stelle x_0),

und $R_{f,x_0}(h)$ eine Funktion mit $\lim\limits_{h\to 0} \frac{R_{f,x_0}(h)}{\|h\|^2} = 0$ ist. x_0 heißt <u>kritischer</u> <u>Punkt</u> von f, falls $\mathrm{Df}(x_0) = 0$ gilt. Aus (6.11) folgt, daß man eine ONB des $\mathbb{R}^n$ finden kann, bzgl. derer $\mathrm{D}^2\mathrm{f}(x_0)(h,h) = \sum\limits_{i=1}^{n} \lambda_i h_i^2$ gilt. Daraus folgt für einen kritischen Punkt x_0 von f:

alle $\lambda_i > 0$ ($\Leftrightarrow \mathrm{D}^2\mathrm{f}(x_0)$ positiv definiert) $\Rightarrow$ x_0 lokales Minimum.
alle $\lambda_i < 0$ ($\Leftrightarrow \mathrm{D}^2\mathrm{f}(x_0)$ negativ definiert) $\Rightarrow$ x_0 lokales Maximum.

Im Gegensatz zum (aus der Schule bekannten) Fall $n = 1$ gibt es für $n > 1$ noch andere wichtige Möglichkeiten für $D^2 f(x_0)$: Es kann z.B. für $n = 2$ $\lambda_1 > 0$ und $\lambda_2 < 0$ gelten. Dann sieht der Graph von f in einer Umgebung von x_0 wie eine "Sattelfläche" aus.

Vorbereitung zum Beweis von Satz (6.11):

Wegen (6.9) können wir annehmen, daß $V = \mathbb{R}^n$, $\langle x, y\rangle = \sum\limits_{i=1}^{n} x_i y_i$ gilt. Ist nämlich $L : \mathbb{R}^n \to V$ orthogonal, so betrachte $\tilde{b} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $\tilde{b}(x,y) := b(L(x), L(y))$. Haben wir die Behauptung für $\tilde{b}$ bewiesen, so folgt sie leicht auch für b. (Als gesuchte ONB von V kann gerade das Bild unter L einer ONB des $\mathbb{R}^n$ genommen werden, die $\tilde{b}$ diagonalisiert.)

Unter leichtem Missbrauch der Bezeichnung verwenden wir B auch als Symbol für die lineare Abbildung $B : \mathbb{R}^n \to \mathbb{R}^n$ mit $B(e_i) = \sum\limits_{j=1}^{n} b_{ij} e_j = \sum\limits_{j=1}^{n} b_{ji} e_j$. Dann gilt:

(6.12) Lemma. $b(x,y) = \langle B(x), y\rangle = \langle x, B(y)\rangle$ für alle $x, y \in \mathbb{R}^n$.

Bem.: Ein $B \in \mathrm{End}(\mathbb{R}^n)$ mit $\langle B(x), y\rangle = \langle x, B(y)\rangle$ für alle $x, y \in \mathbb{R}^n$ heißt "selbstadjungiert".

$$\begin{aligned}
\text{Bew.: } \langle B(x), y\rangle &= \langle B\left(\sum_{i=1}^n x_i e_i\right), \sum_{k=1}^n y_k e_k\rangle \\
&= \sum_{i=1}^n \sum_{k=1}^n x_i y_k \langle B(e_i), e_k\rangle \\
&= \sum_{i=1}^n \sum_{k=1}^n \sum_{j=1}^n x_i y_k b_{ij} \underbrace{\langle e_j, e_k\rangle}_{=\delta_{jk}} = \sum_{i=1}^n \sum_{j=1}^n x_i y_j b_{ij} \\
&= b(x, y)
\end{aligned}$$

Da b symmetrisch ist, folgt daraus:

$$\langle x, B(y)\rangle = \langle B(y), x\rangle = b(y, x) = b(x, y) = \langle B(x), y\rangle.$$

Satz (6.11) ist also äquivalent dazu, daß es für jedes selbstadjungierte $B \in \mathrm{End}(\mathbb{R}^n)$ eine ONB $f_1, \ldots, f_n$ des $\mathbb{R}^n$ aus Eigenvektoren von B gibt. Sind dann $\lambda_1, \ldots, \lambda_n$ die zugehörigen Eigenwerte von B, d.h. $B(f_i) = \lambda_i f_i$ für $i = 1, \ldots, n$, so gilt:

$$b_{ij} = b(f_i, f_j) = \langle B(f_i), f_j\rangle = \lambda_i \langle f_i, f_j\rangle = \lambda_i \delta_{ij}, \text{ wie in (6.11) behauptet.}$$

Im ersten (schwierigeren) Schritt zeigen wir:

1. Schritt: B besitzt einen Eigenwert $\lambda \in \mathbb{R}$.
Sei $\tilde{B} \in \mathrm{End}(\mathbb{C}^n)$ die $\mathbb{C}$-lineare Fortsetzung von $B \in \mathrm{End}(\mathbb{R}^n)$ auf $\mathbb{C}^n$, d.h. $\tilde{B}(z_1, \ldots, z_n) = \left(\sum_{i=1}^n b_{1i} z_i, \sum_{i=1}^n b_{2i} z_i, \ldots, \sum_{i=1}^n b_{ni} z_i\right)$. Wir verwenden nun (ohne das bewiesen zu haben), daß $\mathbb{C}$ algebraisch abgeschlossen ist. Daraus folgt, daß es einen Eigenwert $\lambda \in \mathbb{C}$ von $\tilde{B}$ gibt (vgl. (5.24) und die daran anschließenden Bemerkungen).

Wir wollen zeigen, daß in Wahrheit $\lambda \in \mathbb{R}$ gilt. Sei $v = x + iy \in \mathbb{C}^n \setminus \{0\}$ mit $x, y \in \mathbb{R}^n$ ein EV zum EW $\lambda = \alpha + i\beta$, $\alpha, \beta \in \mathbb{R}$, von $\tilde{B}$. Dann gilt

$$\tilde{B}(x + iy) = B(x) + iB(y) \text{ mit } B(x), B(y) \in \mathbb{R}^n$$

und andererseits wegen $\tilde{B}(v) = \lambda v$:

$$\tilde{B}(x + iy) = (\alpha + i\beta)(x + iy) = (\alpha x - \beta y) + i(\beta x + \alpha y).$$

Also:

$$(*) \qquad B(x) = \alpha x - \beta y \text{ und } B(y) = \beta x + \alpha y$$

Aus $(*)$ folgt mit (6.12):

$$\langle B(x), y\rangle = \alpha\langle x, y\rangle - \beta \| y \|^2 = \langle x, B(y)\rangle = \beta \| x \|^2 + \alpha\langle x, y\rangle.$$

Also $0 = \beta(\| x \|^2 + \| y \|^2)$, woraus wegen $\| x \|^2 + \| y \|^2 > 0$ (da $v \neq 0$!) $\beta = 0$ folgt, d.h. $\lambda = \alpha \in \mathbb{R}$ und, nach $(*)$,

$$B(x) = \lambda x, B(y) = \lambda y.$$

Da $v = x + iy \neq 0$ gilt, sind nicht beide x und y der Nullvektor, d.h. mindestens einer von beiden ist ein Eigenvektor von B (zum EW $\lambda = \alpha \in \mathbb{R}$).

2. Schritt: Induktion über $n = \dim \mathbb{R}^n$.
$n = 1$: Wähle $f_1 := e_1 \in \mathbb{R}^1$. Dann gilt: $b(x_1 e_1, x_1 e_1) = x_1^2 \underbrace{b(e_1, e_1)}_{=:\lambda_1}$.

$n > 1$: Wähle als $f_1 \in \mathbb{R}^n$ einen (nach Schritt 1 existierenden) Eigenvektor von B, der ohne Einschränkung als normiert angenommen werden kann, d.h. $\| f_1 \| = 1$. Betrachte

$$(**) \qquad U := \{x \in \mathbb{R}^n \mid \langle x, f_1 \rangle = 0\}.$$

Dann ist U ein $(n-1)$-dimensionaler Untervektorraum des $\mathbb{R}^n$ (vgl. (3.24)) und die Restriktion $\langle , \rangle \mid U \times U$ von $\langle , \rangle$ auf U ist ein Skalarprodukt auf U. Entscheidend ist nun, daß $B(U) \subseteq U$ gilt: Ist $x \in U$, d.h. $x \in \mathbb{R}^n$ und $\langle x, f_1 \rangle = 0$, so folgt $B(x) \in \mathbb{R}^n$ und

$$\langle B(x), f_1 \rangle \overset{(6.12)}{=} \langle x, B(f_1) \rangle = \langle x, \lambda_1 f_1 \rangle = \lambda_1 \langle x, f_1 \rangle = 0,$$

d.h. $B(x) \in U$. $B \mid U$ ist also ein selbstadjungierter Endomorphismus von U, so daß nach der Induktionsvoraussetzung die Existenz einer ONB $f_2, \ldots, f_n$ von U aus Eigenvektoren von $B \mid U$ ($\Rightarrow$ von B) existiert. Wegen $(**)$ ist dann $f_1, \ldots, f_n$ eine ONB des $\mathbb{R}^n$ aus Eigenvektoren von B.